Optimal Real-Time Bidding for Display Advertising

Weinan Zhang

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

of

University College London.

Department of Computer Science

University College London

June 6, 2016

I, Weinan Zhang, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Real-Time Bidding (RTB) is revolutionising display advertising by facilitating a real-time auction for each ad impression. As they are able to use impression-level data, such as user cookies and context information, advertisers can adaptively bid for each ad impression. Therefore, it is important that an advertiser designs an effective bidding strategy which can be abstracted as a function - mapping from the information of a specific ad impression to the bid price. Exactly how this bidding function should be designed is a non-trivial problem. It is a problem which involves multiple factors, such as the campaign-specific key performance indicator (KPI), the campaign lifetime auction volume and the budget.

This thesis is focused on the design of automatic solutions to this problem of creating optimised bidding strategies for RTB auctions: strategies which are optimal, that is, from the perspective of an advertiser agent - to maximise the campaign's KPI in relation to the constraints of the auction volume and the budget. The problem is mathematically formulated as a functional optimisation framework where the optimal bidding function can be derived without any functional form restriction. Beyond single-campaign bid optimisation, the proposed framework can be extended to multi-campaign cases, where a portfolio-optimisation solution of auction volume reallocation is performed to maximise the overall profit with a controlled risk. On the model learning side, an unbiased learning scheme is proposed to address the data bias problem resulting from the ad auction selection, where we derive a "bid-aware" gradient descent algorithm to train unbiased models. Moreover, the robustness of achieving the expected KPIs in a dynamic RTB market is solved with a feedback control mechanism for bid adjustment.

To support the theoretic derivations, extensive experiments are carried out based on large-scale real-world data. The proposed solutions have been deployed in three commercial RTB systems in China and the United States. The online A/B tests have

demonstrated substantial improvement of the proposed solutions over strong baselines.

Acknowledgements

First, I would like to express my sincere gratitude to my first supervisor Prof. Jun Wang. I got to know Jun in Mar. 2012 when I was reading one of his papers on personalised recommendation techniques published in RecSys 2010. After a careful survey about his research on computational advertising and recommender systems, I decided to only apply his research group in UCL for pursuing my Ph.D. Fortunately, I got the offer and the scholarship from Jun and started my Ph.D. career from Sep. 2012.

During these 3.5 years of Ph.D. research work with Jun, I have been always feeling lucky that I have made the right choice. Jun is a nice and easy-going professor, just like a senior friend of mine. I totally have no pressure but only enjoyment when discussing research problems with him. Jun usually emphasises the novelty of a research topic, including the research problem novelty and methodology contributions. He dedicates quite a lot of time to discuss with Ph.D. students and helps revise their papers very carefully. In addition, he encourages me to pursue research internships in large companies like Google and Microsoft Research as well as technology startups like MediaGamma. Without his invaluable and generous support, I do not believe I could complete 4 internships and publish 16 research papers during 3.5 years.

Luckily, I have got the offer and am going to join Shanghai Jiao Tong University as an assistant professor in Aug. 2016, working on big data mining and artificial intelligence research. I believe Jun's philosophy of being a good Ph.D. supervisor will deeply help me in my future academic career.

Second, I want to show my great respect and appreciation to my second supervisor Prof. Stephen Robertson. When he agreed to be my second supervisor, Steve was about to retire and planned to fully change the shape of his life. I sincerely thank Steve for travelling to London to provide me both high-level research advices and specific solution tips during our meetings and my first-year and transfer vivas.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

"Half the money I spend on advertising is wasted; the trouble is I don't know which half."

— John Wanamaker (11 July, 1838 - 12 December, 1922)

This popular quotation from John Wanamaker, a pioneer of advertising and department stores, illustrates how difficult it was, a hundred years ago, to quantify customer response and so an advertising campaign's potential performance. Over the last twenty years, advancements related to the world wide web have fundamentally changed this situation. The web provides not only new, efficient, ways to connect customers to retailers but also effective feedback mechanisms whereby customer response can be measured. Some of these mechanisms are as follows: observing users' search queries [1], web browsing patterns [2], clicks [3] and conversions [4] etc. With the techniques of user cookie matching and online behaviour tracking, a user's feedback on a single ad impression can be quantitatively evaluated. Based on such quantitative feedback, advertisers are able to refine their campaigns, e.g., by refining the targeting rules and reallocating the budget, to improve their ad campaigns' performance. Therefore, Internet advertising has become more and more efficient and thus one of the most important forms of advertising.

1.1 Background of Internet Advertising Market

According to marketing investigation reports from the Interactive Advertising Bureau (IAB) and eMarketer.com, in 2014, the total Internet advertising spend in the United States reached $49.5 billion, a 15.6% increase from the same figure for 2013. This
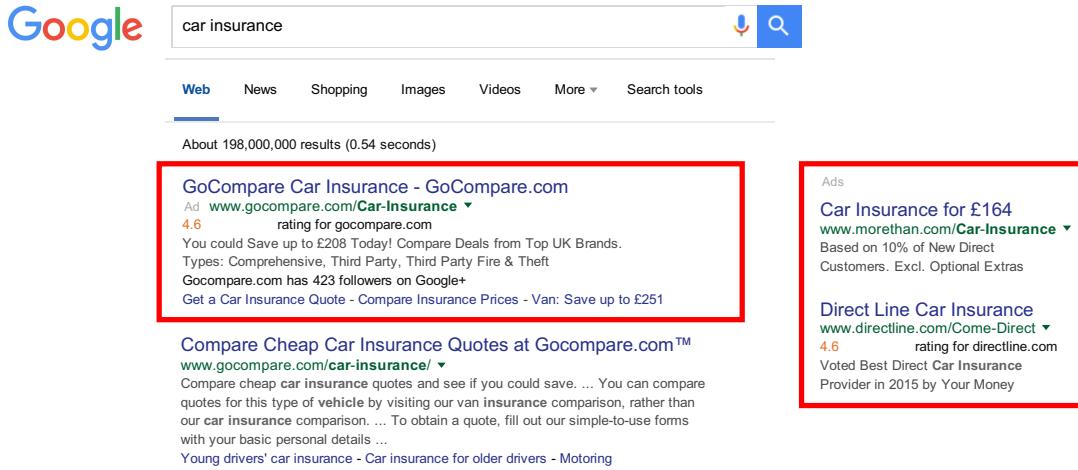
Figure 1.1: An example snapshot of sponsored search: after a user submits a query "car insurance" to the search engine, relevant text ads will be retrieved and shown in the search engine result page along with the retrieved web pages. The match between the ads and the user is mainly based on the relevance of the advertisers' bid keywords to the user's query keywords. Source: Google.

amount is 75.0% of the television advertising spend [5]. The total Internet advertising spend in the United Kingdom, in 2014, reached £7.0 billion, a 13.0% increase from 2013, and it was estimated to have become more than half of the total UK ad spend in 2015 [6]. It is also reported that the global Internet advertising spend will jump 18.0%, in 2015, to reach $170.2 billion: i.e., 29.9% of the total advertising market [7].

There are different ways to categorise the various Internet advertising formats. From the point of view of mechanisms, there are two major categories: sponsored search, accounting for 50.0% of the total Internet advertising spend in the United States in 2014, and display advertising, accounting for 35.2%. The remaining 14.8% was spent on other ad formats such as classifieds [8].

Sponsored search refers to the text ads placed on the top and at the right-hand side of the search engine result pages [9]. An example is shown in Figure 1.1. Sponsored search connects advertisers to users via keywords, i.e., the advertisers bid on a number of keywords and the users search for certain keywords according to their interests at the time. The search engine tries to make the best match it can between the user's search keywords and the advertisers' bid keywords via an auction mechanism [10]. The best matched ads will be ranked according to a score calculated from the ad's relevance and the bid price. The pricing scheme for sponsored search is cost per click (CPC), i.e., the advertiser has to pay the search engine provider only when her ad has been clicked on
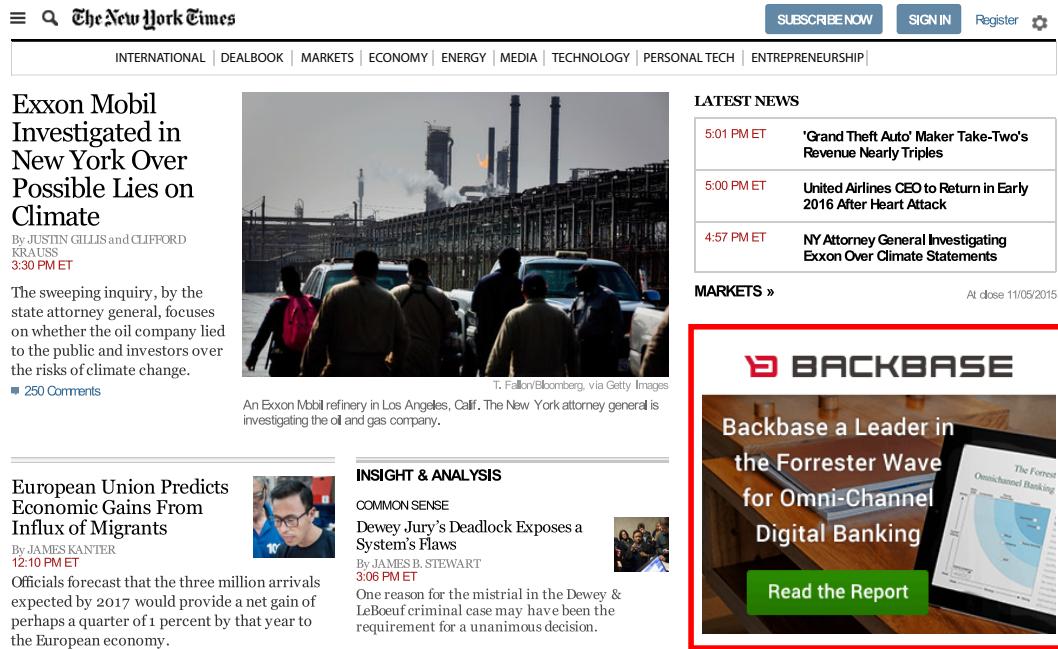
Figure 1.2: An example snapshot of display advertising: when a user browses a web page, some image-based or multimedia ads will be shown along with the web page content. The match between the ads and the users is mainly based on the correspondence between the advertisers' targeting rules and the user's segmentations, such as demographics and interest tags. Source: NYTimes.

by a user.

Display advertising plays in more general scenarios than does sponsored search. The image-based or multimedia ads are displayed to users when they use any online media: web pages, mobile apps and videos, i.e. not just when they are performing a search [11], shown in Figure 1.2 as an example. Without the search keywords to explicitly indicate the user's interest, display advertising tries to match the advertisers and the users via users' demographical and interest segmentations [12]. This represents a more challenging problem than that related to sponsored search. With the rapid growth of mobile applications, display advertising (mobile and non-mobile) has achieved very fast growth in the past five years (from $9.9 billion in 2010 to $17.4 billion in 2014, United States) [13, 5]. Display advertising is set to create much larger business value in the near future.

Before the emergence of real-time bidding techniques in 2009 - these will be discussed in Section 1.2 - the main trading mechanisms for display advertising were negotiation-based contracts and ad network volume aggregations [14]. In this scenario, publishers sell the website volume in batches, either to the premium advertisers di-

rectly or to intermediate ad agencies (ad networks), while advertisers buy the media volume in batches from publishers directly or through ad agencies. There are several obvious inefficiencies involved with such a trading mechanism: (i) There is usually an imbalance between demand side (advertisers) and supply side (publishers) in relation to a particular (often local) market, e.g., a publisher and several advertisers, or an ad agency linked with some advertisers and publishers. Usually, the advertisers cannot buy the expected amount of the targeted volume while the publishers cannot sell all the volume. (ii) As the contract is based on aggregated volume, the price for each ad impression, in the contract, is uniform despite the fact that there is always large utility difference among the ad impressions in a contract volume [15, 4].

1.2 Real-Time Bidding based Display Advertising

Since 2009, Real-Time Bidding (RTB) based display advertising has emerged and has become the new frontier for Internet advertising [16, 17]. Unlike conventional sponsored search or contextual advertising, where an advertiser pre-sets a bid price for each selected keyword for her campaign, RTB allows an advertiser to use computer algorithms to submit a bid for each impression within a very short time frame, often less than 100ms [14]. These bids will be based on the impression-level features, such as user cookie and context information. A real-time auction is hosted by an intermediary, called the ad exchange [16], which will select the ad with the highest bid for display to the user. RTB has fundamentally changed the landscape of Internet advertising and solved the above mentioned problems of conventional display advertising because (i) allowing per-impression transactions scales the buying process across a large number of available ad inventories including the leftovers; (ii) the real-time audience data encourages behavioural targeting and makes a significant shift towards buying that is focused on user data [18] rather than contextual data. With its fine-grained user targeting and auction mechanism, RTB has significantly improved the campaign return-on-investment (ROI) and become an essential Internet advertising paradigm.

Demand-Side Platforms (DSPs) are thus created to help advertisers manage their campaigns and optimise their real-time bidding activities. The interaction process among the DSP and the other main components of the RTB eco-system is summarised into the following steps in Figure 1.3: (0) when a user visits an ad-supported site (e.g.,
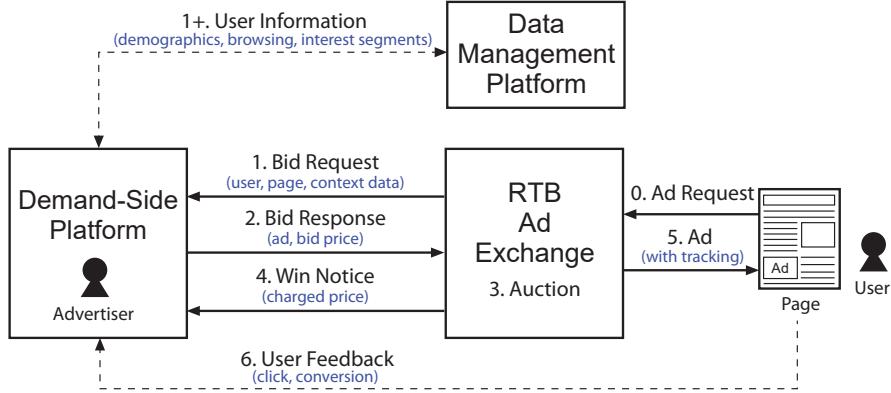
Figure 1.3: A brief illustration of the interactions between user, ad exchange and the advertiser's DSP bidding agent. The solid-line arrows represent real-time interactions, while the dashed-line arrows represent the interactions that are not necessarily in real-time. The step (1+) is optional, i.e., the DSP may use their own user segmentation information instead of calling the third-party DMP for this user information. Besides, this figure is mainly from a DSP's perspective, thus some components, such as supply-side platform (SSP) and ad server are omitted for simplicity.

web pages, streaming videos and mobile apps), each ad placement will trigger a call for an ad (ad request) to the ad exchange. (1) The ad exchange sends the bid requests, for this particular ad impression, to each advertiser's DSP bidding agent, along with the other available information such as the user cookie and context information. (2) With the information of the bid request and each of its qualified ads,[1] the bidding agent calculates a bid price. Then the bid response (a pair consisting of the ad and the bid price) is sent back to the exchange to take part in the auction. (3) Having received the bid responses from the advertisers within a predefined time window, the ad exchange hosts an auction and picks the ad with the highest bid as the auction winner. (4) Then the winner is notified of this result and the price which will be charged by the ad exchange. (5) Finally, the winner's ad will be shown to the visitor along with the regular content of the publisher's site. It is commonly known that a long page-loading time will greatly reduce the user's' experience, in terms of quality [16]. Thus, advertiser bidding agents are usually required to return a bid in a very short time frame (e.g., 100ms). (6) The user's feedback (e.g., click and conversion) on the displayed ad is tracked and finally sent back to the winning advertiser. A more detailed discussion concerning the whole RTB eco-system is given in [14].

From the above interaction processes, it is obvious that, for DSPs, the most signif-

---

[1]The qualification of each ad for the bid request is based on ad size matching, the campaign's target rules, etc.
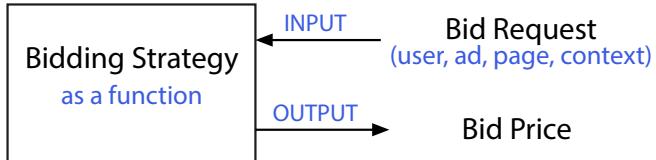
Figure 1.4: A bidding strategy can be abstracted as a function mapping from the given bid request (in a high dimensional feature space) to a bid price (a non-negative real or integer number).

icant module is the bidding agent which enables advertisers to perform per-impression bidding in real-time via computer algorithms [19]. The bidding strategy used here can be abstracted as a function which takes in the information of a specific ad display opportunity, i.e., the bid request, and outputs the corresponding bid price for each qualified ad, as illustrated in Figure 1.4. How this bidding function should be designed involves multiple factors, including the auction type [20], campaign's lifetime, ad auction volume, the campaign budget and the campaign-specific key performance indicator (KPI) - such as ad click number or advertising revenue. Calculating all these factors is non-trivial. To the best of our knowledge, most studies on bidding strategies are still restricted to truthful bidding for second price auctions [21, 4] and there is no previous work which describes a system which directly learns a bidding strategy in order to optimise campaign performance in RTB display advertising.

1.3   Research Problems and Contributions

In this thesis, we focus on designing automatic algorithms which optimise the bidding strategies for impression-level auctions, from the perspective of an advertiser or a DSP, to improve the supported single or multiple campaigns' advertising performance within the auction volume and budget constraints. Advertising performance is measured by specific KPIs, such as the number of campaign clicks, effective cost per click (eCPC) and profit etc. In the model learning stage, the data bias problem caused by the ad auction selection is carefully handled in order to facilitate unbiased model learning and optimisation. Moreover, we explicitly consider the robustness of achieving the expected KPIs [21] in relation to dynamic RTB market competition by embedding a feedback control mechanism in the bidding agent. The overview of the research problems and their connections are shown in Figure 1.5.
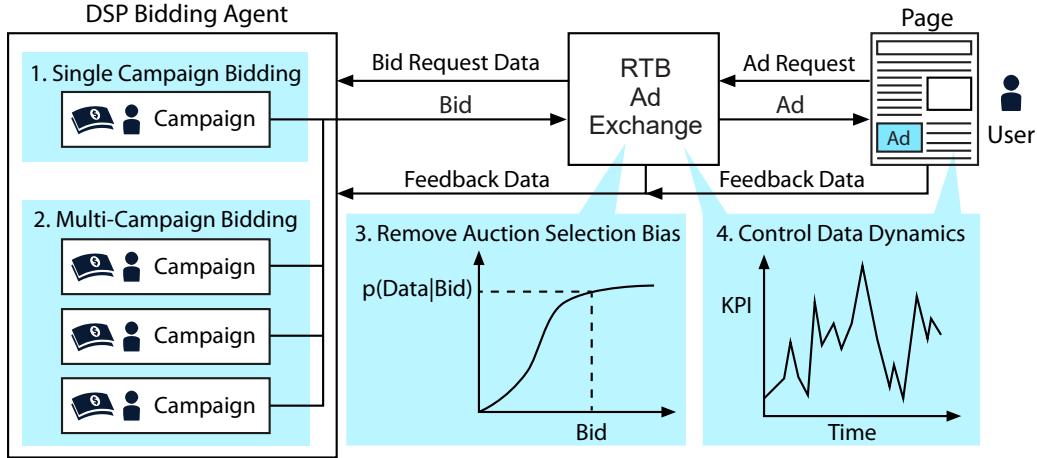
Figure 1.5: An overview of the research problems studied in this thesis and their links: (1) single-campaign bid optimisation; (2) multi-campaign bid optimisation; (3) handling auction selection bias for unbiased model learning; (4) feedback control mechanisms for dealing with data dynamics.
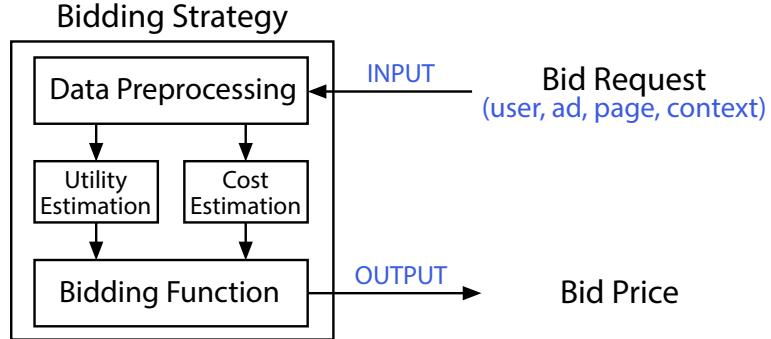
Figure 1.6: The proposed quantitative bidding: the logic of the bidding function only depends on two (sets of) factors, i.e., the estimated utility and the cost of the ad display opportunity.

1.3.1 Single-Campaign Bidding

The fundamental research problem addressed by this thesis is bidding strategy optimisation for a single campaign, given its lifetime auction volume and budget. In such a performance-driven advertising scenario, each ad display opportunity, i.e., bid request, is *quantified* where its *utility*, e.g., the probability of a user clicking on the displayed ad [22] or the expected revenue from this ad impression [23, 4], and *cost*, e.g., the cost of winning this ad impression in the auction [24], are carefully estimated. Based on the estimated utility and cost of each bid request, the concept of quantitative bidding is proposed. This means that the logic of the bidding function should only depend on two factors: the estimated utility of the ad display opportunity and the estimated cost to win it. All other information can be regarded as independent with the bid price conditioned
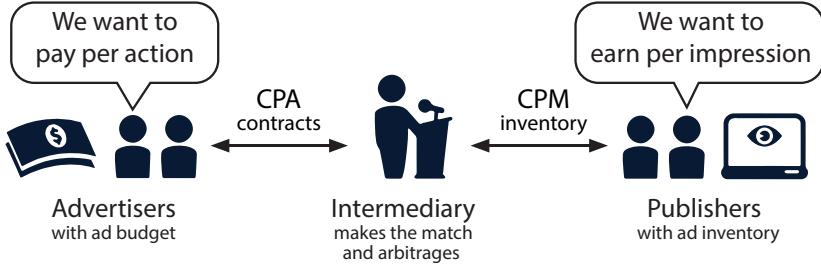
Figure 1.7: The DSP (or other intermediary roles) in display advertising ecosystem may face the risk of buying CPM ad inventory from publishers and earning CPA payoffs from advertisers. This is also possibly a good opportunity to make profits by mining statistical arbitrage opportunities.

only by these two factors[2], as illustrated in Figure 1.6. For example, a sneakers advertiser would like to bid high on users with ages between 15 and 30; this is motivated by the fact that users in such a segment are more likely to be converted to purchase the sneakers after seeing the ads. This is quantified as a higher conversion rate. This is analogous with a high frequency trading strategy in a stock/option market where the trading action is wholly based on the quantified risks and the returns for each asset, regardless of the specific asset attributes or fundamentals [25].

Using the estimated utility and cost, the optimal bidding function to maximise the specific KPI under the target campaign budget and auction volume constraints is derived. This is a functional optimisation framework [26] for each campaign, as will be shown in Chapter 3. In such a framework, the optimal bidding function can be directly derived without any prior restriction on the function form. To our knowledge, such a methodology for bid optimisation has not been studied before in previous RTB display advertising research.

1.3.2 Multi-Campaign Bidding

Furthermore, for a DSP serving multiple campaigns, each of which has a value of conversion, an important research problem would be how to optimise the bidding strategy so as to maximise the total profit from all these campaigns. More interestingly, such a problem can be equivalently studied as a multi-campaign statistical arbitrage problem. Suppose the campaigns all have cost-per-action (CPA) contracts, i.e., the advertisers will pay a predefined value only when the DSP brings a conversion via advertising. In

---

[2]All the information needed to determine the bid has been reflected in the utility and cost factors, just like the conditional independence in probabilistic graphic models.
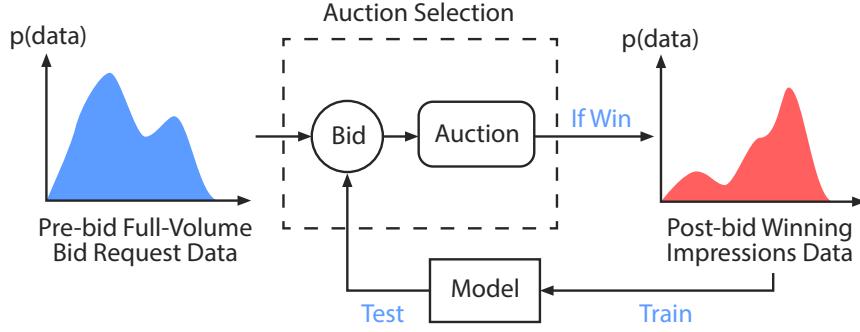
Figure 1.8: In a DSP, all the learning models serve the bidding agent to perform bidding on the pre-bid full-volume data with the true data distribution. However, the user feedback, e.g., clicks, conversions, and the market price of a particular potential ad impression can be observed only if the DSP wins the corresponding auction. Thus, the post-bid winning impression data with feedback labels are censored and biased - this being caused by auction selection. As a result, the learning models trained on such feedback data are biased when subsequently performing predictions or bidding on the pre-bid full-volume data in the next stage.

such cases, the DSP could take the risk of buying ad impressions by the cost-per-mille (CPM) pricing scheme and get the CPA return. This turns out to be a novel statistical arbitrage mining problem for display advertising, as illustrated in Figure 1.7. The research problem is to design a DSP-level bidding strategy to maximise the profit of the DSP with a reasonable total budget and try to avoid negative profit cases.

In our proposed solution framework, in Chapter 4, the profit margin for each campaign with the derived bidding strategy can be estimated, and in relation to this, a portfolio-based solution [27] of auction volume (and budget) reallocation can be performed to optimise the DSP-level profit. We propose to maximise the total expected profit across multiple campaigns under a constrained risk by alternatively learning the auction volume allocation across the campaigns and the DSP-level bidding function in an EM-fashion (Expectation Maximisation). To the best of our knowledge, such an EM-fashion optimisation solution has not been proposed in any previous literature relating to computational advertising.

1.3.3 Unbiased Learning and Optimisation on Censored Data

On the model learning side, the labels of each data instance, e.g., the user's feedback (click or conversion) and the charged price for the ad impression, can be observed by the advertiser only if her bid is high enough to win the auction [14], as shown in Figure 1.8. Such an ad auction selection process causes a strong data bias which in turns results

in model bias in supervised learning tasks such as user response prediction, market price estimation[3], and optimisation tasks, such as bid optimisation. Most existing work on RTB model learning and optimisation fail to consider such data and model bias problems [22, 4, 24, 21, 19].

To deal with such data bias problem, we propose a general learning framework which explicitly models the underlying probability of generating each observed training data instance and incorporate this into the model learning or optimisation process. Specifically, by estimating the market price distribution [28], it is feasible to calculate the auction-winning probability of each auction with the historic bid. Based on an importance sampling method, the unbiased data distribution can be recovered. The derived learning algorithm is called "bid-aware" gradient descent, which incorporates the historic bid price into the learning of each training instance to provide much less biased and more effective learning models for RTB display advertising. Specifically, for bid optimisation tasks, as will be shown in details in Chapter 5, we find the historic bid price not only influences the learning rate for each data instance, but also the gradient directions.

1.3.4 Feedback Control for Handling Data Dynamics

Besides the ability to deal with KPI optimisation on biased data, the robustness of achieving the expected KPI is very important for advertisers. However, in practice such robustness is non-trivial to guarantee due to high dynamics from the impression-level bidding competition and user response behaviour [29]. To illustrate this, Figure 1.9 plots the four major KPIs — CPM (cost per mille), AWR (auction winning ratio), eCPC (effective cost per click) and CTR (click-through rate) — over time for two example campaigns in a real-world RTB dataset. All four KPIs fluctuate heavily across the time under a widely-used bidding strategy [19]. Such instability may result in the risk of unsatisfactory RTB advertising overall performance.

To address this problem, a feedback control mechanism specifically for RTB is proposed in Chapter 6. Using this feedback control mechanism, the advertiser can set a reference value for a specific KPI and the bidding agent will dynamically adjust

---

[3]According to [28]. From an advertiser's perspective, the market price refers to the highest bid price in an ad auction from all other competitors. The advertiser needs to bid higher than the market price to win the auction. Market price estimation is also referred to as bid landscape forecasting [24].
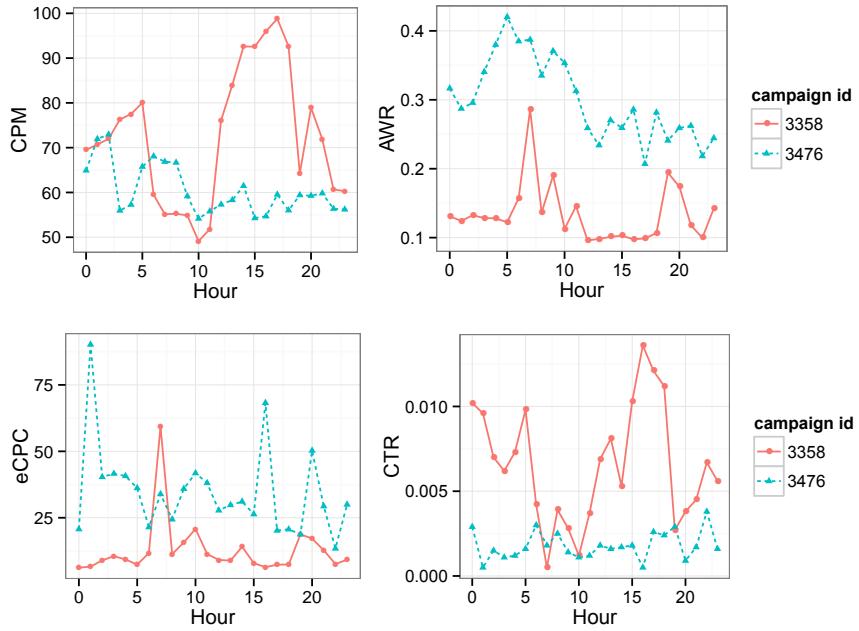
Figure 1.9: The instability of CPM (cost per mille), AWR (auction winning ratio), eCPC (effective cost per click), and CTR (click-through rate) for two example campaigns without a controller. Dataset: iPinYou.

the bids to effectively control the KPI to this reference value. Furthermore, with the assumption that, via feedback control, the reference KPI can be achieved, an optimisation framework is built to calculate the optimal reference value to set so as to optimise the campaign click performance.

1.3.5 Summarised Contributions

The scientific contributions of this thesis are fourfold.

First, the RTB display advertising bidding strategy is abstracted as a function mapping from the bid request features to a bid price. For a single campaign, a general functional optimisation framework is proposed to directly solve the optimal bidding function to maximise the predefined KPI with constraints of the campaign's lifetime auction volume and the budget - without any assumption of the bidding function's form.

Second, for a DSP serving multiple campaigns, a joint optimisation framework is proposed to optimise both the DSP-level bidding strategy and the auction volume allocation across the campaigns to maximise the total profit from these campaigns. In a special case, if the campaigns use the CPA pricing scheme and the DSP buys CPM ad inventories, then there exist statistical arbitrage opportunities, and the proposed framework is capable of maximising the DSP-level profit with a controlled risk.

Third, a general training framework based on importance sampling is proposed to handle the training data bias problem caused by the RTB auction selection. The novel "bid-aware" gradient descent algorithm is derived and is shown to work flexibly in relation to various supervised learning and optimisation tasks in RTB display advertising.

Fourth, a feedback control mechanism, embedded in the bidding agent, is proposed to address the KPI instability problem caused by the highly dynamic user feedback and market competition situation which pertains to RTB display advertising. With the efficacy of the feedback control mechanism, an alternative click optimisation framework is further formulated.

Besides these significant scientific innovations, extensive and repeatable experiments on large-scale real-world data have been performed to verify the effectiveness of each proposed solution. More importantly, these proposed solutions are quite flexible in relation to any models of the bid request utility and cost estimation, which means that there would be only small engineering costs involved in their deployment to different real-world platforms. All the proposed bidding strategies, training schemes and feedback mechanisms have been (once) deployed in one of three commercial DSP systems located in China and the United States and have provided significant performance improvements over strong baselines.

In summary, the scientific and empirical contributions of this research are significant in terms of moving towards optimal RTB display advertising performance.

1.4 Thesis Structure

The rest of this thesis is organised as follows. In Chapter 2, we perform a literature review of related techniques and their connections to our research. In Chapter 3, we study the fundamental single-campaign bid optimisation problem and propose a general functional optimisation framework to derive the optimal bidding function. In Chapter 4, we discuss the advanced multi-campaign bid optimisation in the statistical arbitrage mining scenario. In Chapter 5, we propose a bid-aware learning/optimisation framework to reduce the model bias caused by the ad auction selection. In Chapter 6, to achieve robust advertising performance against RTB data dynamics, we propose a feedback mechanism to control the advertising KPIs via adaptive bid adjustment. Finally, we conclude this thesis and discuss future research plans in Chapter 7.

1.5 Supporting Publications

The selected publications which support the chapters of this thesis are listed as below. The full publication list is available in Appendix A.

1) Weinan Zhang, Tianxiong Zhou, Jun Wang, and Jian Xu. Bid-aware gradient descent for unbiased learning with censored data in display advertising. In *KDD*, 2016. (**Main text of Chapter 5**)

2) Weinan Zhang, Ulrich Paquet, Katja Hofmann. Collective noise contrastive estimation for policy transfer learning. In *AAAI*, 2016. (**Techniques partially supporting Chapter 5**)

3) Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. Feedback control of real-time display advertising. In *WSDM*, 2016. (**Main text of Chapter 6**)

4) Weinan Zhang and Jun Wang. Statistical abitrage mining for display advertising. In *KDD*, 2015. (**Main text of Chapter 4**)

5) Weinan Zhang, Shuai Yuan, and Jun Wang. Optimal real-time bidding for display advertising. In *KDD*, 2014. (**Main text of Chapter 3**)

6) Weinan Zhang, Shuai Yuan, Jun Wang, and Xuehua Shen. Real-time bidding benchmarking with iPinYou dataset. *arXiv:1407.7073*, 2014. (**Data analysis supporting Chapters 3, 4, 5 and 6**)

7) Weinan Zhang, Jun Wang, Bowei Chen, and Xiaoxue Zhao. To personalize or not: a risk management perspective. In *RecSys*, 2013. (**Techniques partially supporting Chapter 4**)

Chapter 2

Related Work

This chapter covers various techniques which form the background of the research problems or the basis of proposed solution models. In Section 2.1, the background of display advertising is discussed. In Section 2.2, the related research work on keyword bid optimisation in sponsored search is discussed, which provides a comparison to the bid optimisation techniques in RTB display advertising discussed in Section 2.3. The latter supports the modelling parts of Chapters 3 and 4 and the experiment comparison in Chapters 3-6. Later in Section 2.4, financial risk management models are discussed, which serve as basis of the proposed risk-sensitive statistical arbitrage models in Chapter 4. Then a survey of research work on learning and evaluation on biased or censored data is provided in Section 2.5 to support the technique background of the solutions in Chapter 5. In Section 2.6, the feedback control theory is discussed, which supports the basis of the proposed feedback control system in Chapter 6.

2.1 Display Advertising

Display advertising is one of the major types of online advertising that comes in several forms, including banner ads, rich media and more. Unlike text-based ads in sponsored search, display advertising relies on elements such as images, audio and video to communicate an advertising message with the users [11]. It plays a crucial role in the online marketing for both branding and performance-driven campaigns. According to the IAB's report in 2015 [5], display advertising accounted for $18.9 billion (38% of total) revenue during 2014.

Before the emergence of the auction-based RTB market, the transactions between advertisers and publishers of display advertising were mostly made by direct negoti-

ations or through ad networks [14]. The user targeting settings were normally based on inferred user demographics [30]. Some transactions guaranteed the ad impression volume within a certain period, typically several weeks or months, called guaranteed delivery [31], and some were based on CPM pricing scheme but did not guarantee the delivered volume, called non-guaranteed delivery [15]. Before the popularisation of RTB in 2011 [17], most research work on display advertising optimisation was about ad inventory allocation across campaigns on behalf of publishers in order to maximise the overall revenue or other advertising KPI with the guaranteed delivery constraints [32, 33, 34, 35] using linear programming techniques. The authors in [31] further proposed an automatic model for pricing the guaranteed-delivery contracts based on the prices of the targeted individual user visits in a spot market.[1] Generally speaking, such a conventional display advertising trading mechanism is of large granularity on user targeting and the advertisers cannot perform acquisition on demand thus it is of relatively low efficiency.

With the rise of ad exchange and RTB in 2011, a lot of work emerged on auction-based optimisation for display advertising [17]. RTB mechanism enables advertisers to bid for an individual ad impression with a specific user in a specific context [14], thus immediately solves the inefficiency problem of conventional ad trading mechanism of display advertising.

On the advertiser side, the bid optimisation for campaign performance improvement is studied. The authors in [36] proposed a budget pacing scheme embedded in a campaign conversion revenue optimisation framework to maximise the campaign revenue. The authors in [19] focused on a bidding function formulation to maximise the campaign clicks. Bid landscape forecasting models [24] were studied to estimate the campaign's impression volume and cost given a bid price. A detailed discussion of research work on bid optimisation of RTB display advertising will be given in Section 2.3.

On the publisher side, the placement-level reserve price optimisation was studied in [37], where the authors considered the reaction between the publishers' reserve

---

[1]Here the term "spot market" comes from finance, which originally means a public financial market where the commodities are traded for immediate delivery. In display advertising, the traded commodities are ad display opportunities. The "spot market" here means display advertising marketplaces with non-guaranteed delivery contracts or RTB ad auctions.

price and the advertisers' bid price and modelled this problem from a game-theoretic prospective. The authors in [38] suggested that the publisher could act as a bidder on behalf of its guaranteed contracts so as to make smart inventory allocations among the guaranteed and non-guaranteed delivery contracts. One step further, the authors in [39] proposed a mathematical model of allocating and pricing the future ad inventory between guaranteed delivery contracts and RTB spot markets. Their solution assumed that advertisers are risk-averse and prefer guaranteed ad inventories if there is positive profit. Also the advertisers' purchase behaviour is based on the option price and the starting date of the guaranteed contracts.

2.2 Bid Optimisation for Sponsored Search

As auctions have been a major trading mechanism of online advertising, bid optimisation becomes a crucial problem for advertisers [40, 41, 42, 19]. Nonetheless, most research has been so far limited to keyword auction in the context of sponsored search [10, 43, 44]. Typically, under the scenario of pre-setting the keyword bids (not impression level), the keyword utility, cost and volume are estimated and then an optimisation process is performed to maximise the advertisers' objectives (KPIs) [45, 46, 47, 48]. Specifically, the keyword utility is normally evaluated as the expected CTR on a specific ad slot in the search engine result page of this keyword; the cost is estimated as the market competitiveness on bidding this keyword; the volume of a keyword is based on the estimation of the volume of users' relevant search queries [48, 45]. Given a campaign budget as the cost upper bound, optimising the advertiser performance is defined as a budget (allocation) optimisation problem [40, 49]. Furthermore, the authors in [50, 51] focused on the bid generation and optimisation on broad matched keywords, where the user's query keywords do not exactly match the bid keywords. In the solution, query language features were leveraged to infer the relevance between the search query and bid keywords, which then helped estimate the optimal bid price of the broad match. Extending the bid optimisation to the multi-campaign level, the authors in [52] proposed to jointly optimise the keyword-level bid and account-level budget allocation under a multi-campaign sponsored search account.

Some recent work focused on periodically changing the pre-setting keyword auction price, taking into account the remaining budget and lifetime. For instance, in

[28, 53], Markov decision process was used to perform online decision in tuning the keyword bid price, where the remaining auction volume and budget acted as states and the bid price setting as actions. In [47] authors proposed to calculate a bid allocation plan during the campaign lifetime, where the bid price on each keyword was set in different discrete time unit by considering the market competition and the CTR on different ad positions. However, none of the work evaluated per-impression auction because in sponsored search all the bids are associated with keywords and impression-level features are seldom considered, especially for advertisers and their agencies.

Moreover, in sponsored search bid optimisation, a search engine plays as two roles: setting the keyword bids as well as hosting the auctions. The objective function could be diverted to optimise the overall revenue for the search engine [54, 55, 56, 57], rather than the performance of each advertiser's campaigns.

2.3   Bid Optimisation for RTB Display Advertising

The bid optimisation problem for RTB display advertising is fundamentally different from that for sponsored search. First, the bids are not determined by pre-defined keywords in a search session, but are based on impression-level features in a general web or mobile page view [14]. In addition to setting up their target rules, advertisers or DSPs need to estimate the value of each ad impression that is being auctioned in real time and accordingly make the bid decision for each ad auction [19]. As such, the bid optimisation is never just to find a bid value for a keyword or a part of user volume but to design an effective bidding function to calculate the bid values for billions of different ad auctions daily. Second, in RTB, CPM pricing scheme is generally used [19, 14], which is different from the CPC pricing scheme in sponsored search. Winning an impression directly results in the cost. Thus, the dependencies over various effectiveness measures such as eCPC[2], CPM, campaign lifetime volume and budget constraints need to be jointly studied in a single framework. Third, DSPs act on behalf of the advertisers and optimise the served campaigns' performance while the search engine primarily serves for itself and the main goal is to maximise the total revenue. For DSPs, the observed auction data and user feedback data is partial – only the data

---

[2]Effective cost per click (eCPC) - The cost of a campaign divided by the total number of clicks delivered.

related to the served campaigns is observed, while for search engines the whole data from all campaigns can be observed [52].

With the emergence of ad exchanges for display advertising in 2009, dynamic bidding strategies start to be investigated. In [42], the authors proposed an algorithm that learned winning bids distribution from full or partial information of auctions in display advertising. The algorithm then made bidding decisions to achieve the best delivery (i.e., number of impressions) with the budget constraint. In [21], the bid price from each campaign could be adjusted by the publisher or the supply-side platform in real time and the target was to maximise the publisher's revenue. Borrowing the idea of the optimal truth-telling bidding in sponsored search [10], a basic and reasonable bidding strategy is to bid the estimated true value for each ad impression. For performance-driven campaigns, the predefined true value is normally based on actions, such as a click or a conversion. As such, the expected true value for a specific impression is estimated as the action value times the auction rate, e.g., click value times CTR (click-through rate) and conversion value times CVR (conversion rate) [4, 21]. However, the truth-telling bidding strategy is optimal only when the budget and auction volume are not considered. With the campaign lifetime auction volume and budget constraints, the optimal bidding strategies are probably not truth-telling. Extending from the truth-telling bidding strategy, the authors in [19] proposed the generalised bidding function with a linear relationship to the predicted CTR for each ad impression being auctioned. Compared to [19], the analytic solution from our proposed functional optimisation framework (in Chapter 3) shows that an optimal bidding function could be non-linear. The non-linearity is closely related to the bidding landscape (i.e., the market price distribution [28]), but is loosely correlated with the prior distribution of the bid request features. In a recent study [58], the authors proposed to set the bid price for each ad impression proportional to the lift (instead of the absolute value) of the user conversion rate after seeing the ad. Their empirical study showed such a lift-based bidding strategy would bring more user conversions to the advertisers but the DSP would be unfortunately attributed with fewer user conversions because of the disadvantages of the last-touch conversion attribution mechanism [59].

As a new advertising paradigm, other problems related to the bidding in RTB display advertising have also been studied. In [36], the authors focused on the pacing

problem, where the target was to smoothly deliver the campaign budget. In [4], the sparsity problem of conversion rate estimation was handled by modelling the conversions at different selected hierarchical levels. In [2], the authors studied the evaluation measures of the display advertising performance and they found the site visit turned to be a better proxy than the user click. In addition, there is some work on the ad exchange communication problem [60, 16]. More discussion on related research of RTB can be found in [14].

2.4 Risk Management Models and Applications

Risk is a consequence of action taken in spite of uncertainty [61]. And the objective of risk management is to assure uncertainty does not to some extent deflect the business from its goals [62, 63]. Computational advertising, as one of the advertising paradigms, is associated with a certain level of risk of deficit, no matter how high ROI expectation it could bring. The risk comes from the dynamics of advertising market and the user online behaviour [64].

In Chapter 4, a risky business in RTB display advertising, called statistical arbitrage mining, will be studied, where the intermediary agent has the possibility of a loss when running campaigns with contracts of performance-based pricing schemes. As such, risk management is a key feature in the effective solution.

In this section, modern portfolio theory, the quantitative approaches of risk management, and its applications will be discussed. Then, the statistical arbitrage problem in finance and its links with the research in Chapter 4 will be discussed.

2.4.1 Modern Portfolio Theory

Modern portfolio theory (MPT), Harry Markowitz's Nobel Prize work [65], originates from modelling uncertainty of the return of multiple financial asset combinations. It is desirable to have quantitative methods to measure such uncertainty (or risk) and model it into the decision making of investment [66]. In MPT, the variance of the return of each asset is modelled as its risk. Then the risk and expected return of a portfolio of invested assets can be quantified based on the funding allocation, the mean return of the assets and the covariance matrix of them [65, 66]. MPT utilises the mean-variance analysis to make an investment portfolio for any tradeoff between the risk and the

expected return, or w.r.t. a reference investment such as bank deposits [67]. MPT is a model-free quantitative framework that can incorporate different risk and return estimation models. With such advantages, MPT has been adopted in almost everywhere of financial investment [68].

Recently, the ideas of risk management have been introduced to the information retrieval area to improve the model robustness or capture the users' satisfaction psychologically [69, 70, 71, 72, 73, 27, 74].

Risk-sensitive models were proposed to measure the uncertainty of document ranking performance in web search tasks. The authors in [69] proposed an asymmetric loss between the estimated relevance and the true relevance of a query-page pair to explicitly punish more on the overestimation cases. The authors in [70] studied the risk-sensitive evaluation measure when comparing a new document ranking model to the existing baseline model. Specifically, the downside risk factor of the new model was defined as the averaged performance reduction due to using the new model compared to the baseline model. The risk factor was then used for model selection with different risk-averse levels. The risk-sensitive optimisation models for query expansion were studied in [71], where the authors used the variance of relevance of the expanded queries as the risk and an objective with the balanced relevance expectation and risk to make the expanded queries of both high quality and low risk of irrelevance. In [72], the idea of leveraging MPT to optimise document ranking was formally presented. For a given query, the mean and variance of the relevance of each document were defined as the expected return and risk, respectively; the probability of the user noticing each document of the search engine result page was regarded as the funding allocation. The optimisation objective was again the balance between the risk and return with the risk-averse level as the hyperparameter of the model.

Moreover, for recommender systems, the risk of delivering an item or a list of items is also significant. However, quite little work explicitly models the risk to improve the performance of recommender systems. In [73], the idea of MPT for ranking documents [72] was borrowed to collaborative filtering (CF) to adaptively diversify the recommended items. Specifically, in a matrix factorisation CF framework, the variance of each user's latent feature vector was defined by its averaged divergence from the rated items' latent feature vectors. As such, the mean and variance of a rating

from a user to an item could be derived. Then a risk-return balanced objective was optimised to rank the items. As a result of risk reduction, the items were naturally diversified even though the item categorical information was never used in the training stage. Furthermore, the authors in [27] proposed a risk-sensitive switching model between personalised and non-personalised recommendation lists, which were regarded as two portfolios. Probabilistic matrix factorisation [75] was adopted to naturally model the mean and variance of user-item ratings. Compared with [27], [73] mainly focused on diversification of recommended items instead of carefully modelling the individual item risk and the decision making of whether to personalise. Furthermore, the risk-sensitive approaches were extended to modelling privacy and security factors in recommendation. The risk concepts were naturally applied in financial item recommendation scenario [76, 77], where the uncertainty of the return from each investment opportunity (or startup company) was modelled into the recommendation objective of an MPT framework. Moreover, in [74], the security risk of each mobile application was considered and the proposed recommendation objective balanced the mobile application popularity and the users' security preferences using MPT.

To our knowledge, there is almost no work adopting MPT into the profit optimisation with risk diversification in online advertising. In [78], the authors studied a batch-mode ad selection from a publisher's perspective in order to maximise the total profit. MPT was adopted to model the risk of selected ads and their correlations. Compared to the RTB display advertising scenario with bidding and auction selection, the studied scenario in [78] is closer to that in recommender systems. In Chapter 4, a novel way of applying MPT is proposed and it is naturally integrated into our bidding strategy designing and optimisation framework to balance the risk and expected reward of a package of ad campaigns in a competitive RTB display advertising market.

2.4.2 Statistical Arbitrage

In financial markets, as a trading strategy, statistical arbitrage is a quantitative approach to security trading. It utilises statistical methods with high-frequency trading systems to detect statistical mispricing of securities caused by market inefficiency to make profit with a large number of transactions [79]. For the statistical arbitrage opportunities, the payoff is stochastic with a positive expectation and a certain variance, where the

optimal action is mainly based on portfolio selection [80, 81].

Drawing an analogy with the statistical arbitrage of security pairs trading in finance [82], for the statistical arbitrage mining problem in RTB display advertising as will discussed in Chapter 4, the campaign's CPA contract and its performance in RTB spot markets can be regarded as a pair of correlated securities. Statistically speaking, if the campaign's performance in an RTB spot market ensures that the average cost to acquire a conversion (i.e., eCPA) is lower than the payoff from the CPA contract, then a statistical arbitrage opportunity exists. Such an opportunity could also be considered to be caused by informational inefficiency of the advertising market where the advertisers fail to lower their CPA payoff when their campaigns in the RTB spot market have a good performance.

The authors in [83] studied auction mechanisms considering arbitrage between CPC and CPM pricing schemes. The study aimed at designing an auction mechanism on behalf of the ad exchange and yielding truthful bidding from advertisers and truthful CTR reporting from arbitrageurs. By contrast, the research in Chapter 4 aims at developing a statistical method for mining and exploiting arbitrage opportunities between CPA and CPM on behalf of a DSP.

2.5 Learning on Biased or Censored Data
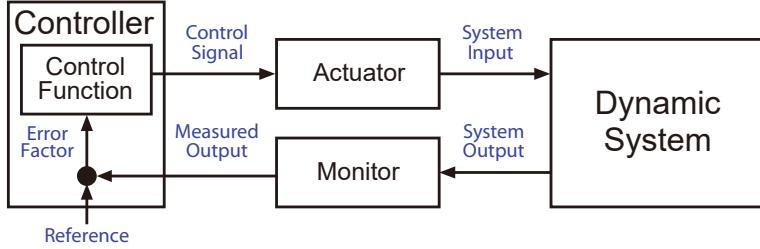
As pointed out in [84], direct online evaluation and optimisation for a new learning model solution are expensive and risky, which is also a dilemma in online advertising [85]. By contrast, it is cheap and risk-free if the model can be optimised and evaluated using offline historic data that was previously collected using another model. The authors in [86] proposed to use historic data for unbiased offline evaluation of news article recommendation model by historic data replay and rejection sampling. Prerequisites of this approach are that the previous model generating the training data (called exploration model) is known, and that it has sufficiently explored all actions in the support of the evaluated policy [87]. For cases where historic data is collected using a biased (non-uniform) or non-stationary policy, the authors in [88] suggested an adaptive rejection sampling approach. For cases where the exploration model is unknown, an evaluation scheme with the estimated propensity scores and a lower bound of the data observation probability was proposed in [89]. In the CTR estimator learning and bid optimisation

based on the historic data with the observed market prices and user feedback in RTB display advertising, the exploration model is known as the historic bid price is available for each bid request.

Handling the missing data is a well-studied problem in machine learning with various applications [90], such as recommender systems and online advertising.

A classic application in recommender systems is item recommendation with implicit feedback [91, 92], where the observations are only (implicit) positive instances, such as the users' movie watching (not rating) records and web browsing history etc., while the unobserved positive and negative instances are mixed together. The authors in [91] proposed a uniform, a user-oriented and an item-oriented sampling methods of negative items to build the data matrix for learning collaborative filtering models. The authors in [92] further proposed user response models to learn the missing data distribution instead of regarding it as completely random observations. An intuitive motivation of this work is that the observation of a user-item rating behaviour depends on the rating score. For example, a user is more likely to rate an item if she likes it. The proposed missing data model re-estimated the latent distribution of the data and the learned models were demonstrated to be more effective in the empirical study.

In online advertising, the market price data and user feedback data are censored because of the auction selection, i.e., the data instances associated with higher bid prices are more likely to win the auction, thus to be observed [24]. For the market price distribution modelling problem, the authors in [28] leveraged a non-parametric survival model to incorporate the partial information from the losing cases, i.e., only knowing the market price is higher than the bid price rather than the exact value, for less unbiased estimation in sponsored search. In a recent work [93], the authors proposed to use censored linear regression to model both the observed winning data and the censored losing data. Specifically, the likelihood of the observed market price was modelled with Gaussian noise, and the likelihood of the censored data was modelled with the probability of the predicted market price higher than the bid price. In order to collect comprehensive data observations to train less biased CTR estimation models, the authors in [94] added a covariance factor periodically to the Baysian Probit regression weight vectors to incorporate dynamics in the prediction stage. However, they failed to consider the data bias in the learning or optimisation stage. Compared

Figure 2.1: Feedback control loop.

to the discussed work, the main novelty of the research in Chapter 5 lies in directly incorporating the auction-selection data bias into the learning process to perform the "bid-aware" gradient descent training algorithm to learn an unbiased model.

2.6 Feedback Control Theory and Applications

As defined in [95], a "dynamic system" refers to the system with its behaviour changes over time, normally in response to different external force or stimulation as input signal. In a situation with two or more dynamic systems interacting with each other, their dynamics are strongly coupled. This situation refers to "feedback". Feedback control theory deals with the reaction and control of dynamic systems from feedback and ambient noise [96].

Figure 2.1 briefly shows the interactions between the controller and the dynamic system. The usual objective of feedback control theory is to control a dynamic system so that the system output follows a desired control signal, called the reference, which may be a fixed or changing value. To do this, a controller is designed which monitors the output and compares it with the reference. The difference between actual and desired output, called the error factor, is applied as feedback from the dynamic system to the control system. With the specific control function, the controller outputs the control signal, which is then transformed by the actuator into the system input signal sent back to the dynamic system. These processes form a feedback control loop between the dynamic system and the controller.

There were many instances of feedback control in the ancient history [97]. The most typical example was to exploit the feedback mechanism to improve the accuracy of water clocks [98]. In modern age, with the invention of steam engine in 18th century and the later formal concept proposal in 1930's from Bell Labs [99], automatic control emerged and was subsequently employed routinely with other system components such

as actuators and sensors [100]. Now feedback control techniques have been widely used in various engineering applications for maintaining some signals at the predefined or changing reference values, such as robot navigation [101] and indoor temperature maintenance [102].

Recently, feedback control techniques have been leveraged in intelligent systems. There are a few research papers on recommender systems leveraging feedback controllers for performance improvement and maintenance. In [103], a rating updating algorithm based on the proportional-integral-derivative (PID) controller was developed to exclude unfair ratings in order to build a robust reputation system. The authors in [104] applied a self-monitoring and self-adaptive approach to perform a dynamic update of the training data fed into the recommender system to automatically balance the computational cost and the prediction accuracy. Furthermore, the authors in [105] adopted the more effective and well-studied PID controller to the data-feeding scheme of training recommender systems, which was proven to be practically effective in their studied training task.

Compared to the work of controlling the recommender system performance by changing the number of training cases, the studied control task in RTB (Chapter 6) is more challenging, with various dynamics from advertising environment such as the fluctuation of market price, auction volume and user behaviour patterns. In [21], the authors discussed multiple aspects in a performance-driven RTB system, where the impression volume control was one of discussed aspects. Specifically, a waterlevel-based controller and a model-based controller were implemented to control the impression volume during each time interval. In [106], feedback control was used to perform budget pacing in order to stablise the conversion volume. Compared to [21, 106], the research work in Chapter 6 is a more extensive study focused on the feedback control techniques to address the practical instability problem in RTB. Besides waterlevel-based controller, the more sophisticated PID controller is extensively investigated. Regarding the controlled KPIs, the control tasks on both eCPC and AWR are studied, which are crucial KPIs for performance-driven campaigns and branding-based campaigns, respectively. In addition, an effective model to calculate the optimal eCPC reference to maximise the campaign's clicks using feedback controllers is proposed.

Chapter 3

Single-Campaign Optimal Real-Time Bidding

3.1   Background and Motivation

Demand-Side Platforms (DSPs) help advertisers manage their campaigns and perform their real-time bidding activities. Figure 3.1 briefly illustrates the role of a DSP in the RTB eco-system and some of its important modules. In RTB display advertising, once a user visits a web page and an ad impression is to be created, an ad request for the impression is immediately triggered by the publisher (usually the Supply-Side Platform, a.k.a. SSP, a technology platform to manage publishers' ad inventories) and then sent to the DSPs via an ad exchange. On behalf of an advertiser, the DSP will compute a bid for this impression and return a bid response to the exchange, where a second price auction is usually held to select the winner. Finally the winner is notified and her ad is displayed to the user through the publisher.

More specifically, after receiving a bid request, the DSP will find all eligible ad creatives from all campaigns[1] and compute a bid for each of them. The DSP uses both contextual [107] (e.g. domain, web page, keywords, time and date, geographical location, weather, language, operating system, browser, etc.) and behavioural [108] (e.g. search, browsing, and purchase history, occupation, income, sentiment, etc.) data to compute a bid. It is common and usually encouraged that advertisers buy user interest segments from third-party data providers [18], e.g., a data management platform

---

[1]The eligibility of the creative and campaign means their target combinations match the bid request, such as placement size, user demographics, geographical location, language, etc. It acts as pre-filtering rules before the bidding process.
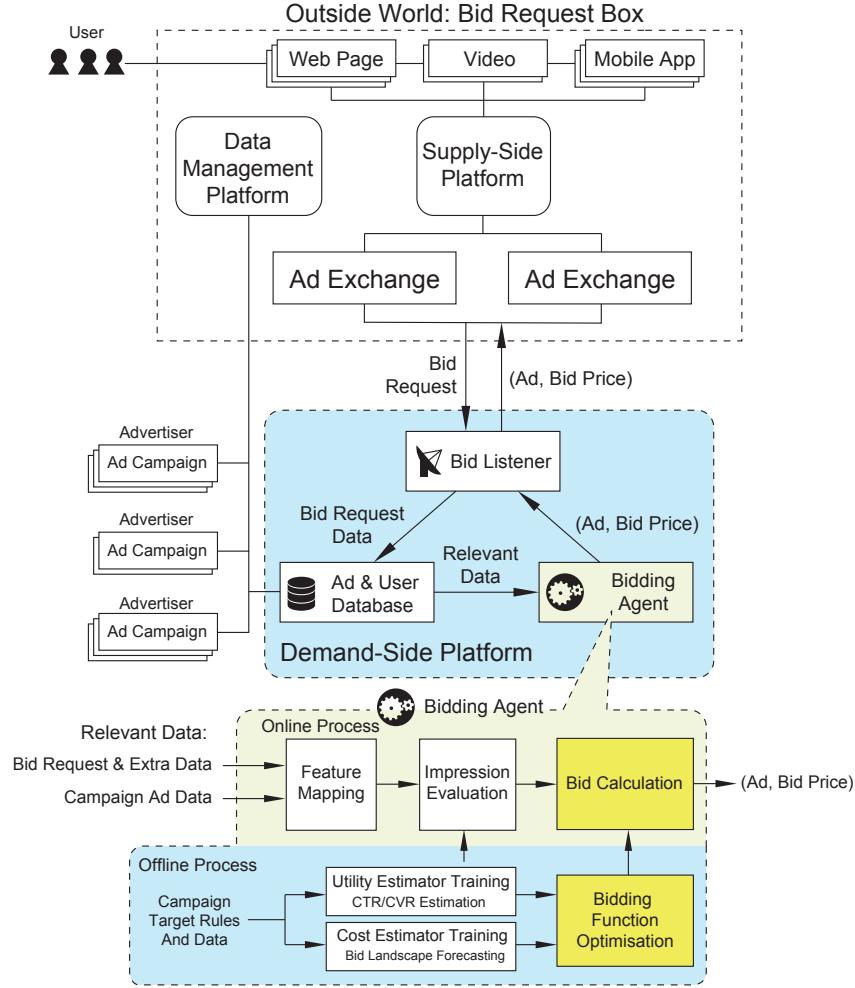
Figure 3.1: An illustration of a demand-side platform and its bidding agent in RTB display advertising.

(DMP). Note that although we confine our work to the cost-per-mille (CPM) pricing scheme which is commonly adopted in RTB, other less popular models are also available (e.g. cost-per-click and cost-per-action).

This bid calculation (see the bidding engine in Figure 3.1) is the most important problem for a DSP. The solution to this problem is referred to as a *bidding strategy*. In pure second price auctions [109] for strategic competitors, theoretically the dominant strategy for advertisers is truth-telling: to bid their private values [10]. When facing a bid request, a DSP will evaluate the value of the impression i.e. to estimate the click-through/conversion rate (CTR/CVR) and multiply it by the value of a click/conversion [4]. Many advertisers simply set this value as their bid [19, 4] and keep using it throughout a campaign's lifetime. However, when computing a bid, practical constraints need to be taken into account including the bid landscape (the auction winning rate against

the bid price), total budget and the campaign's remaining lifetime. With such practical constraints, the optimal bidding strategy is not truth-telling any more. These considerations enable the DSP to optimise towards the overall performance of a campaign (usually quantified via a Key Performance Indicator, KPI, e.g. the number of clicks, conversions or total revenue) using stochastic methods rather than assuming advertisers are strategic and have a private "true" value per impression [28].

In this chapter, the impression-level bidding strategy is formulated as a function that maps the individual impression evaluation to the bid value. A novel functional optimisation framework is proposed to find the optimal bidding function: (i) given the budget constraint and the campaign's lifetime, and (ii) taking into account various data statistics such as the bid landscape [24] and the prior distribution of the bid request features. The derived analytic solution indicates that the auction winning function (from bid landscape) plays a more critical role in shaping the bidding function, whereas the distribution of the features is less correlated. Simple winning functions derived from practical bidding data result in optimal bidding functions that are *non-linear* and in a *concave* form. Unlike the linear function previously proposed [19], the derived bidding function encourages to raise bids for impressions with low estimated value because compared to higher evaluated ones, those are more cost-effective and the chance of winning them are relatively higher. We also show that the linear bidding function can also be derived from our proposed functional optimisation framework under the (strong) assumption of the winning function being linear. Apart from the theoretic insights, both offline experiments on a real dataset and online experiments on a commercial DSP show that the proposed bidding strategies outperform the strong baselines that have been considered in previous work.

To summarise, the contributions of this chapter are listed as follows. (i) We propose a novel functional optimisation framework to find the optimal bidding strategy given a single campaign in RTB display advertising. (ii) Based on the auction winning function built from the data, the derived optimal bidding function from our framework is in the concave form against the KPI of each impression, which to our knowledge has not been studied in previous literature. (iii) Extensive offline and online experiments are conducted to verify the effectiveness of our proposed bidding strategies.

Table 3.1: Notations and descriptions.

Notation	Description
x	The bid request represented by its features.
$p_x(x)$	The probability density function (p.d.f.) of x.
$\theta(x)$	The predicted KPI if winning the auction of x. It could be the CTR, CVR etc.
$p_\theta(\theta)$	The probability density function of KPI θ.
B	The campaign budget.
T	The estimated number of bid requests during the lifetime of the campaign.
$b(\theta(x), x)$	The bidding strategy is defined as function $b()$. Assume a generative process: $x \rightarrow \theta \rightarrow b$, so $b(\theta(x), x) \equiv b(\theta(x))$. See the text. We occasionally use notation b to refer to a specific bid value.
$w(b(\theta(x)), x)$	The probability of winning the bid request x with bid price $b(\theta(x))$. We approximate it by the dependency assumption: $x \rightarrow \theta \rightarrow b \rightarrow w$, so $w(b(\theta(x)), x) \equiv w(b(\theta(x)))$. See the text for details.
$c(b(\theta(x)), x)$	The expected cost after winning the bid request x with bid price $b(\theta(x))$. Like $w(b(\theta(x)), x)$, approximated by the dependency assumption: $x \rightarrow \theta \rightarrow b \rightarrow c$, so $c(b(\theta(x)), x) \equiv c(b(\theta(x)))$. See the text for details.

3.2   Problem Definition

Each view of a publisher page triggers an ad auction for each ad slot on that page, which, in real time, generates a bid request for each targeting campaign. Given a bid request, along with features covering user, ad, context, and auction information, the DSP bidding engine decides whether to participate this auction, and if participates, returns a bid for this auction. The bid price depends on many factors. It is not only influenced by the predicted KPI value of the ad impression being auctioned, such as CTR and CVR, that the advertiser wants to achieve, but most importantly, related to many other factors such as the budget constraint, the probability of auction winning, and estimated cost after winning this particular ad impression. In this section, we formulate the problem of optimally generating real-time bids as a functional optimisation problem and propose a novel optimisation framework by taking all these factors into account. We show that solving it leads to a practical bidding function.

To launch a campaign in display advertising, the advertiser uploads their ad creatives, sets the targeting rules (e.g. the user segmentation, time, location) and the corresponding budget for the lifetime of the campaign.

After the targeting rules are set, before optimising the bid, the advertiser would first spend a small amount of budget to bid random impressions in order to learn some statistics (to serve as the training data for the initial model). For instance, as studied in

[110, 24] the auction volume forecast (e.g. bid landscape prediction) module is usually employed to estimate auction statistics w.r.t. the current setting and budget constraint. More specifically, we denote the estimated number of bid requests for the targeting rules during the lifetime as T and the campaign budget as B. Also, each bid request is represented by a high dimensional feature vector $\boldsymbol{x}$, where its entries consist of two sets of features: one is exacted from the campaign ad itself, and the other is related to the impression being auctioned, e.g., the cookie information of the underlying user, location, time, user terminal, browser, the contextual information about the web page, etc.[2] We use $p_x(\boldsymbol{x})$ to denote the prior distribution of the feature vectors which match the campaign targeting rules. For each campaign, the advertiser can use the historic bidding and feedback data to predict the KPI for the ad impression. We denote the predicted KPI of a bid request $\boldsymbol{x}$ as $\theta(\boldsymbol{x})$. Note that different advertisers might consider different KPIs. For example, if the goal of a campaign is to maximise the direct visits, i.e. the total number of clicks, then $\theta(\boldsymbol{x})$ denotes the predicted CTR for that impression. If the goal of a campaign is for conversions, then $\theta(\boldsymbol{x})$ denotes the predicted CVR for that impression. Moreover, we denote $p_\theta(\theta)$ as the prior distribution of the predicted KPI per bid request. The notation descriptions are given in Table 3.1.

Once the major statistics are gathered, the problem now is to design an optimal bidding strategy such that a certain KPI objective over the budget will be maximised. For the sake of clarity, the number of acquired clicks during the campaign's lifeime is studied as the objective here, while a natural extension to covering alternative KPIs and their experimental results are given later in Section 3.5.5.

Mathematically, the optimal bid generation problem is formulated as a functional optimisation problem:

$$b()_{\text{ORTB}} = \underset{b()}{\arg\max} \ T \int_{\boldsymbol{x}} \theta(\boldsymbol{x}) w(b(\theta(\boldsymbol{x}), \boldsymbol{x}), \boldsymbol{x}) p_x(\boldsymbol{x}) d\boldsymbol{x} \tag{3.1}$$

$$\text{subject to} \ T \int_{\boldsymbol{x}} c(b(\theta(\boldsymbol{x}), \boldsymbol{x}), \boldsymbol{x}) w(b(\theta(\boldsymbol{x}), \boldsymbol{x}), \boldsymbol{x}) p_x(\boldsymbol{x}) d\boldsymbol{x} \leq B.$$

where $b(\theta(\boldsymbol{x}), \boldsymbol{x})$ denotes the bidding function we intend to obtain. It depends on the

---

[2]Commonly, $\boldsymbol{x}$ is encoded as a high-dimensional sparse binary vector, where each entry is set as 1 only when the corresponding field takes the particular category, e.g., city=London, weekday=Sunday etc. Furthermore, the feature space could be compressed by hashing tricks [111, 112].

feature vector $\boldsymbol{x}$ and the estimated CTR $\theta(\boldsymbol{x})$. $w(b,\boldsymbol{x})$ denotes the estimated winning rate for a bid price b given the feature $\boldsymbol{x}$ of the impression auction. In Eq. (3.1), the product of $\theta(\boldsymbol{x})$ and $w(b,\boldsymbol{x})$ produces the probability of click given an impression auction. Marginalising it over the feature space yields the expected click per auction. Note that in practice, the impression auctions arrive sequentially, so one can potentially make a sequential bidding rule by taking a feedback loop and employing a dynamical optimisation model such as partially observable Markov decision processes (POMDPs) [113]. However, generally these models are computationally expensive thus not feasible in our case, where bid decisions are usually required to be returned within 100ms. We, thus, take a two-stage approach, i.e., learning statistics such as $p_x(\boldsymbol{x})$ and T, then optimising the bids. As a practically feasible assumption, we consider a simple static model and follow a widely accepted assumption in the previous bid optimisation work [53, 21]: each time the feature vector is independently generated from an identical distribution (i.i.d. assumption). Such assumption is convincing at least in the case of a short period, where the market bid landscape and general user behaviour patterns can be regarded as unchanged.

The constraint is formulated by the cost function $c(b(\theta(\boldsymbol{x}),\boldsymbol{x}),\boldsymbol{x})$ for each bid request and the campaign total budget B. RTB normally applies the second price auction (pay the second highest bid) for the majority of ad volume. However, due to the reserve price setting, the cost is quite often higher than the second highest bid [37, 14]. In some first-price auctions or second-price auctions with high soft reserve price setting, the winner still pay the bid price. In order to deal with these various auction cases together, the cost function in our research is defined as the upper bound of the cost, i.e., the bid price $b(\theta(\boldsymbol{x}),\boldsymbol{x})$. Specifically, the product of the cost and the winning rate produces the expected cost per auction. The functional optimisation framework is rewritten as

$$b()_{\text{ORTB}} = \underset{b()}{\arg\max} \; T \int_{\boldsymbol{x}} \theta(\boldsymbol{x})w(b(\theta(\boldsymbol{x}),\boldsymbol{x}),\boldsymbol{x})p_x(\boldsymbol{x})d\boldsymbol{x} \tag{3.2}$$

$$\text{subject to } T \int_{\boldsymbol{x}} b(\theta(\boldsymbol{x}),\boldsymbol{x})w(b(\theta(\boldsymbol{x}),\boldsymbol{x}),\boldsymbol{x})p_x(\boldsymbol{x})d\boldsymbol{x} = B.$$

Marginalising it over the feature space and multiplying by T yields the upper bound of the total cost. Here the constraint is formulated as an equation between the upper

bound of the total cost and the budget because (i) the left part of the constraint is the upper bound of the total cost, which means the total cost should not be higher than the budget; (ii) the upper bound of the total cost should not be lower than the budget, as the left cost could always be utilised to generate value to improve the KPI.

To make the above problem solvable, we consider sequential dependency among the variables for each auction by making the following assumptions:

- Assume $b(\theta(\boldsymbol{x}),\boldsymbol{x}) \equiv b(\theta(\boldsymbol{x}))$. That is: $\boldsymbol{x} \to \theta \to b$. This allows us to largely reduce the functional decision space for the optimisation, while still gaining the dependency of the impression features through the KPI estimation $\theta(\boldsymbol{x})$. This dependency follows the idea of the quantitative method in finance: the bid price for a certain asset only depends on a quantified utility of the asset, which eliminates the dependency from any specific attributes of the asset. The previous work in [19] also adopts a similar dependency (bid only depends on CTR). In auction theory, the bidding is also based on the valuation of the item [109, 114].

- Assume $w(b,\boldsymbol{x}) \equiv w(b)$. That makes the feature $\boldsymbol{x}$ only influence the winning rate via its generated bid: $\boldsymbol{x} \to \theta \to b \to w$. The assumption is sensible as we found out (shown in Section 3.4.1) that the dependency over the bid request features is far less than the bid price. Previous sponsored search bid optimisation work [47, 52] also makes such an assumption on winning keyword ad slots.

With above two assumptions, the optimisation problem is rewritten as

$$b()_{\text{ORTB}} = \underset{b()}{\arg\max} \ T \int_{\boldsymbol{x}} \theta(\boldsymbol{x}) w(b(\theta(\boldsymbol{x}))) p_x(\boldsymbol{x}) d\boldsymbol{x} \tag{3.3}$$

$$\text{subject to} \ T \int_{\boldsymbol{x}} b(\theta(\boldsymbol{x})) w(b(\theta(\boldsymbol{x}))) p_x(\boldsymbol{x}) d\boldsymbol{x} = B.$$

Furthermore, since there is a deterministic relationship between $\boldsymbol{x}$ and $\theta(\boldsymbol{x})$,

$$d\theta(\boldsymbol{x}) = ||\nabla\theta(\boldsymbol{x})|| d\boldsymbol{x}, \tag{3.4}$$

the relationship between their p.d.f. is also determined:[3]

$$p_\theta(\theta(\boldsymbol{x})) = \frac{p_x(\boldsymbol{x})}{||\nabla\theta(\boldsymbol{x})||}. \tag{3.5}$$

Thus we can focus on the KPI variable θ via performing integration by substitution:

$$
\begin{aligned}
\int_{\boldsymbol{x}} \theta(\boldsymbol{x})w(b(\theta(\boldsymbol{x})))p_x(\boldsymbol{x})d\boldsymbol{x} &= \int_{\boldsymbol{x}} \theta(\boldsymbol{x})w(b(\theta(\boldsymbol{x})))p_\theta(\theta(\boldsymbol{x}))||\nabla\theta(\boldsymbol{x})||d\boldsymbol{x} \\
&= \int_{\theta(\boldsymbol{x})} \theta(\boldsymbol{x})w(b(\theta(\boldsymbol{x})))p_\theta(\theta(\boldsymbol{x}))d\theta(\boldsymbol{x}) \\
&= \int_{\theta} \theta w(b(\theta))p_\theta(\theta)d\theta, \tag{3.6}
\end{aligned}
$$

and similarly on the cost upper bound:

$$
\begin{aligned}
\int_{\boldsymbol{x}} b(\theta(\boldsymbol{x}))w(b(\theta(\boldsymbol{x})))p_x(\boldsymbol{x})d\boldsymbol{x} &= \int_{\boldsymbol{x}} b(\theta(\boldsymbol{x}))w(b(\theta(\boldsymbol{x})))p_\theta(\theta(\boldsymbol{x}))||\nabla\theta(\boldsymbol{x})||d\boldsymbol{x} \\
&= \int_{\theta(\boldsymbol{x})} b(\theta(\boldsymbol{x}))w(b(\theta(\boldsymbol{x})))p_\theta(\theta(\boldsymbol{x}))d\theta(\boldsymbol{x}) \\
&= \int_{\theta} b(\theta)w(b(\theta))p_\theta(\theta)d\theta. \tag{3.7}
\end{aligned}
$$

Rewriting the integration w.r.t. θ leads to the final functional optimisation problem as follows:

$$b()_{\text{ORTB}} = \underset{b()}{\arg\max} \; T\int_{\theta} \theta w(b(\theta))p_\theta(\theta)d\theta \tag{3.8}$$

$$\text{subject to} \; T\int_{\theta} b(\theta)w(b(\theta))p_\theta(\theta)d\theta = B.$$

With a reliable KPI estimation model $\theta(\boldsymbol{x})$, the optimisation framework is now built on the KPI distribution of the bid request for a specific campaign: every time a bid request with the utility θ is generated, the bid price $b(\theta)$ is based on the utility θ, and the probability of winning the auction $w(b(\theta))$ and the corresponding cost are based on the bid price.

---

[3]The intuition behind Eq. (3.5) can be illustrated by a linear function $f(x) = ax$, where $a > 0$ and x follows a uniform distribution in $[0,1]$, i.e., $p(x) = 1$, and thus $f(x)$ follows a uniform distribution in $[0,a]$, i.e., $p(f(x)) = 1/a$. Thus it is easy to see the correctness of Eq. (3.5) in this simple example.

3.3 Optimal Solutions of Bidding Functions

The Lagrangian of the objective function (Eq. (3.8)) is

$$\mathscr{L}(b(\theta),\lambda) = \int_\theta \theta w(b(\theta)) p_\theta(\theta) d\theta - \lambda \int_\theta b(\theta) w(b(\theta)) p_\theta(\theta) d\theta + \frac{\lambda B}{T}, \quad (3.9)$$

where λ is the Lagrangian multiplier. Based on calculus of variations, the Euler-Lagrange condition of $b(\theta)$ is

$$\theta p_\theta(\theta) \frac{\partial w(b(\theta))}{\partial b(\theta)} - \lambda p_\theta(\theta) \left[w(b(\theta)) + b(\theta) \frac{\partial w(b(\theta))}{\partial b(\theta)} \right] = 0, \quad (3.10)$$

$$\Rightarrow \quad \lambda w(b(\theta)) = \left[\theta - \lambda b(\theta) \right] \frac{\partial w(b(\theta))}{\partial b(\theta)}, \quad (3.11)$$

where we can see that the KPI probability density $p_\theta(\theta)$ has been eliminated and the *form* of bidding function $b(\theta)$ only depends on the winning function $w(b(\theta))$.[4] This is mainly because both the objective and the constraint take the integration over the distribution of $p_\theta(\theta)$. Different winning functions result in different optimal bidding functions. The winning function should be directly built from bid landscape forecasting [24] based on the market price observations.

Here we present two winning functions which are typical and fit the curves of real-world data. And we derive the optimal bidding function form for each winning function.

3.3.1 Winning & Bidding Function 1

As depicted in Figure 3.5 from our experiment on a real data[5], the winning rate $w(b)$ consistently has an (approximately) concave shape: when the bid price is low, adding a unit bid will increase the winning rate more than when the bid is already high. Thus a simple winning function is in the form of

$$w(b(\theta)) = \frac{b(\theta)}{c + b(\theta)}, \quad (3.12)$$

---

[4]Later we will show that the optimal value of λ depends on $p_\theta(\theta)$, but λ is only a parameter in $b(\theta)$; thus $p_\theta(\theta)$ does not change the general form of $b(\theta)$.

[5]All the price numbers are presented with the unit of CNF (i.e. 0.01 CNY). And the bid price is counted on CPM.

(a) Winning function 1.   (b) Bidding function 1.
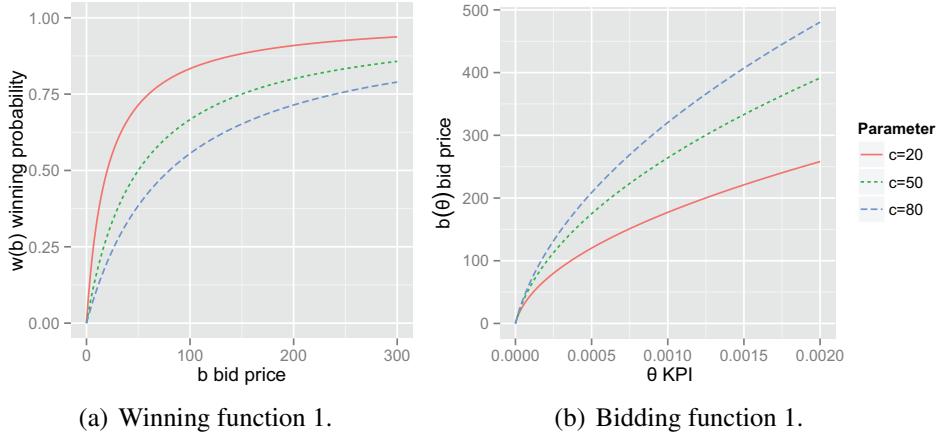
Figure 3.2: Winning function 1 and corresponding optimal bidding function $b_{\mathrm{ORTB_1}}(\theta)$.

where c is a parameter tuned to fit the data. An illustration of the winning function with different c's is given in Figure 3.2(a).

Taking a derivative w.r.t. the bid gives:

$$\frac{\partial w(b(\theta))}{\partial b(\theta)} = \frac{c}{(c+b(\theta))^2}. \tag{3.13}$$

Taking Eq. (3.12) and (3.13) into Eq. (3.11) gives:

$$\frac{\theta c}{(c+b(\theta))^2} - \lambda \left[\frac{b(\theta)}{c+b(\theta)} + c\frac{b(\theta)}{(c+b(\theta))^2} \right] = 0 \tag{3.14}$$

$$\Rightarrow \left(b(\theta)+c \right)^2 = c^2 + \frac{\theta c}{\lambda}. \tag{3.15}$$

Solving the above equation gives the final optimal bidding function:

$$b_{\mathrm{ORTB_1}}(\theta) = \sqrt{\frac{c}{\lambda}\theta + c^2} - c. \tag{3.16}$$

Under the assumption of the winning function 1 in the form of Eq. (3.12), the optimal bidding function $b_{\mathrm{ORTB_1}}(\theta)$ is in a concave form: a square root function form. Figure 3.2(b) gives an illustration of this bidding function with different parameter c's, fixing $\lambda = 5.2 \times 10^{-7}$.

3.3.2 Winning & Bidding Function 2

For some campaigns with competitive targets, or the targeted publishers/SSPs setting a high reserve price, the winning probability will not increase rapidly when the bid

(a) Winning function 2. (b) Bidding function 2.
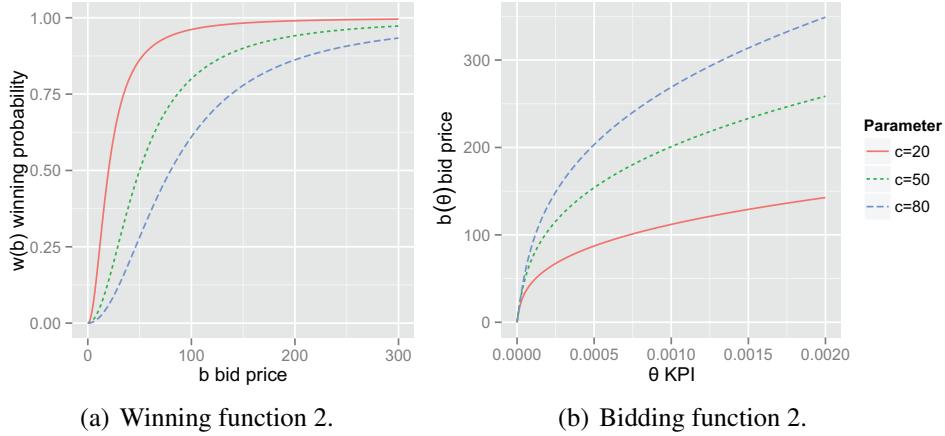
Figure 3.3: Winning function 2 and corresponding optimal bidding function $b_{\mathrm{ORTB}_2}(\theta)$.

price is around 0; only after the bid price becomes larger than some non-zero value the winning probability starts to dramatically increase. Such a case usually occurs in high-profile ad slots [52]. To get this feature, we propose an alternative winning function:

$$w(b(\theta)) = \frac{b^2(\theta)}{c^2 + b^2(\theta)}, \tag{3.17}$$

$$\frac{\partial w(b(\theta))}{\partial b(\theta)} = \frac{2b(\theta)c^2}{(b(\theta) + c)^2}, \tag{3.18}$$

where the parameter c controls the increasing point of the curve[6]. An illustration is given in Figure 3.3(a).

Following the same token in Section 3.3.1, we solve Eq. (3.11) using the winning function in Eq. (3.17) and its derivative in Eq. (3.18) to get the second optimal bidding function:

$$b_{\mathrm{ORTB}_2}(\theta) = c \cdot \left[\left(\frac{\theta + \sqrt{c^2\lambda^2 + \theta^2}}{c\lambda} \right)^{\frac{1}{3}} - \left(\frac{c\lambda}{\theta + \sqrt{c^2\lambda^2 + \theta^2}} \right)^{\frac{1}{3}} \right]. \tag{3.19}$$

Fixing $\lambda = 5.2 \times 10^{-7}$, the bidding functions with different c's are shown in Figure 3.3(b). Again the $b_{\mathrm{ORTB}_2}(\theta)$ is a concave function.

The proposed optimisation framework is a general one: Eq. (3.11) shows that different winning functions would lead to different optimal bidding functions. The framework can adapt to various ad markets with different winning functions (bid landscapes).

---

[6]Actually we can take a more general form of the winning function: $w(b(\theta)) = b^\alpha(\theta)/(c^\alpha + b^\alpha(\theta))$. We investigate the case of $\alpha = 1, 2$ in this research. When α is larger than 2, there is no analytic solution of $b_{\mathrm{ORTB}}(\theta)$.
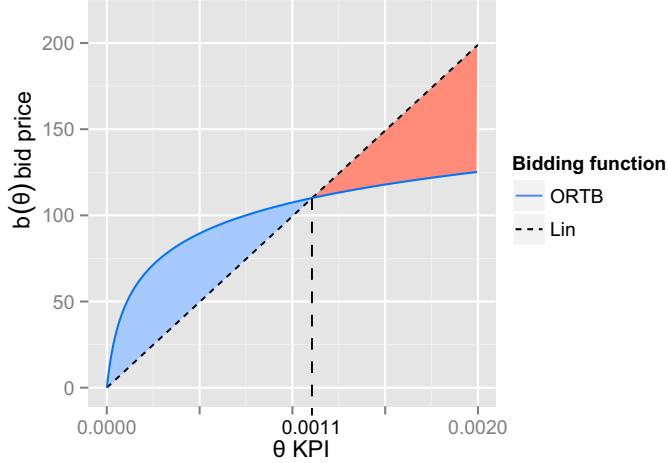
Figure 3.4: Concave bidding function bids higher on the cases with predicted CTR less than 0.0011 (example) and vice versa, which allocates more budget on the cases with low bid price and thus low cost.

Here we estimate the winning functions from real data (Figure 3.5) and limit our study to the RTB markets only.

3.3.3   Discussions on Derived Bidding Functions

Unlike the linear form bidding function in the previous study [19, 4] (denoted as LIN), the derived bidding functions in this research (denoted as ORTB) Eq. (3.16) and (3.19) suggest a non-linear concave form mapping from predicted KPI to the bid value under a budget constraint for RTB. As shown in Figure 3.4, compared with LIN, ORTB bids higher when the estimated KPI is low, which means ORTB allocates more budget on the low reward and low cost cases.

The strategy of bidding more on cheap impressions comes from the shape of the winning functions. In Figure 3.5 we can find that for all the campaigns, when the bid price increases from zero, the winning probability will have a high growth rate first and after the bid price surpasses a region, the winning probability starts to converge to 1. As such, the strategy of ORTB will earn much higher winning probability while only increase a little cost upperbound because of the concavity of winning rate w.r.t. the bid price.

3.3.4   Optimal Solution of λ

The bidding functions in Eq. (3.16) and (3.19) also take λ as the parameter. Denoting the bidding function explicitly with λ as $b(\theta, \lambda)$. To calculate the optimal λ, the Euler-

Lagrange condition of λ from Eq. (3.9) is

$$\frac{\partial \mathscr{L}(f,\lambda)}{\partial \lambda} = 0$$

$$\Rightarrow \int_{\theta} b(\theta,\lambda) w(b(\theta,\lambda)) p_{\theta}(\theta) d\theta = \frac{B}{T}. \tag{3.20}$$

Given the formula of $b(\theta,\lambda)$, the solution λ can be found. However, in many cases such as using winning functions from Eqs. (3.12) and (3.17), there is no analytic solution of λ. Also we can see the solution depends on the probability density of predicted KPI $p_{\theta}(\theta)$. Alternatively, one can find the numeric solution using the bidding log data and practically solve it using efficient numeric calculation.

Here we take a rather pragmatic approach by regarding λ as a tuning parameter for the bidding functions and learn it from the data. Eq. (3.20) can be rewritten as

$$\int_{\theta} \left(b(\theta,\lambda) w(b(\theta,\lambda)) - \frac{B}{T} \right) p_{\theta}(\theta) d\theta = 0, \tag{3.21}$$

which can be solved with a tighter but practically feasible minimisation problem

$$\min_{\lambda} \int_{\theta} \frac{1}{2} \left(b(\theta,\lambda) w(b(\theta,\lambda)) - \frac{B}{T} \right)^2 p_{\theta}(\theta) d\theta. \tag{3.22}$$

If there are a very large number N of observations of θ for the optimised campaign, Eq. (3.22) can be reasonably approximated as

$$\min_{\lambda} \sum_{k=1}^{N} \frac{1}{2} \left(b(\theta_k,\lambda) w(b(\theta_k,\lambda)) - \frac{B}{T} \right)^2, \tag{3.23}$$

where θ_k is the predicted CTR of the kth data instance in the training data. The techniques like (mini-)batch descent or stochastic gradient descent can be used to solve λ by the following iteration:

$$\lambda \leftarrow \lambda - \eta \sum_{k=1}^{N} \left(b(\theta_k,\lambda) w(b(\theta_k,\lambda)) - \frac{B}{T} \right) \cdot \tag{3.24}$$

$$\left(\frac{\partial b(\theta_k,\lambda)}{\partial \lambda} w(b(\theta_k,\lambda)) + b(\theta_k,\lambda) \frac{\partial w(b(\theta_k,\lambda))}{\partial \lambda} \right),$$

until convergence. Usually, just like Eqs. (3.16) and (3.19), as $b(\theta,\lambda)$ has a mono-

tonic relationship with λ and $w(b(\theta, \lambda))$ monotonically increases against $b(\theta, \lambda)$, $b(\theta_k, \lambda)w(b(\theta_k, \lambda))$ has a monotonic relationship with λ. For example, with the bidding function as Eq. (3.16) and the winning function as Eq. (3.12), the factor $b(\theta_k, \lambda)w(b(\theta_k, \lambda))$ decreases monotonically against λ, which makes it quite easy to find the optimal solution. Therefore, with larger per-case budget B/T, the solution of λ becomes smaller, which corresponds to a higher bid price. The experiment will demonstrate the trend of optimal λ corresponding to different per-case budget B/T (Figures 3.11 and 3.12).

3.3.5   Special Case Discussion: Linear Winning Function

It is interesting to show that our proposed functional optimisation framework is able to derive the widely adopted linear bidding functions [19, 4] with some special setting of winning functions.

Suppose the winning function is linear w.r.t. the bid price in the interval $[0, c]$, with 0 winning probability at 0 bid and with 1 winning probability with bid price no lower than a constant parameter c:

$$w_{\text{LIN}}(b) = \frac{b}{c}. \tag{3.25}$$

The taking Eq. (3.25) into Eq. (3.11) leads to

$$\lambda \frac{b(\theta)}{c} = \left[\theta - \lambda b(\theta)\right] \frac{1}{c} \tag{3.26}$$

$$\Rightarrow b(\theta) = \frac{\theta}{2\lambda}, \tag{3.27}$$

where the derived optimal bidding function is linear w.r.t. to the predicted KPI. The derivation of λ solution follows Section 3.3.4, with the bidding function as in Eq. (3.27).

$$\int_\theta \frac{\theta}{2\lambda} \cdot \frac{\theta}{2\lambda c} p_\theta(\theta) d\theta = \frac{B}{T} \tag{3.28}$$

$$\Rightarrow \lambda = \frac{1}{2}\sqrt{\frac{T}{Bc} \int_\theta \theta^2 p_\theta(\theta) d\theta} \tag{3.29}$$

Thus the analytic solution of the optimal linear bidding function is

$$b_{\mathrm{LIN}}(\theta) = \frac{\theta}{\sqrt{\frac{T}{Bc} \int_{\vartheta} \vartheta^2 p_{\theta}(\vartheta) d\vartheta}}, \tag{3.30}$$

where the term $\int_{\vartheta} \vartheta^2 p_{\theta}(\vartheta) d\vartheta$ is determined by the overall distribution of predicted KPI, which can be directly estimated from the training data. The per auction budget term B/T determined the overall bidding scale. The higher B/T, the higher bid scale, which is reasonable. The parameter c indicates the market competitiveness: the higher c, the lower winning probability given a bid price b, i.e., the higher market competitiveness, thus the higher bid should be performed in order to win impressions and exhaust the budget.

From above derivation, we know that under the assumption of linear winning function the derived optimal bidding function is with the linear form w.r.t. the predicted KPI, which is widely used in industry. However, as from Figure 3.5 we can see the winning functions shown by the real-world data is always non-linear, which challenges the effectiveness of the linear bidding functions. In the experiment part, we will extensively investigate the comparison between our proposed non-linear bidding functions $b_{\mathrm{ORTB_1}}(\theta)$, $b_{\mathrm{ORTB_2}}(\theta)$ and the linear bidding function $b_{\mathrm{LIN}}(\theta)$.

3.4 Experimental Setup

The derived bidding functions are tested both by offline evaluation (Section 3.5) using a real-world dataset and via online A/B testing (Section 3.6) over a commercial DSP with real advertisers and impressions. In this section, we introduce the experiment setup and report the results from our data analysis.

3.4.1 Dataset and Analysis

iPinYou dataset description. We use the real-world bidding feedback log from iPinYou as our dataset[7]. It records more than 64.8 million bid requests and 15.4 million impressions and the user feedback of 9 campaigns from different advertisers during 10 days in 2013. The dataset disk size is 35GB. For each bid request, the log contains

---

[7]iPinYou is a mainstream DSP company in China. The dataset comes from the global RTB algorithm competition by iPinYou [115] and has been publicly released for research on our website: http://data.computational-advertising.org.

Table 3.2: iPinYou dataset statistics. Here CNY means Chinese Yuan, while CNF means Chinese Fen, which is 0.01 CNY.

Camp. ID	Bids (M)	Imps (K)	Clicks	Convs	Cost (CNY)	AWR	CTR	CPM (CNF)	eCPC (CNF)
1458	14.70	3,083.1	2,454	1	2,124.0	20.97%	0.080%	68.89	86.55
2259	2.99	835.6	280	89	777.5	27.97%	0.034%	93.06	277.70
2261	2.16	687.6	207	0	616.1	31.84%	0.030%	89.6	297.64
2821	5.29	1,322.6	843	450	1,180.8	24.99%	0.064%	89.28	140.07
2997	1.02	312.4	1,386	0	196.9	30.69%	0.444%	63.02	14.21
3358	3.75	1,742.1	1,358	369	1,609.4	46.44%	0.078%	92.38	118.51
3386	14.09	2,847.8	2,076	0	2,190.7	20.21%	0.073%	76.92	105.52
3427	14.03	2,593.8	1,926	0	2,102.4	18.48%	0.074%	81.06	109.16
3476	6.71	1,970.4	1,027	26	1,560.9	29.35%	0.052%	79.22	151.98
Total	64.75	15,395.3	11,557	935	12,358.8	23.78%	0.075%	80.28	106.94

the information from the user (e.g. the user segmentation), advertiser (e.g. the creative format and size), publisher (e.g. the auction reserve price, ad slot, page domain and URL) and the context (e.g. the time, region, the browser and operation system). For each bid request, there is an auction winning bid price, and the user feedback (click, conversion) is recorded if the campaign won the auction. More details of the dataset is shown in Table 3.2 and the references [115, 29].

According to the data publisher [115], the last 3-day for each campaign is split into the test dataset while the data in the previous period is the training data. The training data is mainly used to train the CTR estimator and tune the bidding function parameters. The test data is used to evaluate the compared DSP bidding strategies.

Data analysis on winning prices. Figure 3.5 depicts the auction winning rate w.r.t. the bid price for all 9 campaigns. It can be observed that the data of all the campaigns follows a similar pattern: as the bid price increases, the campaign's auction winning rate increases dramatically; when the bid price gets larger (e.g. more than 100), the increasing of winning rate slows down and finally the winning rate converges to 1. Thus, it is much more reasonable to employ concave functions like Eq. (3.12) and (3.17) to model such relationships than linear winning functions like Eq. (3.25). For each campaign, we fit the winning functions with the parameter c leading to the least square error with the real curve.

Next, we study the dependency between the bid request features and the winning bid prices (also named as market prices in [28]). Figure 3.6 gives the box plots [116] of
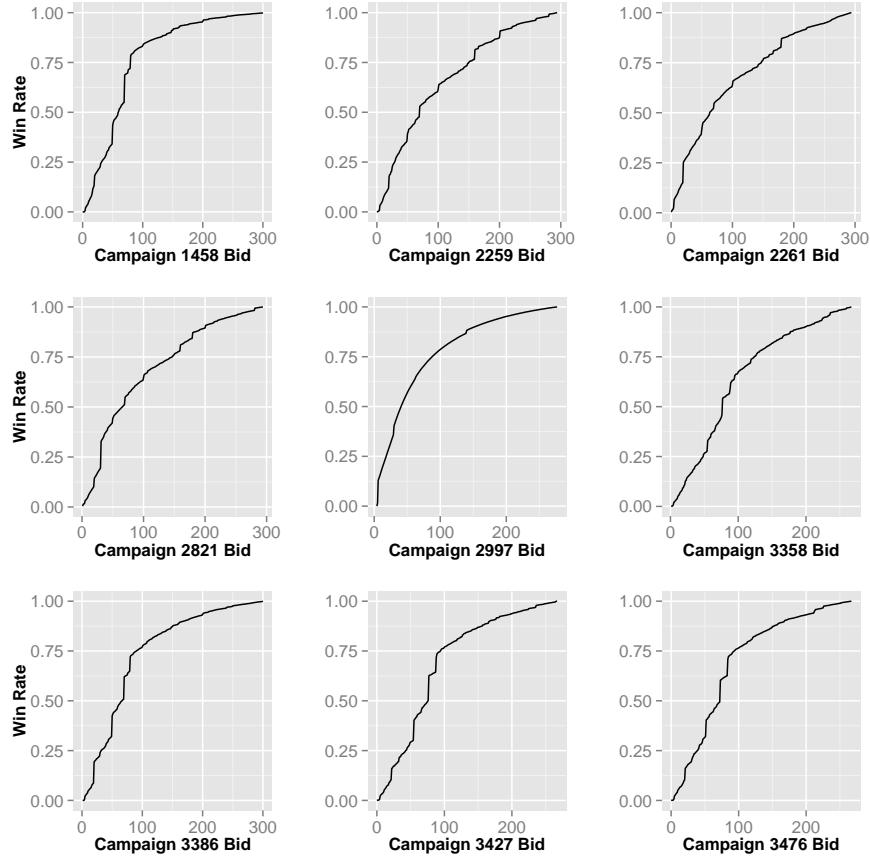
Figure 3.5: Relationship between auction winning rate and bid value for different campaigns. There may be some "angles" in the winning functions of some campaigns. The reason is there are some "needles" in market price p.d.f. because of the constant bidding by some advertisers.

winning price distribution against the features such as hour, weekday, user browser, operation system and location regions of bid requests to campaign 1458 (other campaigns follow the similar pattern). Compared with the clear changing relationship with the bid price shown in Figure 3.5, Figure 3.6 shows that the winning price distributions do not have obvious dependency on the categorical feature value. It suggests that the bid price is the key factor influencing the campaign's winning rate in its auctions. Once we have known the bid value, the auction winning rate is less sensitive to the other bid request features. Thus it is practically reasonable to simplify $w(b, \boldsymbol{x}) \equiv w(b)$ as proposed in Section 3.2.

3.4.2 Evaluation Measures

The task of a DSP is to optimise each campaign's KPI (such as clicks, conversions etc.) given the budget. Therefore, the KPI is the primary evaluation measure of the bidding
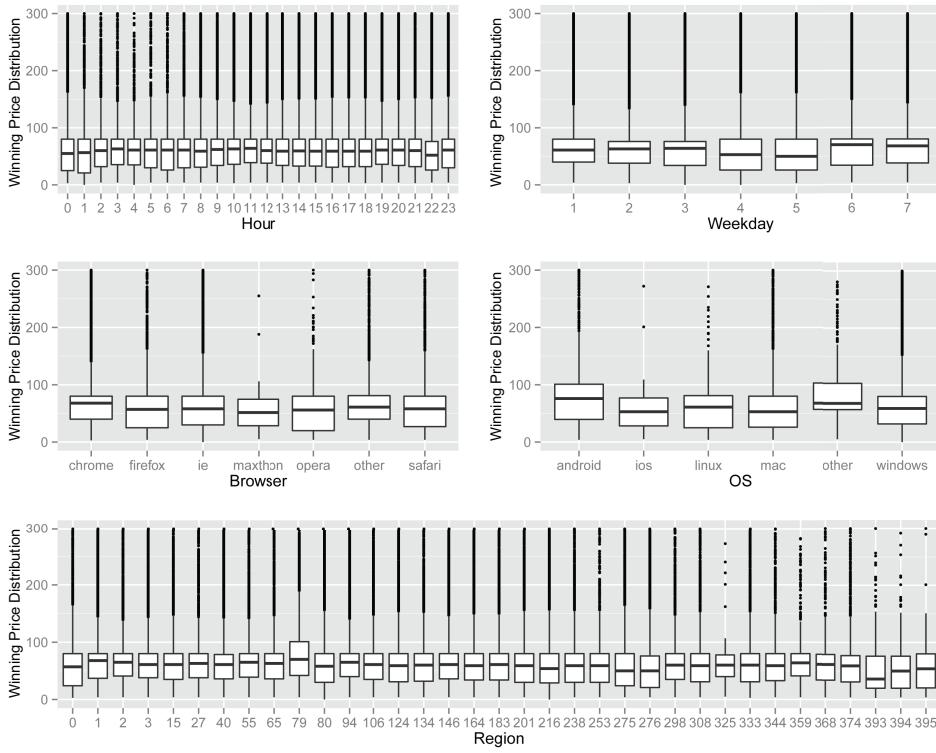
Figure 3.6: Winning price distribution against different features for campaign 1458. Note that these are the overall market price distributions w.r.t. different features, rather than individual cases.

strategies in the experiment. Specifically, the campaign acquired click number is set as the primary KPI in the experiment, while other statistics such as CPM, and eCPC are also monitored and reported. In addition, the optimisation on an alternative KPI, i.e. the combination of achieved clicks and conversions, is also investigated, as will be discussed in Section 3.5.5.

3.4.3   KPI Estimator Training

For each campaign, we use its impression/click/conversion log data to train a KPI estimator for each bid request. Particularly, if the KPI is click, then this task turns to be the well-known CTR estimation [94, 1]. Regression models such as random forest and gradient boosting regression trees can be applied here. Since this work mainly focuses on the bidding strategy instead of the KPI estimator model, the logistic regression is applied as the CTR estimator as it is a widely used choice [1]. The loss is the cross entropy between the predicted click probability and the ground-truth result. In addition, L2 regularisation is used. For an alternative KPI in the experiment, the KPI estimator training will be discussed in Section 3.5.5.
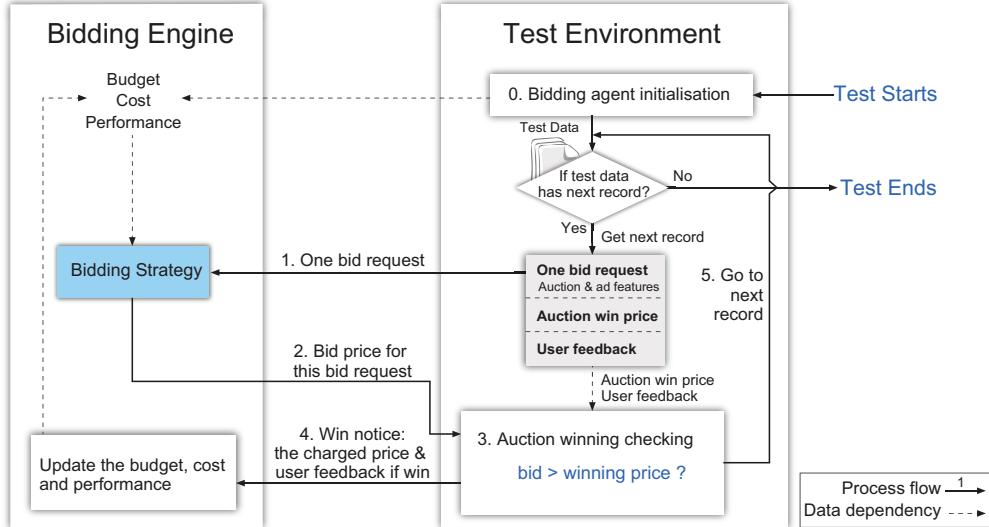
Figure 3.7: Evaluation flow chart.

Features are extracted from the log data to train the CTR estimator model. Specifically, we extract 29,408 first-order binary features and based on that generate 657,756 second-order binary features, which yields the total of 687,164 features for our training.[8]

3.4.4 Test Evaluation Setup

Evaluation flow. The evaluation flow is depicted in Figure 3.7. Given the bidding strategy and a budget for the test period for a particular campaign, the offline evaluation can be conducted via going through its test data. The test data is a list of records. Each record consists of the features of one bid request, the auction winning price and the user feedback information. Specifically, receiving the bid request features of each record by the timestamp, the bidding strategy generates a bid price for it (if the cost is beyond the budget, just returns 0, i.e. skips the remaining bid requests). If the bid price is higher than the auction winning price of this record, the campaign wins the auction, gets its ad shown. The corresponding user feedback (click) and the corresponding charged price of the record are then referenced to update the performance and cost. After that, if there is no more bid requests in the test data, then the evaluation is over, with the final performance returned.[9]

It is worth mentioning that using user feedback logs for offline evaluation has limi-

---

[8]Normally, each feature index does not need to be updated but the whole feature index dictionary needs to augment as there are usually more and more new features occurs over time.

[9]The code of evaluation framework: `https://github.com/wnzhang/optimal-rtb`.

tation. In the case of RTB offline evaluation, user feedback only occurs for the winning auctions (having ad impressions); there is no user feedback for the losing bids. We thus do not know whether the user will click or convert even if we bid enough high to win the originally losing ad auction in our offline experiment. In this thesis, we follow the convention of the offline evaluations from sponsored search [94], recommender systems [117] and web search [118] that the objects (auctions) with unseen user feedback are ignored (i.e., considered as auction failure cases). To complement with the offline evaluation, we will further show the online test performance on a commercial DSP in Section 3.6. Note that such a problem will be focused and mathematically solved in Chapter 5.

Budget constraints. It is easy to see that if the budget is set as the same as the original total cost in the test log, then just simply bidding as high as possible for each case will exactly run out the budget and get all the logged clicks. In the experiment, to test the performance against various budget constraints, for each campaign, the evaluation test uses 1/64, 1/32, 1/16, 1/8, 1/4 and 1/2 of the original total cost respectively in the test log as the budget. Likewise, in the training stage, the budget used is 1/64, 1/32, 1/16, 1/8, 1/4 and 1/2 of the original total cost of the training data respectively.

3.4.5 Compared Bidding Strategies

The following baseline and state-of-the-art bidding strategies are compared in the experiment. The parameters of each bidding strategy are tuned based on the training data.

Constant bidding (CONST). The bidding agent bids a constant value for all the bid requests to the campaign. The parameter is the specific constant bid price.

Random bidding (RAND). The bidding agent randomly chooses a bid value in a given range. The parameter is the upper bound of the random bidding range.

Bidding below max eCPC (MCPC). As discussed in [4], given the advertiser's goal on max eCPC, i.e., the upper bound of effective cost per click, the bid price for a bid request on an impression is obtained by multiplying the max eCPC and its predicted CTR. Here the max eCPC for each campaign is calculated by dividing its total cost with achieved number of clicks in the training data. No parameter for this bidding strategy.

Table 3.3: Bidding strategy attributes.

Bidding Strategies	CONST	RAND	MCPC	LIN	ORTB
Consider budget conditions	$\checkmark$	$\checkmark$		$\checkmark$	$\checkmark$
Evaluate per impression value			$\checkmark$	$\checkmark$	$\checkmark$
Consider winning function					$\checkmark$

Linear-form bidding w.r.t. predicted CTR (LIN). In the previous work [19], the bidding agent bids the value linearly proportional to the predicted CTR for each bid request. The formula can be generally written as

$$b_{\mathrm{LIN}}(\theta) = b_0 \frac{\theta}{\theta_0}, \tag{3.31}$$

where θ_0 is the average predicted CTR under a target condition (e.g., a user-inventory pair or a campaign's volume) and b_0 is the basic bid price for this target condition. b_0 is the parameter to tune in the experiment. Note that this bidding function is a generalisation of our derived linear bidding function in Eq. (3.30) as the bidding scale can be tuned.

Optimal real-time bidding (ORTB$_1$ and ORTB$_2$). These are the derived optimal bidding strategies in the proposed framework, as shown in Eq. (3.16) and (3.19). The parameters are c and λ for both bidding strategies, where c is obtained by fitting the winning probability and λ is tuned using the training data.

Table 3.3 summarises the attributes of the different strategies. MCPC is not budget-aware to make sure it spends all the budget. MCPC, LIN and ORTB perform bidding based on impression-level evaluation. Taking winning functions into account, ORTB is the most informative strategy. In Section 3.5 we will analyse the impact of these attributes in the final performance.

3.5 Offline Empirical Study

From the offline empirical study, the following questions are to be answered. (i) Does the derived non-linear bidding function outperform the state-of-the-art linear one? (ii) What are the characteristics of the proposed ORTB algorithms and how do the parameters impact the performance and what are their relationships with the budget conditions?
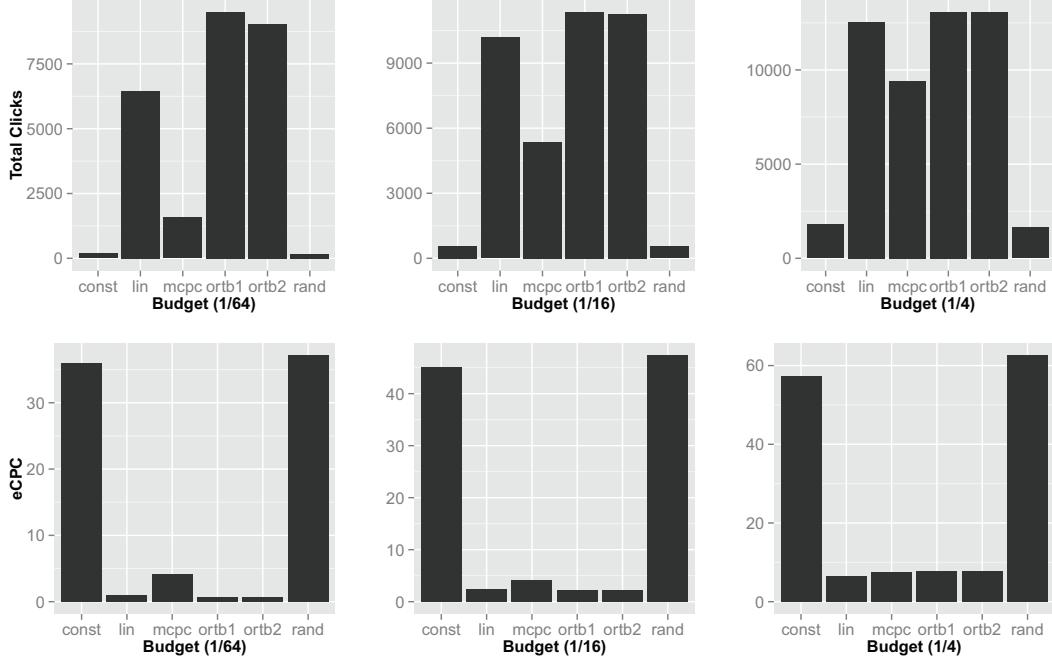
Figure 3.8: Overall performance comparison. The higher clicks and lower eCPC, the better.

3.5.1 Performance Comparison

The performance comparison on total clicks and eCPC under different budget conditions are reported in Figure 3.8. It can be observed that (i) under every budget condition, the proposed bidding strategies ORTB$_1$ and ORTB$_2$ lead the best performance on total clicks, which verifies the effectiveness of the derived non-linear forms of the bidding functions. (ii) Except ORTB, LIN is the best algorithm in the comparison. This algorithm represents the widely used DSP bidding strategies in industry [19]. (iii) MCPC is aware of the upper bound cost for each bid request, and dynamically changes its bid according to the estimated CTR. However, compared to ORTB and LIN, MCPC has no adaptability to different budget conditions. For example, when the budget is relatively low for the bid request volume, MCPC will still bid based on the originally set max eCPC, while ORTB and LIN can adaptively lower the bid to earn the impressions and clicks with higher ROI. (iv) RAND and CONST provide very low performance even though their parameters are tuned under different budget conditions. (v) Also from the eCPC performance we can see RAND and CONST spend much more money to get one click than the case-value-aware strategy MCPC, LIN and ORTB. The last two points suggest the benefit of real-time bidding based display advertising: evaluating the value for each bid request (impression level) plays a significant role in the performance.

Table 3.4: Click improvement of ORTB$_1$ over LIN for each campaign under different budget conditions.

Camp.	1/64	1/32	1/16	1/8	1/4	1/2
1458	0.68%	1.97%	-0.46%	1.04%	1.25%	0.26%
2259	84.07%	49.89%	9.13%	-1.99%	5.67%	3.51%
2261	85.67%	51.10%	18.50%	6.27%	2.67%	0.00%
2821	58.13%	23.78%	8.82%	1.93%	0.59%	0.65%
2997	428.26%	118.78%	103.91%	63.68%	18.82%	6.50%
3358	0.70%	0.00%	0.44%	3.38%	0.48%	0.90%
3385	27.16%	2.33%	1.45%	0.43%	6.50%	1.11%
3427	24.46%	-1.19%	3.11%	2.84%	0.06%	3.14%
3476	49.54%	10.85%	0.33%	5.17%	3.82%	1.42%
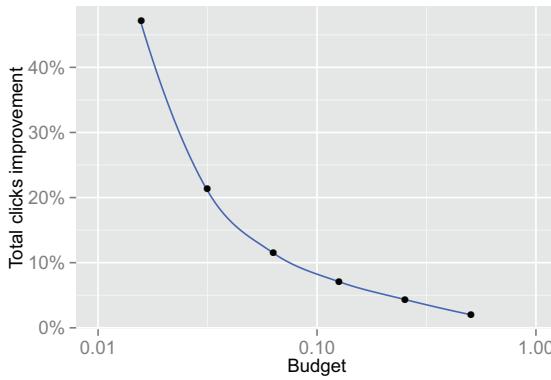
Figure 3.9: Overall click improvement of ORTB$_1$ over LIN under different budget conditions.

Table 3.4 gives detailed performance improvement on total clicks of ORTB$_1$ over LIN under different campaigns and budget conditions. Among the listed 54 settings, ORTB$_1$ wins LIN in 49 (90.7%) settings, ties in 2 (3.7%) settings, and loses in 3 (5.6%) settings. This shows ORTB$_1$ is fairly robust and the outperformance is stable.

3.5.2 The Impact of Budget Constraints

It is of great interest to investigate how the bidding strategy adapts the changing of the budget constraints. In the experiment, the test budget is set as 1/64, 1/32, 1/16, 1/8, 1/4 and 1/2 of the original total cost in the history log respectively. Figure 3.9 depicts the percentage improvement on total clicks of ORTB$_1$ over LIN w.r.t. the budget constraints. As can be observed, (i) when the budget is quite low (e.g. 1/64 of the original total cost), the click improvement of ORTB$_1$ over LIN is quite high (more than 45%). This indicates that the proposed bidding strategy performs particularly well under very limited budget conditions. Intuitively, when the budget is quite low (budget on per bid request B/T is low), a good bidding strategy should spend relatively low

on each bid request. Compared with the linear LIN, ORTB$_1$ allocates more budget on the low cost cases due to the concave form of the bidding function (see Figure 3.2(b)). This is sensible because from the winning rate functions in Figure 3.5 we know that lowering the high bid actually does not drop the winning probability too much. By contrast, highering the low bid a little will increase the winning probability a lot. (ii) When the test budget gets higher, the improvement percentage gets lower. This is reasonable: when there is more budget per bid request, the strategy will appropriately reallocate budget from the low cost cases to high cost cases because the high cost cases also mean high value (CTR). Thus the concave degree of the curve in Figure 3.2(b) will be lower. The curve will relatively approximate to (but not fully change to) the linear form. An extreme case is that when the test budget is set the same as the original total cost in the test log, the improvement is zero. This is because under such a condition for every bidding strategy (except MCPC), just bidding as high as possible will exactly run out the budget and get every impression and click in the test log.

3.5.3   Clicks vs. Impressions

Figures 3.10(a) and 3.10(b) show the total clicks and eCPC of each bidding strategy. It can be observed that both click number and eCPC increase as the budget increases. This is reasonable because when the budget is low, the optimal setting for each bidding strategy will push the budget to the low cost cases, while the budget is higher, the left budget will be allocated on the relative expensive cases (but with lower ROI), which will higher the eCPC. MCPC is unaware of the budget condition, and its slight eCPC fluctuation purely depends on the data. Another point is that when the budget is set very low (1/64, 1/32, 1/16 of the original spend), the eCPC of LIN and ORTB is lower than MCPC, and when the budget increases over 1/4 of the original spend, the eCPC of LIN and ORTB starts to be higher than MCPC.

Figures 3.10(c) and 3.10(d) plot the total impressions and CPM of each bidding strategy. It can be observed that while ORTB strategies generate the highest clicks, they also produce comparable numbers of impressions against the others. This certainly benefits advertisers who aim to maximise their traffic (clicks) while still want to maintain a good exposure (impressions).

(a) Click number against budget

(b) eCPC against budget

(c) Impression number against budget

(d) CPM against budget
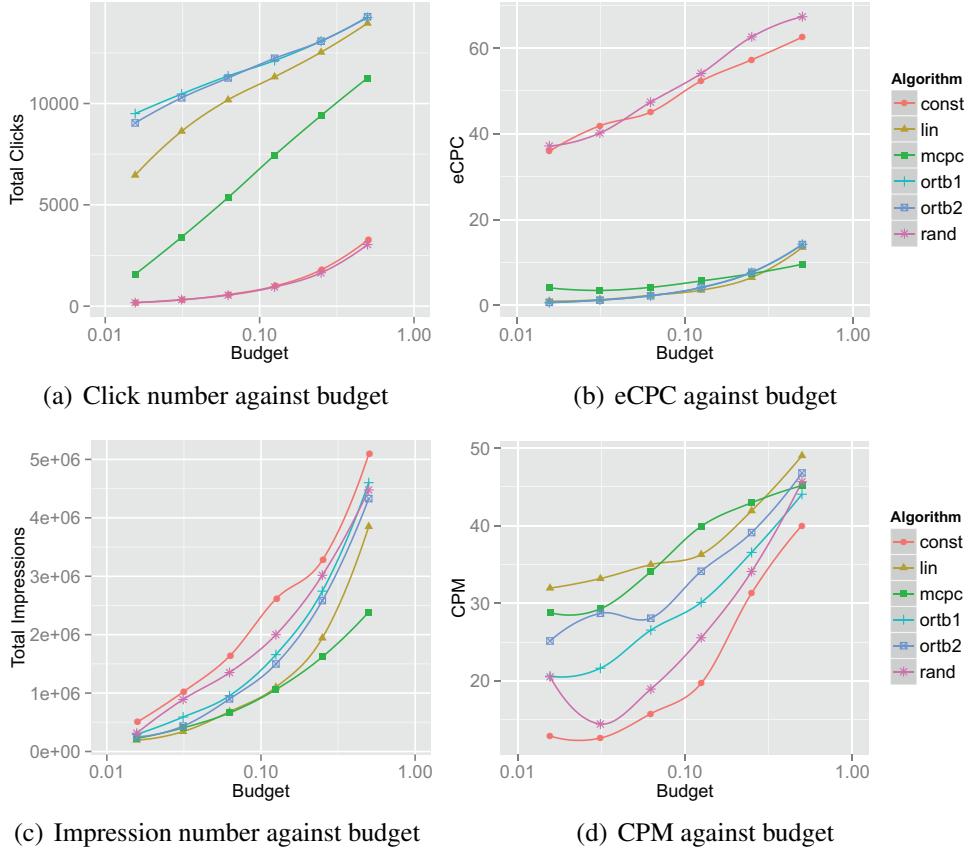
Figure 3.10: Performance on different measures with different budget conditions.

3.5.4 Parameter Tuning

As explained previously, although parameter λ can be directly solved numerically, for efficiency, in the experiment, we still investigate the performance against different value of λ to understand its influences. Figures 3.11 and 3.12 show the corresponding ORTB$_1$ and ORTB$_2$ click performance for the campaign 1458[10] when tuning its parameter λ. For each λ, we also try different values of the other parameter c, which makes multiple points to each x-value for volatility checking. However, practically c is directly learned via best fitting the auction winning rate data for each campaign. From Figure 3.11 and 3.12, we can see that when the given budget is low, the optimal value of λ is high. This is very reasonable. From Eq. (3.16) and (3.19), we can see that the parameter λ controls the general scale of the bidding price: when λ is higher, the bid price gets lower. Thus it is reasonable that when the budget is more limited, the general bidding price level should get lower, which corresponds to the higher optimal value of λ.

---

[10]The trends on other campaigns are quite similar.
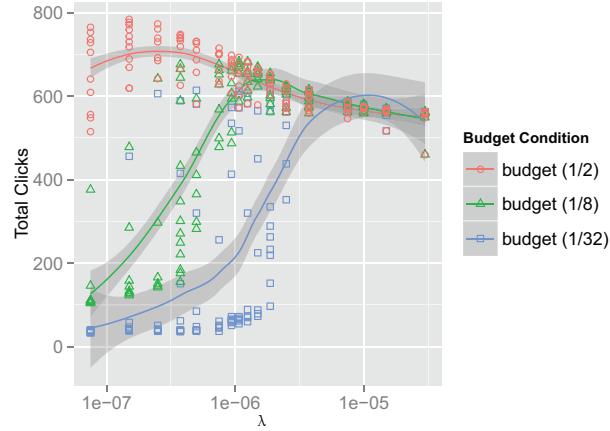
Figure 3.11: Parameter tuning on λ of ORTB$_1$ under different budget conditions.
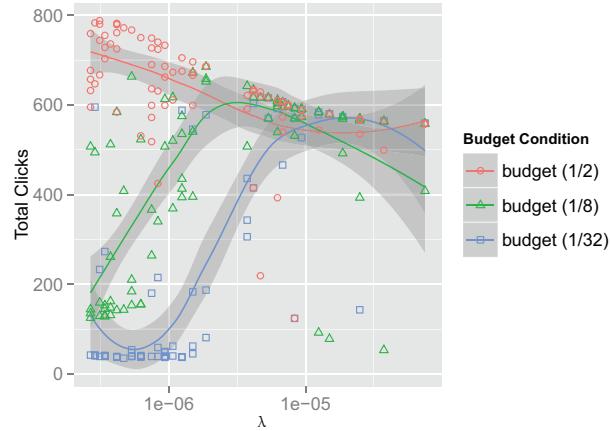
Figure 3.12: Parameter tuning on λ of ORTB$_2$ under different budget conditions.

3.5.5   Results for an Alternative KPI

As discussed in Section 3.2, the optimisation framework is flexible to incorporate different KPIs as optimisation target. Apart from the number of clicks as the main objective studied in this chapter, we also test an alternative KPI as the target by considering conversions, namely, a linear combination of the click number and conversion number, with the parameter k controlling the importance of conversion:

$$\text{KPI} = \#\text{click} + k \cdot \#\text{conversion}. \tag{3.32}$$

This objective is practically useful [119] since the conversion is a quite important measure for performance campaigns [19], and such a linear combination can alleviate the sparsity problem of conversion counting [4]. In the studied iPinYou dataset, only campaign 2821 and 3358 have sufficient conversion records. Thus we choose them as our optimising campaigns and we set $k = 5$ in our experiment. Specifically, two lo-
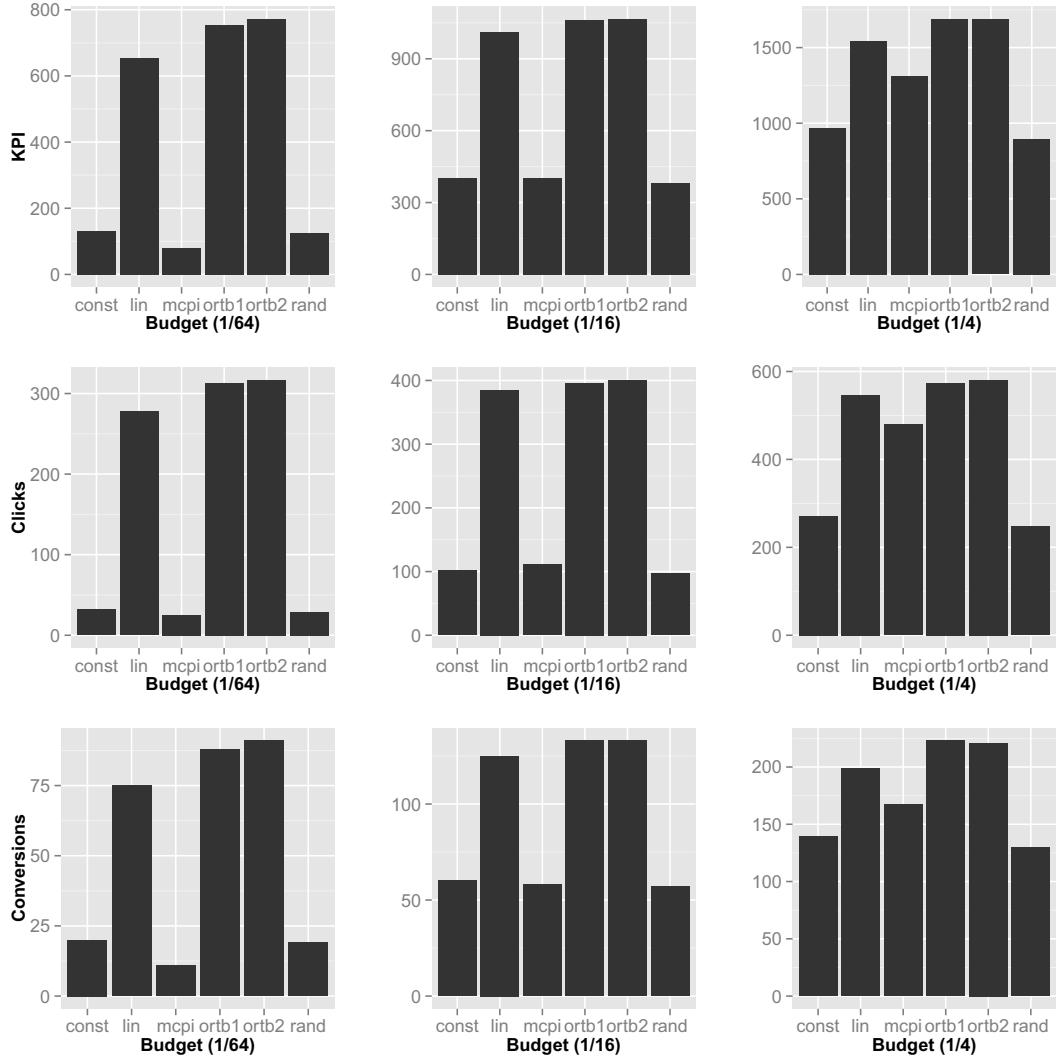
Figure 3.13: Performance comparison with the new KPI (Eq. (3.32)) as optimisation target.

gistic regression models are trained to learn and predict the CTR and CVR (pCTR and pCVR) for each bid request and the predicted KPI (pKPI) can be calculated by $\text{pKPI} = \text{pCTR} + k \cdot \text{pCVR}$, which is the value of θ in the bidding function.

Figure 3.13 gives the overall KPI performance and the specific clicks/conversions obtained by each bidding strategy[11]. From the results we see that ORTB strategies still outperform other compared bidding strategies, which verifies the effectiveness of ORTB on the alternative KPI optimisation. Particularly, on 1/64 budget condition, ORTB$_2$ achieves higher KPI and conversion numbers than ORTB$_1$, this is mainly because the winning function 2 in Eq. (3.17) fits these two optimised campaigns better than winning function 1 in Eq. (3.12).

---

[11]Here MCPC is renamed as MCPI (i.e. bidding under max cost-per-unit-of-KPI).
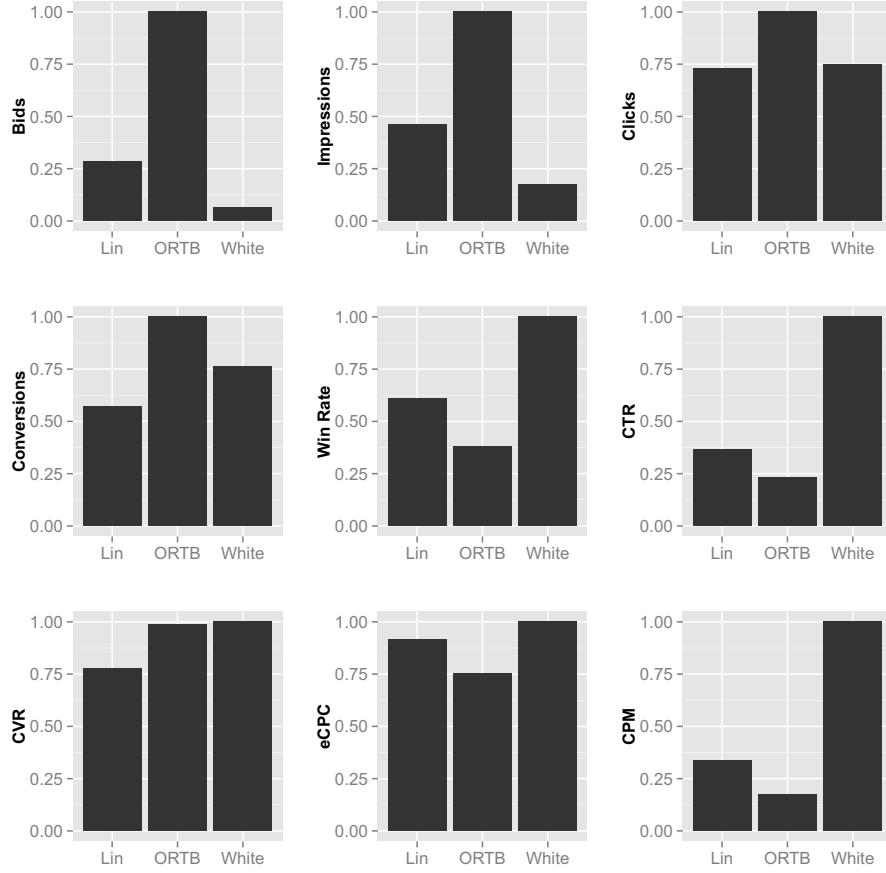
Figure 3.14: Relative performance for online test.

3.6   Online A/B Test

With the strategy of emphasising more on lower CPM impressions than the linear strategy, we conducted an online experiment on iPinYou Optimus platform, which is currently the largest DSP in China. We tested on three campaigns during consecutive three days in December 2013. Another two compared bidding strategies are: (i) White list (WHITE): the bidding agent keeps a list of feature combination rules and bids a high constant value only when the case satisfies any of the white list rules; (ii) LIN, as discussed before. The target KPI in the optimisation is click. Compared with LIN, WHITE acts like the strategy with a step function in [19]: when the predicted CTR is higher than a threshold (matching the white list), the bidding agent bids a much high value, otherwise bids 0. For each bid request, each of the three algorithms has the equal possibility to perform the bidding of the bid request that iPinYou DSP receives. The performance comparison with various measures is reported in Figure 3.14. Due to the company's policy about data sensitivity, we follow [52] to only present the relative performance here.

From the comparison we can have following observations. (i) ORTB bids much more cases than the other two algorithms (bids lower than the auction reserve price are not counted), and gets the most impressions, clicks and conversions. Also ORTB achieves the lowest eCPC, which indicates it is also the most cost effective algorithm. (ii) ORTB obtains the lowest CPM. These two points show ORTB allocates more budget to the cheap cases. As a result, ORTB bids more cases with lower CPM. (iii) Because of the low CPM on low predicted CTR cases, ORTB has the lowest auction winning rate and CTR, but this is not that important since the optimisation target is the total click number. (iv) WHITE acts oppositely compared to ORTB. It only bids high on a subset of cases, which results in its low bidding number and high CPM. Because of the carefully crafted white list, the cases matching the white list do have high CTR. (v) LIN plays moderately, which is not surprising. In sum, the online test demonstrates the effectiveness of the proposed bidding strategies which allocate more budget on the cheap cases to perform more bids with lower CPM.

3.7 Summary

In this chapter, we formally presented the single-campaign real-time bidding strategy optimisation problem. Then we proposed a novel functional optimisation framework to find the optimal bidding function in a general function space. Given the winning functions fitting the real-world data, the derived optimal bidding functions are non-linear and concave w.r.t. the predicted KPI, which has not been studied in the previous literature about real-time bidding strategies. The widely adopted linear bidding function can also be derived from our framework when assuming the winning function to be linear. In the experiments, we compared the derived bidding strategies with other baselines and state-of-the-art linear bidding strategies under different budget and KPI settings. Both the offline and online experiments show that the proposed optimal bidding strategies were also practically the most effective.

This chapter acts as the core part of this thesis. It studies the fundamental problem of single-campaign bid optimisation in RTB display advertising. Based on the research work in this chapter, several directions for the further work are described as follows.

- As an intermediary agent like a DSP or an ad agency, how to optimise the overall performance over multiple campaigns it manages, particularly when these cam-

paigns are set with the CPA pricing scheme? This research problem and the proposed solution will be presented in Chapter 4.

- As discussed in Section 3.4.4, the offline training and evaluation based on historic data could be biased due to the data censorship from ad auction selection. In Chapter 5, the probability of observing each data instance will be estimated and a novel training and optimisation scheme will be proposed to yield less unbiased CTR estimator and bidding strategy.

- As mentioned in Section 3.2, the RTB ad market and the user behaviour are highly dynamic, which probably result in instability of advertising performance. In Chapter 6, a feedback control mechanism will be proposed, which is incorporated into the bidding agent to adaptively adjust the bid scale to maintain a certain KPI close to a predefined reference value.

3.8   Chapter Appendix: Game Theoretic Analysis

In this chapter appendix, we conduct a theoretic analysis of the optimal bidding functions in a symmetric game of repeated auctions with budget constraints. We will derive the analytic solutions of the optimal bidding function in the equilibrium of the first and second price auctions, respectively. Then based on the derived solutions, we notice that there actually exists a *tragedy of the common* situation for advertisers in RTB display advertising.

3.8.1   Problem Settings

Monotonously Increasing Bidding Function. In a clean game theoretic analysis setting [114], we suppose there are n ($n \geq 2$) advertisers with the same bidding strategy $b(\theta)$ which takes in the estimated CTR θ and outputs the bid price b. It is reasonable to assume $b(\theta)$ monotonically increases w.r.t. CTR θ, i.e.,

$$b(\theta_1) > b(\theta_2) \Leftrightarrow \theta_1 > \theta_2. \tag{3.33}$$

Later we will show the derived bidding functions are indeed monotonically increasing w.r.t. CTR θ.

Each time when an impression is auctioned, for each advertiser the CTR θ follows the same p.d.f. $p_\theta(\theta)$ independently (i.e., i.i.d.) and denote its c.d.f. as $F(\theta)$:

$$F(\theta) = \int_0^\theta p_\theta(\vartheta)d\vartheta, \tag{3.34}$$

$$\frac{\partial F(\theta)}{\partial \theta} = p_\theta(\theta). \tag{3.35}$$

The Winning Probability in a Symmetric Game. In such a setting, the winning probability of advertiser 1, without loss of generality, given her ad CTR θ is actually the probability of her ad CTR, is the largest one among the n advertisers:

$$w(\theta) = P(\theta > \theta_2, \theta > \theta_3, \ldots, \theta > \theta_n) = F(\theta)^{n-1}. \tag{3.36}$$

The Expected Utility. The expected utility given the CTR is denoted as $u(\theta)$. The specific form of $u(\theta)$ depends on the campaign KPI. For example, if the KPI is the click number, then

$$u_{\text{clk}}(\theta) = \theta. \tag{3.37}$$

If the KPI is the campaign's revenue, denoting the advertiser's true value on each click as r, then

$$u_{\text{rev}}(\theta) = r\theta. \tag{3.38}$$

The Expected Cost. The expected cost if winning the auction with a bid b is denoted as $c(b)$. In the RTB ad market we have the first price auction cost equal to the bid price

$$c_1(b) = b, \tag{3.39}$$

and the second price auction cost equal to the expected market price when it is lower than the bid price

$$c_2(b) = \frac{\int_0^b z p_z(z)dz}{\int_0^b p_z(z)dz}. \tag{3.40}$$

3.8.2   Equilibrium Bidding Function in the First Price Auctions

Suppose each of n advertisers wants to maximise her click number from T ad auctions with a budget B via designing a bidding function $b(\vartheta)$, where ϑ is a signal related to CTR θ. Suppose the advertiser knows the CTR θ but she tells her friend the CTR is ϑ and asks the friend to participate the auction for her, i.e., to perform the bid $b(\vartheta)$. Then its winning rate is $F(\vartheta)^{n-1}$ (see Eq. (3.36)). The optimisation problem with the first price auction cost function $c_1(b)$ is formulated as

$$\max_{b()} \ T \int_\theta \theta F(\vartheta)^{n-1} p_\theta(\theta) d\theta \qquad (3.41)$$

$$\text{subject to} \ T \int_\theta b(\vartheta) F(\vartheta)^{n-1} p_\theta(\theta) d\theta = B. \qquad (3.42)$$

Here the constraint could also be set as $T \int_\theta b(\vartheta) F(\vartheta)^{n-1} p_\theta(\theta) d\theta \leq B$, i.e., the cost not higher than the budget. However, it is obvious to see that when the cost is strictly lower than the total budget, the left budget could always generate non-negative expected utility, i.e., the optimal situation is always when the total budget is exhausted. Therefore, the above equality constraint is reasonable.

The Lagrangian is

$$\mathcal{L}(\vartheta, \lambda) = T \int_\theta (\theta - \lambda b(\vartheta)) F(\vartheta)^{n-1} p_\theta(\theta) d\theta - \lambda B \qquad (3.43)$$

and its derivative w.r.t. ϑ is

$$\frac{\partial \mathcal{L}(\vartheta, \lambda)}{\partial \vartheta} = T \int_\theta \left(-\lambda \frac{\partial b(\vartheta)}{\partial \vartheta} F(\vartheta)^{n-1} + (\theta - \lambda b(\vartheta))(n-1) F(\vartheta)^{n-2} p_\theta(\vartheta) \right) p_\theta(\theta) d\theta.$$
$$(3.44)$$

In a symmetric equilibrium, the objective is maximised when using the true signal, i.e., at $\vartheta = \theta$ [114]. It is intuitive that the utility can be maximised when the advertiser's friend uses the true CTR in the bidding. Therefore, we have

$$\left. \frac{\partial \mathcal{L}(\vartheta, \lambda)}{\partial \vartheta} = 0 \right|_{\vartheta = \theta}, \qquad (3.45)$$

which is

$$\int_\theta \left(-\lambda \frac{\partial b(\theta)}{\partial \theta} F(\theta)^{n-1} + (\theta - \lambda b(\theta))(n-1)F(\theta)^{n-2} p_\theta(\theta) \right) p_\theta(\theta) d\theta = 0, \quad (3.46)$$

$$\Rightarrow \lambda \frac{\partial b(\theta)}{\partial \theta} F(\theta)^{n-1} = (\theta - \lambda b(\theta))(n-1)F(\theta)^{n-2} p_\theta(\theta) \quad (3.47)$$

for all θ. To solve Eq. (3.47), we leverage the fact

$$\frac{\partial}{\partial \theta} b(\theta) F(\theta)^{n-1} = \frac{\partial b(\theta)}{\partial \theta} F(\theta)^{n-1} + b(\theta)(n-1)F(\theta)^{n-2} p_\theta(\theta) \quad (3.48)$$

Taking Eq. (3.47) into Eq. (3.48), we have

$$\frac{\partial}{\partial \theta} b(\theta) F(\theta)^{n-1} = \frac{\theta}{\lambda}(n-1)F(\theta)^{n-2} p_\theta(\theta). \quad (3.49)$$

Thus

$$b(\theta) F(\theta)^{n-1} = \int_0^\theta \frac{\vartheta}{\lambda}(n-1)F(\vartheta)^{n-2} p_\theta(\vartheta) d\vartheta + C, \quad (3.50)$$

where C is the constant independent with θ. From Eq. (3.50) we know when $\theta \to 0$, $b(\theta)F(\theta)^{n-1} \to 0$, so the constant $C = 0$. Therefore, we obtain the optimal bidding function

$$b(\theta) = \frac{1}{\lambda F(\theta)^{n-1}} \int_0^\theta \vartheta(n-1)p_\theta(\vartheta)F(\vartheta)^{n-2} d\vartheta. \quad (3.51)$$

From Eq. (3.51) we can see that there is a Lagrangian multiplier λ to solve, which is the difference of our solution from the classic first auction solution with no budget constraint [114].

To solve λ, we set the Lagrangian derivative to be 0 (also with $\vartheta = \theta$).

$$\frac{\partial \mathscr{L}(\theta, \lambda)}{\partial \lambda} = 0, \quad (3.52)$$

which just results in the constraint equation

$$T \int_\theta b(\theta) F(\theta)^{n-1} p_\theta(\theta) d\theta = B. \quad (3.53)$$

Taking our bidding function form Eq. (3.51) into Eq. (3.53), we have

$$\int_{\theta} \frac{1}{\lambda} \int_0^{\theta} \vartheta(n-1)p_{\theta}(\vartheta)F(\vartheta)^{n-2}d\vartheta \cdot p_{\theta}(\theta)d\theta = \frac{B}{T} \tag{3.54}$$

$$\Rightarrow \lambda = \frac{T}{B} \int_{\theta'} \int_0^{\theta'} \vartheta(n-1)p_{\theta}(\vartheta)F(\vartheta)^{n-2}d\vartheta \cdot p_{\theta}(\theta')d\theta'. \tag{3.55}$$

Here we use the notation θ' in the integration to avoid the notation conflict against the θ in Eq. (3.51). Finally, taking Eq. (3.55) into Eq. (3.51), we obtain the solution of the optimal budget constrained bidding function in equilibrium of the first price auction:

$$b(\theta) = \frac{B}{TF(\theta)^{n-1}} \cdot \frac{\int_0^{\theta} \vartheta(n-1)p_{\theta}(\vartheta)F(\vartheta)^{n-2}d\vartheta}{\int_{\theta'} \int_0^{\theta'} \vartheta(n-1)p_{\theta}(\vartheta)F(\vartheta)^{n-2}d\vartheta \cdot p_{\theta}(\theta')d\theta'}. \tag{3.56}$$

Analytic Solution with a Special Case. If we assume the uniform distribution of CTR in $[0,1]$

$$p_{\theta}(\theta) = 1, \tag{3.57}$$

$$F(\theta) = \theta, \tag{3.58}$$

then we can get the analytic solution of the optimal bidding function from Eq. (3.56) as

$$b(\theta) = \frac{B}{T\theta^{n-1}} \cdot \frac{\int_0^{\theta} \vartheta(n-1)\vartheta^{n-2}d\vartheta}{\int_{\theta'} \int_0^{\theta'} \vartheta(n-1)\vartheta^{n-2}d\vartheta d\theta'} \tag{3.59}$$

$$= \frac{B}{T\theta^{n-1}} \cdot \frac{\theta^n}{\int_{\theta'} \theta'^n d\theta'} \tag{3.60}$$

$$= \frac{B(n+1)\theta}{T}. \tag{3.61}$$

We can see from Eq. (3.61), with the simple uniform CTR distribution assumption, the optimal bid has a linear relationship w.r.t. the CTR, the number of competitors plus 1, and the average budget on each auction. More discussions will be given in Section 3.8.4.

3.8.3   Equilibrium Bidding Function in the Second Price Auctions

In the second price auction, we first introduce a theorem. Regard θ as a random variable and $b = b(\theta)$ as a dependent random variable. Furthermore, we define $F_b(b)$ as the

c.d.f. of b, i.e., the probability of performing a bid less than b:

$$F_b(b) = \int_0^b p_b(a)da. \tag{3.62}$$

Note that

$$F_b(b(\theta)) = P(b(\theta) > b(\theta_2)) = P(\theta > \theta_2) = F_\theta(\theta) \tag{3.63}$$

because $b(\theta)$ monotonously increases w.r.t. θ. Here we add subscripts to the c.d.f. functions F_b, F_θ and the following F_z to make differences.

Therefore, for the market price variable z, defined as the highest bid across $n-1$ competitors, its c.d.f $F_z(z)$ is

$$F_z(z) = F_b(z)^{n-1}, \tag{3.64}$$

and thus its p.d.f. is

$$p_z(z) = (n-1)F_b(z)^{n-2}p_b(z). \tag{3.65}$$

Now we derive the optimal bidding function in an equilibrium with the second price auction cost function $c_2(b)$. Suppose the click value is r and each advertiser wants to maximise her revenue across T auctions and with campaign budget B.[12] Again, if the bidding is based on a signal ϑ related with the CTR θ, the optimisation framework is

$$\max_{b()} \ T\int_\theta r\theta F_\theta(\vartheta)^{n-1}p_\theta(\theta)d\theta \tag{3.66}$$

$$\text{subject to } T\int_\theta \int_0^{b(\vartheta)} zp_z(z)dz \cdot p_\theta(\theta)d\theta = B. \tag{3.67}$$

---

[12]Optimising the revenue is equivalent to optimising the click number since the monetary value on each click are regarded as the same, i.e., r. Here we just want to make the dimensions of objective and constraint be the same (i.e., price) for readability, and also show different utility functions in our discussion.

The Lagrangian $\mathscr{L}(\vartheta, \lambda)$ is

$$\mathscr{L}(\vartheta, \lambda) = T \int_\theta \left(r\theta F_\theta(\vartheta)^{n-1} - \lambda \int_0^{b(\vartheta)} z p_z(z) dz \right) p_\theta(\theta) d\theta - \lambda B. \tag{3.68}$$

We can calculate its gradient w.r.t. ϑ

$$\frac{\partial \mathscr{L}(\vartheta, \lambda)}{\partial \vartheta} = T \int_\theta \left(r\theta(n-1)F_\theta(\vartheta)^{n-2} p_\theta(\vartheta) - \lambda b(\vartheta) p_z(b(\vartheta)) \frac{\partial b(\vartheta)}{\partial \vartheta} \right) p_\theta(\theta) d\theta. \tag{3.69}$$

Similarly to the derivations in first price auctions, the objective should be maximised when the signal used for bidding is based on true signal $\vartheta = \theta$, thus

$$\left. \frac{\partial \mathscr{L}(\vartheta, \lambda)}{\partial \vartheta} = 0 \right|_{\vartheta = \theta} \tag{3.70}$$

$$\Rightarrow r\theta(n-1)F_\theta(\theta)^{n-2} p_\theta(\theta) - \lambda b(\theta) p_z(b(\theta)) \frac{\partial b(\theta)}{\partial \theta} = 0, \tag{3.71}$$

for all θ.

As $b(\theta)$ monotonously increases w.r.t. θ, then their p.d.f.s $p_\theta(\theta)$ and $p_b(b)$ have the following relationship:

$$p_\theta(\theta) = p_b(b(\theta)) \frac{\partial b(\theta)}{\partial \theta}. \tag{3.72}$$

Taking Eqs. (3.65), (3.72) and (3.63) into Eq. (3.71), we have

$$r\theta(n-1)F_\theta(\theta)^{n-2} p_\theta(\theta) = \lambda b(\theta)(n-1)F_b(b(\theta))^{n-2} p_b(b(\theta)) \frac{\partial b(\theta)}{\partial \theta} \tag{3.73}$$

$$= \lambda b(\theta)(n-1)F_b(b(\theta))^{n-2} p_\theta(\theta) \tag{3.74}$$

$$= \lambda b(\theta)(n-1)F_\theta(\theta)^{n-2} p_\theta(\theta) \tag{3.75}$$

$$\Rightarrow b(\theta) = \frac{r\theta}{\lambda}. \tag{3.76}$$

From Eq. (3.76), we can see that the optimal bidding function in equilibrium of the second price auctions is linear w.r.t. the CTR θ. Similar with the case of the first price auctions, in the second price auctions the difference between the optimal bidding function derived in our multi-auction budget constrained scenario and the classic single-

auction non-budget scenario is the ratio λ.

To solve λ, we set the Lagrangian derivative to be 0, also with $\vartheta = \theta$.

$$\frac{\partial \mathscr{L}(\theta, \lambda)}{\partial \lambda} = 0, \tag{3.77}$$

which just results in the constraint equation. With Eqs. (3.65), (3.72) and (3.63), we have

$$\int_\theta \int_0^{\frac{r\theta}{\lambda}} z p_z(z) dz \cdot p_\theta(\theta) d\theta = \frac{B}{T} \tag{3.78}$$

$$\Rightarrow \int_\theta \int_0^{\frac{r\theta}{\lambda}} z(n-1) F_\theta \left(\frac{z\lambda}{r}\right)^{n-2} p_\theta \left(\frac{z\lambda}{r}\right) \frac{\lambda}{r} dz \cdot p_\theta(\theta) d\theta = \frac{B}{T} \tag{3.79}$$

$$\Rightarrow \int_\theta \int_0^\theta \frac{r\vartheta}{\lambda} (n-1) F_\theta(\vartheta)^{n-2} p_\theta(\vartheta) d\vartheta \cdot p_\theta(\theta) d\theta = \frac{B}{T} \tag{3.80}$$

$$\Rightarrow \frac{r}{\lambda} = \frac{B}{T \int_\theta \int_0^\theta \vartheta (n-1) F_\theta(\vartheta)^{n-2} p_\theta(\vartheta) d\vartheta \cdot p_\theta(\theta) d\theta}. \tag{3.81}$$

Thus we obtain the form of the optimal bidding function

$$b(\theta) = \frac{B\theta}{T \int_{\theta'} \int_0^{\theta'} \vartheta (n-1) F_\theta(\vartheta)^{n-2} p_\theta(\vartheta) d\vartheta \cdot p_\theta(\theta') d\theta'}. \tag{3.82}$$

Analytic Solution with a Special Case. Again, if we assume the uniform distribution of CTR in $[0, 1]$

$$p_\theta(\theta) = 1 \tag{3.83}$$

$$F(\theta) = \theta, \tag{3.84}$$

thus we can get the analytic solution of the optimal bidding function from Eq. (3.82) as

$$b(\theta) = \frac{B\theta}{T} \cdot \frac{1}{\int_{\theta'} \int_0^{\theta'} \vartheta (n-1) \vartheta^{n-2} d\vartheta \cdot d\theta'} \tag{3.85}$$

$$= \frac{B\theta}{T} \cdot \frac{1}{\int_{\theta'} \frac{n-1}{n} \theta'^n \cdot d\theta'} \tag{3.86}$$

$$= \frac{B\theta}{T} \cdot \frac{n(n+1)}{n-1}. \tag{3.87}$$

We can see from Eq. (3.87) the optimal bid price has a linear relationship with the CTR and the average budget per auction. Also it monotonously increases w.r.t. n when

$n \geq 2$.

3.8.4   Discussion: Tragedy of the Commons in RTB

We define the advertising performance comparison scheme as: first, comparing the achieved utilities (e.g., clicks), the higher utility the better; then comparing the cost for achieving such a utility, if the utility are the same, the lower cost the better.

From Eqs. (3.61) and (3.87) we can see the bid increases w.r.t. the number of auction competitors. The reason is that each advertiser wants to maximise the objective (click number or revenue) given the cost not higher than the budget. When there are n advertisers in an auction, the averaged winning probability of each advertiser is $1/n$. As a result, each advertiser will try to spend out the budget to win higher objective value. However, such an equilibrium is not efficient. In fact it results in a very low social welfare situation among advertisers: they exhaust all their budget to win the same utility, i.e., $1/n$ impressions and clicks.

A much better situation would be, every advertiser just spends B/k budget and still gets the same utility. Extremely, when $k \to +\infty$, the social welfare (revenue minus cost) among advertisers are maximised. In such a case, every advertiser will just bid 0 and the winner is randomly selected and pays 0 for each ad impression. However, this situation is never realistic since the participants of an auction will always compete with each other instead of cooperating, otherwise such a kind of trading cannot be regarded as an auction. In such an unstable situation, each advertiser will higher her bid, which is expected to acquire more utility given the current market situation (i.e., $F_z(z)$). Finally, the system will get to the equilibrium of Eqs. (3.61) and (3.87) where every advertiser spends out the budget.

This is actually the *tragedy of the commons* [120], which is a reflection of multi-player cases of *prisoner's dilemma* [121] in RTB campaign competition with budget constraint. The advertisers spend more money to acquire the same advertising utility that they could acquire when they cooperated, although the RTB mechanism enables them to effectively reallocate their budget to targeted volume. The publishers (or ad exchanges) earn the upper bound they can earn as the advertisers have spent out their budget.

A simple example of tragedy of the commons in RTB display advertising is shown

Table 3.5: An example of tragedy of the commons in RTB.

	all others bid low	any of others bid high
bid low	$(r - b_{\text{low}})/n$	0
bid high	$r - b_{\text{low}}$	$(r - b_{\text{high}})/n_{\text{high}}$

in Table 3.5. There are n advertisers in an auction and suppose there is only two possible bid actions, i.e., b_{low} and b_{high}, and the impression value is r. Suppose $b_{\text{low}} < b_{\text{high}} < r$. The left column shows the actions for an advertiser; the first row shows the possible situations of other advertisers; the table entries are the corresponding reward (or welfare) for the advertiser corresponding to her action taken in a situation. In the situation when all others bid b_{low}, if the advertiser bids b_{low} as well, then with $1/n$ probability she will win the auction and make $r - b_{\text{low}}$ profit; if she bids b_{high}, she will surely win the auction and make $r - b_{\text{low}}$ profit. In the situation when any of others bid b_{high}, if the advertiser bids b_{low}, she will lose the auction and earn nothing; if she bids b_{high}, then with $1/n_{\text{high}}$ probability she will win the auction and make $r - b_{\text{high}}$ profit, where n_{high} is the number of advertisers bidding b_{high}. Obviously, the bidding-high strategy dominates the bidding-low strategy. When every advertiser performs the bidding-high strategy, $n_{\text{high}} = n$, which makes $(r - b_{\text{high}})/n_{\text{high}} < (r - b_{\text{low}})/n$, the system reaches an equilibrium of low social welfare.

Some possible solutions for advertisers would be (i) to set the cost related objectives, such as ROI; (ii) to set the upper bound of their bid prices. However, the success of reaching such better solutions still needs all advertisers to cooperate, just like the cooperation of prisoners to achieve the higher reward for each one.

Chapter 4

Multi-Campaign Statistical Arbitrage Mining

4.1 Background and Motivations

Real-Time Bidding (RTB) has emerged to be a frontier for Internet advertising [16, 17]. It mimics stock spot exchanges and utilises computers to *programmatically* buy display ads in real-time and per impression via an instant auction mechanism between buyers (advertisers) and sellers (publishers) [14]. Such automation not only improves efficiency and scales of the buying process across lots of available inventories, but, most importantly, encourages performance driven advertising based on targeted clicks, conversions etc., by using real-time audience data, as has been shown in Chapter 3.

As a result, ad impressions become more and more *commoditised* in the sense that the effectiveness (quality) of an ad impression does not rely on where it is bought or whom it belongs to any more, but depends on how much it will benefit the campaign target (e.g., underlying web users' satisfactions and their direct responses)[1].

According to the Efficient Market Hypothesis (EMH) in finance, in a perfectly "efficient" market, security (such as stock) prices should fully reflect all available information at any time [122]. As such, no *arbitrage* opportunity exists, i.e., one can neither buy securities which are worth more than the selling price, nor sell securities worth less than the selling price without making riskier investment [79]. However, due to the heavily-fragmented, non-transparent ad marketplaces and the existence of

---

[1]The discussion in this chapter is limited to performance-driven campaigns and direct responses such as clicks and conversions only, whereas for the purpose of branding, the quality of publishers still play an important role in defining the ad inventory quality.

various ad types, e.g., sponsored search, display ads, affiliated networks, and pricing schemes, e.g., cost per mille (CPM), cost per click (CPC), cost per action (CPA), the ad markets are not informationally efficient. In other words, two display opportunities with similar or identical targeted audiences and visit frequency may sell for quite different prices. Such a price discrepancy caused by market inefficiency create some *arbitrage* opportunities for some intermediaries in the advertising ecosystem: they buy ad inventories at a relatively low price and sell them at a higher price. For example, if an advertiser undertakes a campaign to sell travel insurances, investing in few highly visited web pages might be more costly than consolidating the same quality display opportunities from a wide range of unpopular personal blogs about travels. In such a case, if an ad agency sells the advertiser the ad inventories with the high price as from the popular websites and buys such inventories from many unpopular media at a low price, it makes arbitrage.

While exploiting such price discrepancies is still debatable in the advertising field, the following four arbitrage situations exist:

I **Inter-exchange arbitrage.** Multiple ad exchanges exist. As the supply and demand vary across exchanges for the same user types or targeting rules, there exist intermediate agencies that act as a buyer with low bid in exchange A and as a seller with high reserve price in exchange B in order to make profits [123]. The resell of the ad inventories should be accomplished within very short time (below 50ms).

II **Guaranteed delivery and spot market arbitrage.** Some DSPs offer advertisers the contracts with guaranteed delivery [31] while buying ad inventories over an RTB exchange with non-guaranteed spot prices [124]. Conversely, some ad agencies buy inventories in advance in bulk for fixed "preferential rates" from private marketplaces, and then charge a client for their campaigns with the spot prices.

III **Publisher volume I/O arbitrage.** One can purchase traffic to a particular web page and subsequently make more from ad revenue than the initial inbound click cost. An extreme case is a homepage purely dedicated to host ads: the Million

Dollar Homepage[2].

IV **Pricing scheme arbitrage.** In RTB, different counter-parties prefer different pricing schemes in order to reduce their risk of deficit [125]. CPM is commonly used for RTB auctions and preferred by publishers because it is likely to generate stable income from the site volume. By contrast, advertisers focusing on performance are likely to follow CPA and CPC pricing schemes as they are directly related to return on investment (ROI) [126]. As such, if the CPM cost to yield a user conversion is less than the CPA payoff for the conversion, an intermediate agency can earn a positive profit.

Scientifically, this is of great interest as it presents a new type of data mining problem, which demands a principled mathematical formulation and novel computational solution to mine and exploit arbitrage opportunities in RTB display advertising. Commercially and socially, principled ad arbitrage algorithms would not only ensure the business more smooth and risk free (e.g., III & IV), but also make the ad markets more transparent and informationally more efficient (e.g., I, II & IV) by connecting otherwise segmented markets to correct the misallocation of risks and prices, and eventually reach to an "arbitrage free" equilibrium.

In this chapter, we formulate Statistical Arbitrage Mining (SAM) and present a solution in the context of display advertising. We focus on modelling discrepancies between CPA-based campaigns and CPM-based ad inventories (IV above), while the arbitrage models for the remaining cases can be obtained analogically. The studied arbitrage is a stochastic one due to the uncertainty of market supply/demand and users response behaviour, e.g, clicks and conversions. The probability distribution of the arbitrage profit from an ad display opportunity is estimated by user response predictors [4] and the bid landscape forecasting models [24], trained on historic large-scale data. Essentially, the proposed statistical arbitrage miner is a campaign-independent RTB bidder, which assesses the arbitrage opportunity for an incoming CPM bid request against a *portfolio* of CPA campaigns, then selects a campaign and provides a bid accordingly. Different from Chapter 3 on single-campaign RTB bidding strategies, this chapter introduces the concept of *meta-bidder*, which performs the bidding for a

---

[2]http://www.milliondollarhomepage.com/

portfolio of ad campaigns, similar to a hedge fund holding a set of valid assets in financial markets. In the proposed SAM framework, (i) functional optimisation is utilised to seek for an optimal bidding function to maximise the expected arbitrage profit across multiple campaigns managed by the meta-bidder, and (ii) a portfolio-based risk management solution is leveraged to reallocate the bidding volume and the budget across multiple campaigns to make a trade-off between arbitrage risk and return. We propose to jointly optimise those two components in an EM fashion with high efficiency to make meta-bidder successfully catch the transient statistical arbitrage opportunities in RTB. Experiments on both large-scale datasets and online A/B tests demonstrate the large improvement of our proposed SAM solutions over the state-of-the-art baselines.

4.2　Problem Definition
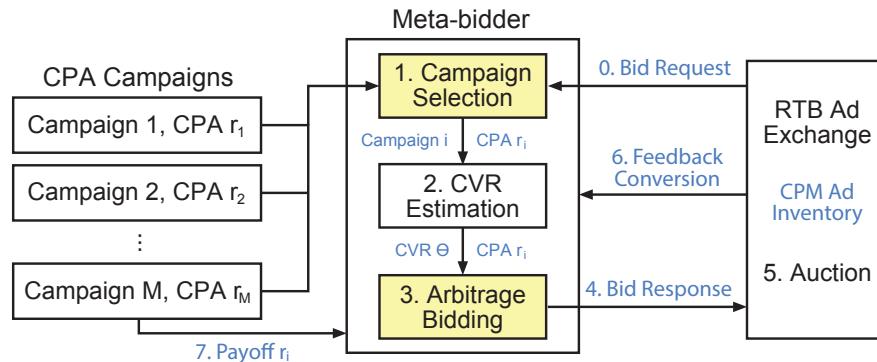
Suppose there is an ad agency acting on behalf of advertisers to run their ad campaigns. To hedge advertisers' risk, quite often an ad agent gets paid on the basis of the performance: receive a payoff each time a placed ad eventually leads to a product purchase (cost-per-action, CPA)[3]. Note that it remains active research to determine whether and how much a purchase action is attributed to the ads previously shown to the user. In this research, we adopt the last-touch attribution model commonly used in the industry – the last ad impression before the user's conversion event is assigned with the full attribution credit [127].

To run the campaigns and place the ads, the agency then goes to the RTB spot market to purchase ad impressions. In RTB, the ad agency pays the cost for each ad impression displayed (cost-per-mille, CPM) on the basis of ad auctions. In essence, the ad agency is an arbitrageur, making a profit so long as the payoff by conversions (CPA) is higher than the ad impression cost (CPM) of acquiring relevant users to making the purchase. Potentially, the ad agency could in parallel run a large number of campaigns from various advertisers to scale up their profit. Note that the ad agencies build their business by taking the risk from the uncertainty of market competitions and user response behaviours. For the entire ad ecosystem, it is healthy as it protects both advertisers and publishers by introducing an intermediate layer that exploits (and

---

[3]A notable example is `mobpartner.com` who explicitly offers payoffs (CPA deals) for anyone who can acquire the needed customers programmatically.

Table 4.1: Notations and descriptions.

Notation	Description
$\boldsymbol{x}$	The bid request represented by its features.
$p_x(\boldsymbol{x})$	The probability density function of $\boldsymbol{x}$.
i	The ith campaign in the DSP portfolio.
M	The number of campaigns in the DSP portfolio.
r_i	The payoff of campaign i for each conversion.
R	The variable of meta-bidder arbitrage profit.
C	The variable of meta-bidder arbitrage cost.
$\theta(\boldsymbol{x}, i)$	The predicted CVR if i wins the auction of $\boldsymbol{x}$. Occasionally θ is used to refer to a specific CVR.
$p_\theta^i(\theta)$	The probability density function of CVR θ for campaign i.
B	The meta-bidder total budget.
T	The estimated number of bid requests during the arbitrage period.
$b(\theta, r)$	The bidding function which returns the bid. b is also used to refer to a specific bid value.
$w(b)$	The probability of winning a bid request with bid price b.
v_i	The probability of selecting campaign i. For multiple campaigns, the campaign selection probability vector is $\boldsymbol{v} = (v_1, v_2, \ldots, v_M)^T$.

Figure 4.1: An ad agency running a meta-bidder (arbitrageur) for statistical arbitrage mining.

ultimately remove) the discrepancies between market segments (in this case, the two pricing schemes, CPA and CPM).

Traditionally, these arbitrages are accomplished manually. With statistical approaches, it is possible that the above operations can be automatically done by an intelligent meta-bidder across campaigns, where for a certain CPA campaign, the meta-bidder seeks cost-effective ad impressions with high conversion possibility and low market competition.

With the notations summarised in Table 4.1 and an illustration how the SAMer works in Figure 4.1, the research problem is mathematically formulated as below. Suppose there exist M CPA-based campaigns. Each campaign i has set its payoff for a

conversion as r_i. Over period T, the meta-bidder keeps receiving bid requests at time $t \in \{1, \ldots, T\}$, where each bid request is represented with high dimension feature vector $\boldsymbol{x}_t$ and if won, it is charged based on CPM. For each of the incoming bid requests, the Statistical Arbitrage Mining (SAM) problem is to *select a campaign* and *specify its bid* such that over the period T the expected total arbitrage profit (accumulated payoff minus cost) is maximised.

We consider the following process. When a bid request comes, the meta-bidder samples campaign i with probability v_i to participate the RTB auction, where $\sum_{i=1}^{M} v_i = 1$. Once campaign i is selected, the meta-bidder then estimates its conversion rate (CVR), denoted as $\theta(\boldsymbol{x}_t, i)$, i.e., if the ad is placed in this impression, how likely the underlying user will see the ad and eventually convert (purchase) [4]. After that, the meta-bidder generates the bid price via a bidding function $b(\theta, r_i)$ depending on the CVR $\theta(\boldsymbol{x}_t, i)$ and the conversion payoff r_i [128].

Given campaign selection probability $\boldsymbol{v} = (v_1, v_2, \ldots, v_M)^T$ and bidding function $b(\theta, r)$, the meta-bidder's total arbitrage profit is given by summation over bid requests and campaigns:

$$R(\boldsymbol{v}, b(\theta, r)) = \sum_{t=1}^{T} \sum_{i=1}^{M} \Big(\theta(\boldsymbol{x}_t, i) r_i - b(\theta(\boldsymbol{x}_t, i), r_i) \Big) \cdot w(b(\theta(\boldsymbol{x}_t, i), r_i)) v_i, \qquad (4.1)$$

where $w(b)$ is the probability of winning an RTB auction given the bid b. The product $w(b) v_i$ specifies the probability a campaign is selected *and* wins the auction; $(\theta r_i - b)$ is profit for the winning campaign. And the total cost upper bound is

$$C(\boldsymbol{v}, b(\theta, r)) = \sum_{t=1}^{T} \sum_{i=1}^{M} b(\theta(\boldsymbol{x}_t, i), r_i) w(b(\theta(\boldsymbol{x}_t, i), r_i)) v_i, \qquad (4.2)$$

where the bid price b is the maximum possible cost for a campaign to be placed no matter in the first- or second-price auction with hard or soft reserve price [37], which is regarded as the upper bound of the cost as modelled in Chapter 3.

Next, we need to model how likely we will see an ad impression with feature $\boldsymbol{x}_t$ in the future. Assume $\boldsymbol{x}_t \sim p_x(\boldsymbol{x}_t)$, that is for a relatively short period, the bid request feature is drawn from an i.i.d. built from historic data. The whole model needs to be re-trained periodically with the latest data. Detailed empirical study on the re-training

frequency for dynamic arbitrage will be given in Section 4.6.3. Taking the integration over $\boldsymbol{x}$ gives the expected profit:

$$
\begin{aligned}
\mathbb{E}[R(\boldsymbol{v}, b(\boldsymbol{\theta}, r))] &= T \int_{\boldsymbol{x}} \sum_{i=1}^{M} \Big(\boldsymbol{\theta}(\boldsymbol{x}, i) r_i - b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i)\Big) w(b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i)) v_i p_x(\boldsymbol{x}) d\boldsymbol{x} \\
&= T \sum_{i=1}^{M} v_i \int_{\boldsymbol{x}} \Big(\boldsymbol{\theta}(\boldsymbol{x}, i) r_i - b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i)\Big) w(b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i)) p_x(\boldsymbol{x}) d\boldsymbol{x} \\
&= T \sum_{i=1}^{M} v_i \int_{\boldsymbol{\theta}} \Big(\boldsymbol{\theta} r_i - b(\boldsymbol{\theta}, r_i)\Big) w(b(\boldsymbol{\theta}, r_i)) p_{\theta}^i(\boldsymbol{\theta}) d\boldsymbol{\theta},
\end{aligned} \tag{4.3}
$$

where $p_{\theta}^i(\boldsymbol{\theta}(\boldsymbol{x}, i)) = p_x(\boldsymbol{x})/\|\nabla \boldsymbol{\theta}(\boldsymbol{x}, i)\|$ as there is a deterministic relationship between $\boldsymbol{x}$ and its estimated CVR $\boldsymbol{\theta}(\boldsymbol{x}, i)$, also given in Section 3.2. Similarly, the upper bound of the total cost is rewritten as

$$
\begin{aligned}
\mathbb{E}[C(\boldsymbol{v}, b(\boldsymbol{\theta}, r))] &= T \int_{\boldsymbol{x}} \sum_{i=1}^{M} b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i) w(b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i)) v_i p_x(\boldsymbol{x}) d\boldsymbol{x} \\
&= T \sum_{i=1}^{M} v_i \int_{\boldsymbol{x}} b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i) w(b(\boldsymbol{\theta}(\boldsymbol{x}, i), r_i)) p_x(\boldsymbol{x}) d\boldsymbol{x} \\
&= T \sum_{i=1}^{M} v_i \int_{\boldsymbol{\theta}} b(\boldsymbol{\theta}, r_i) w(b(\boldsymbol{\theta}, r_i)) p_{\theta}^i(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned} \tag{4.4}
$$

Finally, the SAM problem is cast as a constrained optimisation problem: to find campaign selection probability $\boldsymbol{v}$ and bidding function $b(\boldsymbol{\theta}, r)$ to maximise the expected arbitrage profit with budget and risk constraints:

$$
b_{\text{SAM}}(), \boldsymbol{v}^*() = \underset{b(), \boldsymbol{v}}{\arg\max} \ \mathbb{E}[R] \tag{4.5}
$$

$$
\text{s.t.} \ \mathbb{E}[C] = B \tag{4.6}
$$

$$
\text{Var}[R] \leq h \tag{4.7}
$$

$$
\boldsymbol{0} \leq \boldsymbol{v} \leq \boldsymbol{1} \tag{4.8}
$$

$$
\boldsymbol{v}^T \boldsymbol{1} = 1, \tag{4.9}
$$

where variance $\text{Var}[R]$ is used to measure the risk of the profit and h is a parameter for an upper tolerable risk. In addition, similar to the cost upper bound constraint setting in Chapter 3, it is reasonable to assume that the ad auction volume is large enough to exhaust the meta-bidder budget.

We propose to solve the problem (Eq. (4.5)) in an EM fashion. In particular, the campaign selection probability $\boldsymbol{v}$ is regarded as the latent factors to infer and the bidding function $b(\theta, r)$ is regarded as the parameter used to maximise the optimisation target. In E-step, we fix the current estimated bidding function $b(\theta, r)$ and solve the optimal campaign selection probability $\boldsymbol{v}$ with the constraints Eqs. (4.7), (4.8), & (4.9). In M-step, we fix the campaign selection probability $\boldsymbol{v}$ and seek for the optimal bidding function $b(\theta, r)$ to maximise the target under the budget constraint Eq. (4.6). When the EM iterations get converged, all the constraints are satisfied and the target is maximised (at least in a local maxima). The following Section 4.3 will describe the detailed solution of optimal bidding function (M-step), and Section 4.4 will discuss the solution of campaign selection probability (E-step).

4.3 Optimal Arbitrage Bidding Function

With the fixed $\boldsymbol{v}$ and the budget constraint at Eq. (4.6), we have a functional optimisation problem in M-step to find the optimal bidding function $b(\theta, r)$ to maximise the profit across the multiple campaigns:

$$\max_{b()} \ T \sum_{i=1}^{M} v_i \int_\theta \Big(\theta r_i - b(\theta, r_i) \Big) w(b(\theta, r_i)) p_\theta^i(\theta) d\theta \tag{4.10}$$

$$\text{s.t.} \ T \sum_{i=1}^{M} v_i \int_\theta b(\theta, r_i) w(b(\theta, r_i)) p_\theta^i(\theta) d\theta = B. \tag{4.11}$$

The Lagrangian is

$$\mathcal{L}(b(\theta, r), \lambda) = T \sum_{i=1}^{M} v_i \int_\theta \Big(\theta r_i - b(\theta, r_i) \Big) w(b(\theta, r_i)) p_\theta^i(\theta) d\theta$$

$$- \lambda T \sum_{i=1}^{M} v_i \int_\theta b(\theta, r_i) w(b(\theta, r_i)) p_\theta^i(\theta) d\theta + \lambda B \tag{4.12}$$

$$= T \sum_{i=1}^{M} v_i \int_\theta \Big(\theta r_i - (1 + \lambda) b(\theta, r_i) \Big) w(b(\theta, r_i)) p_\theta^i(\theta) d\theta + \lambda B. \tag{4.13}$$

Taking its functional derivative w.r.t. $b(\theta, r)$, we have

$$\frac{\partial \mathcal{L}(b(\theta, r), \lambda)}{\partial b(\theta, r)} = T \sum_{i=1}^{M} \left[(\theta r_i - (1 + \lambda) b(\theta, r_i)) \frac{\partial w(b(\theta, r_i))}{\partial b(\theta, r_i)} - (1 + \lambda) w(b(\theta, r_i)) \right] v_i p_\theta^i(\theta).$$

$$\tag{4.14}$$

(a) Winning function
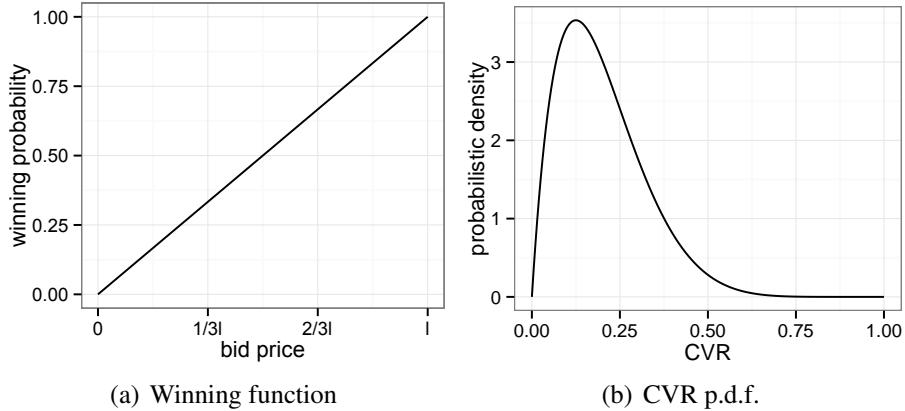
(b) CVR p.d.f.

Figure 4.2: Linear winning function $w(b(\theta))$ and beta CVR pdf $p_\theta(\theta)$.

A sufficient (but not necessary) condition of making this derivative be 0 is

$$\left(\frac{\theta r_i}{1+\lambda} - b(\theta, r_i)\right)\frac{\partial w(b(\theta, r_i))}{\partial b(\theta, r_i)} = w(b(\theta, r_i)), \tag{4.15}$$

for all campaign i. With the specific functional form of winning function $w(b)$ we can derive the optimal SAM bidding function. Below we show solutions in two special cases. These derivations follow the similar (but different) thinking as in Chapter 3. It is necessary to provide the following detailed derivations to make the solution self-contained and comprehensive.

4.3.1 Uniform Market Price Solution

Here we make a simple example of a linear winning function form (see Figure 4.2(a)) based on the assumption of the uniform market price[4] distribution in $[0, l]$:

$$w(b(\theta, r)) = \frac{b(\theta, r)}{l}, \tag{4.16}$$

where the function domain is also $[0, l]$. l is the upper bound of bid price and there is no need to bid higher than l.

Taking Eq. (4.16) into Eq. (4.15), we have the optimal arbitrage bidding function

---

[4]Market price refers to the highest bid price amongst the competitors for each auction [28]. From a bidder's perspective, it can win an auction if the its bid price is higher than the market price on this auction.

b_{SAM_1} as

$$b_{\text{SAM}_1}(\theta, r) = \frac{r\theta}{2(1+\lambda)}. \tag{4.17}$$

To calculate the optimal λ, the Euler-Lagrange condition of λ from Eq. (4.13) is $\partial \mathcal{L}(b(\theta, r), \lambda)/\partial \lambda = 0$, i.e.,

$$\int_{\theta} b(\theta, r) w(b(\theta, r)) p_{\theta}(\theta) d\theta = \frac{B}{T}. \tag{4.18}$$

Taking Eqs. (4.16) and (4.17) into Eq. (4.18) gives

$$\int_{\theta} \left(\frac{r}{2(1+\lambda)} \right)^2 \frac{\theta^2}{l} p_{\theta}(\theta) d\theta = \frac{B}{T} \tag{4.19}$$

$$\Rightarrow \frac{r^2}{4(1+\lambda)^2 l} \int_{\theta} \theta^2 p_{\theta}(\theta) d\theta = \frac{B}{T}. \tag{4.20}$$

Using a new notation $\phi \equiv \int_{\theta} \theta^2 p_{\theta}(\theta) d\theta$, we have

$$\lambda = \frac{r}{2} \sqrt{\frac{T\phi}{Bl}} - 1. \tag{4.21}$$

Substituting Eq. (4.21) into Eq. (4.17) gives the final solution of bidding function

$$b_{\text{SAM}_1}(\theta, r) = \sqrt{\frac{Bl}{T\phi}} \theta, \tag{4.22}$$

where surprisingly the bidding function does not depend on r. This is because the linear forms of $w(b)$ in Eq. (4.16) and $b_{\text{SAM}_1}(\theta, r)$ in Eq. (4.17) make θ factorised out from $r/(1+\lambda)$ in Eq. (4.20), which then removes the factor of $r/(1+\lambda)$. ϕ depends on the probabilistic distribution $p_{\theta}(\theta)$, e.g., the beta distribution BETA(2,8) as shown in Figure 4.2(b), and can be calculated with the empirical data.

4.3.2   Long Tail Market Price Solution

Now consider a more practical winning function used in Section 3.3.1, which is based on a long tail market price distribution $p_z(z) = l/(z+l)^2$ with parameter l. As such,

the winning function is

$$w(b(\theta,r)) = \int_0^{b(\theta,r)} p_z(z)dz = \frac{b(\theta,r)}{b(\theta,r)+l}. \tag{4.23}$$

The real-world data analysis on winning prices as shown in Figure 3.5 demonstrates the feasibility of adopting the winning function in Eq. (4.23) in practice. Taking Eq. (4.23) into Eq. (4.15) gives the optimal arbitrage bidding function b_{SAM_2} as

$$b_{\text{SAM}_2}(\theta,r) = \sqrt{\frac{rl\theta}{1+\lambda} + l^2} - l, \tag{4.24}$$

which is in a concave form w.r.t. CVR θ.

The Solution of λ. The calculation of the optimal value of λ follows almost the same routine as in Section 3.3.4 except for multiple campaigns with probabilistic selection schemes. To make the thesis self-contained, the solution procedures are also provided here although it would be to-some-extent redundant.

To calculate the optimal λ, the Euler-Lagrange condition of λ is Eq. (4.18). With Eq. (4.24), we explicitly regard λ as an input of the bidding function $b(\theta,r,\lambda)$ and rewrite Eq. (4.18) as

$$\sum_{i=1}^{M} v_i \int_\theta b(\theta,r,\lambda)w(b(\theta,r,\lambda))p_\theta^i(\theta)d\theta = \frac{B}{T}. \tag{4.25}$$

In most situations except some special cases like Section 4.3.1, λ has no analytic solution. For numeric solution, since Eq. (4.25) can be rewritten as

$$\sum_{i=1}^{M} v_i \int_\theta \left(b(\theta,r,\lambda)w(b(\theta,r,\lambda)) - \frac{B}{T} \right) p_\theta^i(\theta)d\theta = 0, \tag{4.26}$$

we can obtain a feasible solution of λ by solving the minimisation problem

$$\min_\lambda \sum_{i=1}^{M} v_i \int_\theta \frac{1}{2} \left(b(\theta,r,\lambda)w(b(\theta,r,\lambda)) - \frac{B}{T} \right)^2 p_\theta^i(\theta)d\theta. \tag{4.27}$$

If we have a sufficient number N_i of observations of θ's for each campaign i, we

can write Eq. (4.27) in a discrete form over the observations

$$\min_{\lambda} \sum_{i=1}^{M} v_i \sum_{k=1}^{N_i} \frac{1}{2} \left(b(\theta_k^i, r, \lambda) w(b(\theta_k^i, r, \lambda)) - \frac{B}{T} \right)^2, \tag{4.28}$$

where we can use (mini-)batch descent or stochastic gradient descent to solve λ by the following iteration:

$$\lambda \leftarrow \lambda - \eta \sum_{i=1}^{M} v_i \sum_{k=1}^{N_i} \left(b(\theta_k^i, r, \lambda) w(b(\theta_k^i, r, \lambda)) - \frac{B}{T} \right) \cdot \tag{4.29}$$
$$\left(\frac{\partial b(\theta_k^i, r, \lambda)}{\partial \lambda} w(b(\theta_k^i, r, \lambda)) + b(\theta_k^i, r, \lambda) \frac{\partial w(b(\theta_k^i, r, \lambda))}{\partial \lambda} \right),$$

until convergence. Usually, as $b(\theta, r, \lambda)$ has a monotonic relationship with λ and $w(b(\theta, r, \lambda))$ monotonically increases against $b(\theta, r, \lambda)$, $b(\theta_k^i, r, \lambda) w(b(\theta_k^i, r, \lambda))$ has a monotonic relationship with λ. For example, with the bidding function as Eq. (4.24) and the winning function as Eq. (4.23), the factor $b(\theta_k^i, r, \lambda) w(b(\theta_k^i, r, \lambda))$ decreases monotonically against λ, which makes the optimal solution quite easy to find.

4.4 Optimal Campaign Selection

Fixing the resolved optimal arbitrage bidding function $b(\theta, r)$ from previous M-step, we can optimise the campaign selection probability $\boldsymbol{v}$ and check whether it is better to reallocate the volume for each campaign.

We here introduce the concept of SAM profit margin γ in RTB display advertising. The profit margin is a measure of ROI; it is the ratio of the profit of the advertising, either from one campaign or a set of them (meta-bidder), divided by the advertising cost during the corresponding period:

$$\gamma = \frac{R}{C} = \text{ROI.} \tag{4.30}$$

With the dynamics of the RTB spot market and user response behaviour, the RTB advertising performance measured by ROI is stochastic, thus γ is modelled as a random variable with expectation and variance. By modelling γ_i for each campaign i, the optimal campaign selection can be solved by portfolio-based risk management methods.

4.4.1 Single Campaign

With the optimal arbitrage bidding function $b(\theta, r)$ as derived from the condition in Eq. (4.15), the expectation and variance of the profit margin γ_i for each campaign i can be calculated by

$$\mu_i(b) = \mathbb{E}[\gamma_i] = \mathbb{E}\left[\frac{R_i(\mathbf{v}_{i=1}, b)}{C_i(\mathbf{v}_{i=1}, b)}\right], \tag{4.31}$$

$$\sigma_i^2(b) = \mathbb{E}\left[\frac{R_i(\mathbf{v}_{i=1}, b)^2}{C_i(\mathbf{v}_{i=1}, b)^2}\right] - \mathbb{E}\left[\frac{R_i(\mathbf{v}_{i=1}, b)}{C_i(\mathbf{v}_{i=1}, b)}\right]^2, \tag{4.32}$$

where $R_i(\mathbf{v}_{i=1}, b)$ and $C_i(\mathbf{v}_{i=1}, b)$ are as in Eqs. (4.1) and (4.2) with $v_i = 1$ and $v_j = 0$ for all other campaign j. Both $\mu_i(b)$ and $\sigma_i^2(b)$ can be estimated via Monte Carlo (MC) sampling methods [129]: (i) repeat N times on sampling T bid requests from the training data and calculate $R_i(\mathbf{v}_{i=1}, b)$ and $C_i(\mathbf{v}_{i=1}, b)$, then (ii) calculate the expectation and variance using these N observations of $R_i(\mathbf{v}_{i=1}, b)$ and $C_i(\mathbf{v}_{i=1}, b)$. Sampling-based methods are a type of important solutions to estimating (posterior) distribution of random variables in Bayesian inference [130, 131, 132].

4.4.2 Campaign Portfolio

Suppose there are M campaigns in the meta-bidder with CPA contracts. For each campaign i, as discussed in Section 4.4.1, there is a variable of achieved profit margin γ_i given the bidding function $b(\theta, r)$, and its expectation is $\mu_i(b)$ and standard deviation is $\sigma_i(b)$. As such, the vector of expected profit margins for these M campaigns is

$$\boldsymbol{\mu}(b) = (\mu_1(b), \mu_2(b), \dots, \mu_M(b))^T \tag{4.33}$$

and the covariance matrix for the profit margins of the M campaigns is

$$\boldsymbol{\Sigma}(b) = \begin{pmatrix} \sigma_{1,1}(b) & \sigma_{1,2}(b) & \cdots & \sigma_{1,M}(b) \\ \sigma_{2,1}(b) & \sigma_{2,2}(b) & \cdots & \sigma_{2,M}(b) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{M,1}(b) & \sigma_{M,2}(b) & \cdots & \sigma_{M,M}(b) \end{pmatrix}, \tag{4.34}$$

where each element

$$\sigma_{i,j}(b) = \psi_{i,j}\sigma_i(b)\sigma_j(b), \tag{4.35}$$

where $\psi_{i,j} \in [-1,1]$ is the profit margin correlation factor between campaign i and j, which can be calculated by routine given the profit margin time series of the two campaigns i and j [65].

Such a probabilistic campaign combination is called as *campaign portfolio* in this thesis. With the campaign selection probability $\boldsymbol{v}$, the campaign portfolio expected profit margin and its variance are

$$\mu_p(\boldsymbol{v},b) = \boldsymbol{v}^T\boldsymbol{\mu}(b), \tag{4.36}$$

$$\sigma_p^2(\boldsymbol{v},b) = \boldsymbol{v}^T\boldsymbol{\Sigma}(b)\boldsymbol{v}. \tag{4.37}$$

Generally, the arbitrage profit margin may change w.r.t. the allocated volume: the more bid request volume, the more statistical arbitrage opportunities, and the higher margin. For simplicity, we assume that the profit margin distribution does not change much w.r.t. the auction volume allocated to the campaign during a short period. The empirical results in Section 4.6.2 will demonstrate the feasibility of the assumption.

4.4.3   Campaign Portfolio Optimisation

The E-step of the original optimisation problem Eq. (4.5), with the fixed bidding function and constraint Eqs. (4.7), (4.8), & (4.9), can be rewritten by taking the Lagrangian as

$$\max_{\boldsymbol{v}}\ \boldsymbol{v}^T\boldsymbol{\mu}(b) - \alpha\boldsymbol{v}^T\boldsymbol{\Sigma}(b)\boldsymbol{v}, \tag{4.38}$$

$$\text{s.t. }\ \boldsymbol{v}^T\boldsymbol{1} = 1,\ \ \boldsymbol{0} \leq \boldsymbol{v} \leq \boldsymbol{1},$$

where the Lagrangian multiplier α acts as a risk-averse parameter to balance the expected profit margin and its variance. This optimisation framework is widely used as portfolio selection [65, 68, 72, 27], where maximisation of the objective Eq. (4.38) is equivalent with the maximisation of the mean $\boldsymbol{v}^T\boldsymbol{\mu}(b)$ with constrained variance $\boldsymbol{v}^T\boldsymbol{\Sigma}(b)\boldsymbol{v}$.

Algorithm 4.1 Statistical Arbitrage Mining for Display Advertising

---

Require: Meta-bidder winning function $w(b)$
Require: CTR distribution $p_\theta^i(\theta)$ for each campaign i
  Initialise $b(\theta, r) = r\theta$ and $\boldsymbol{v} = \boldsymbol{1}/M$.
  while not converged **do**
    E-step:
      Get $\boldsymbol{\mu}(b)$ and $\boldsymbol{\Sigma}(b)$ by Eq. (4.33) and Eq. (4.35)
      Solve optimal $\boldsymbol{v}$ by Eq. (4.38)
    M-step:
      Get the bidding function form by $w(b)$ and Eq. (4.15)
      Solve λ by Eq. (4.29)
      Update the SAM bidding function $b(\theta, r)$ by Eq. (4.24)
  end while
  return $\boldsymbol{v}$ and $b(\theta, r)$

When the risk, i.e., the variance of the profit margin, is not considered, α is set as 0. Then the campaign i with the highest $\mu_i(b)$ will always be selected, i.e., $v_i = 1$, while $v_j = 0$ for all other campaigns j.

The Overview of SAM Algorithm. Finally, the overall operations to get the optimal campaign selection probability $\boldsymbol{v}$ and the arbitrage bidding function $b(\theta, r)$ are summarised in Algorithm 4.1. In practice, $\boldsymbol{v}$ and $b(\theta, r)$ will get converged within 5 EM iterations. For E-step, the computationally costly parts are the MCMC methods for evaluating the margin of M individual campaign (Eqs. (4.31) and (4.32)), where the time complexity is $O(MNT)$, and the campaign correlation calculation ($\psi_{i,j}$ in Eq. (4.35)), which is $O(M^2NT)$. For M-step, the bidding function is derived with closed form; the calculation of λ by numeric descent methods Eq. (4.29), which depends on the data values but is normally much efficient $O(T)$. The performance in Section 4.6.3 will demonstrate the capability of the proposed solution for highly efficient re-training in dynamic arbitrage tasks.

4.5 Experimental Setup

4.5.1 Datasets

The experiments[5] are conducted based on two real-world large-scale bidding logs collected from two DSP companies.

---

[5]To make the experiment repeatable, the project code with the public dataset link has been published at `https://github.com/wnzhang/rtbarbitrage`.

iPinYou RTB dataset was published in 2014 after iPinYou's global RTB algorithm competition. This dataset contains the bidding and user feedback log from 9 campaigns during 10 days in 2013, which consists of 64.75M bid records, 19.50M impressions, 14.79K clicks and 16K CNY expense[6]. The train/test set spliting has been given by the data publisher [115], where the last three-day data of each campaign is split as the test data and the rest as the training data. More statistics and analysis of the dataset is available in Section 3.4.1 and Table 3.2.

BigTree RTB dataset is a proprietary dataset from BigTree Times Co., a mobile DSP technology company based in Beijing. This dataset is collected from Nov. 2014 to Feb. 2015 from 3 iOS mobile game campaigns. It consists of 10.85M impressions and 46.38K actions[7] with $0.083 CPA. This dataset is used to train the model and conduct online A/B test on BigTree DSP during Feb. 2015.

Both datasets are in a record-per-line format, where each line consists of three parts: (i) the features for this auction, e.g., the time, location, IP address, the URL/domain of the publisher, ad slot size, user interest segments etc.; (ii) the auction winning price, which is the threshold of the bid to win this auction; (iii) the user feedback on the ad impression, i.e., click, conversion or not.

4.5.2 Evaluation Protocol

Evaluation Procedure. The evaluation procedure adopted in this experiment is similar to the previous work on bid optimisation as described in Chapter 3. The difference from the previous evaluation procedure lies on the campaign sampling process (via v) for each incoming bid request, which is handled by following an offline evaluation scheme similar to a previous work on evaluating interactive systems [86]. As in the historic data, the user's feedback is only associated with the winning campaign of the auction, there is no corresponding user feedback if a different campaign is sampled. As such, based on the bid request i.i.d. assumption made before, for each round, the meta-bidder first samples a campaign i, then passes the next test data record of this campaign to the bid agent for bidding. If there is no more test data left for this campaign, i.e., the

---

[6]Note that this dataset is slightly different from that used in Chapter 3. This is because iPinYou published a new version of the dataset after the experimental work in Chapter 3 had been completed.

[7]According to the advertiser's contract, here the action is defined by users' successful landing on the game's page in the app store.

bid requests are run out, the test ends.

Budget Constraints. Similar to the budget setting in Section 3.4.4, it is easy to see that if the meta-bidder's budget is set as the same as the original total cost in the test log, then simply bidding as much as possible for each auction will exactly run out the budget and get all the logged clicks and conversions. In the offline empirical study, to test the performance against various budget constraints, for each campaign, the evaluation tests are conducted using $1/2, 1/4, 1/8, \ldots, 1/256$ of the original total cost in the test log as the budget, respectively.

Payoff Setting. To set up various difficulties in arbitrage, for our offline experiments, we manually set the CPA payoff for each iPinYou campaign. Specifically, for each campaign i, we set a high and a low CPA payoff in order to test the algorithms' performance under an easy and a hard arbitrage situation, denoted as r_i^{easy} and r_i^{hard}, respectively:

$$r_i^{\text{easy}} = \text{eCPA}_i \times 0.8,$$
$$r_i^{\text{hard}} = \text{eCPA}_i \times 0.2,$$

where eCPA_i is the original average cost for acquiring each conversion of campaign i in the training data without any arbitrage strategy. In addition, the sufficient conversion data in iPinYou is unavailable for 7 out of 9 campaigns. To have more tests done, we thus regard the user clicks as a proxy for the desired actions (page landing) in our offline experiment.

To complement the offline tests, in our online experiments, we directly adopt the CPA payoff specified by the real-world advertisers to test the real business case.

4.5.3 Compared Strategies

Both bidding strategies and campaign selection strategies should be compared and investigated in the empirical study.

4.5.3.1 Bidding Strategies

The following baseline and state-of-the-art bidding strategies are compared in the experiment. Their parameters are tuned on the training data.

Constant bidding (CONST). A constant bid regardless bid requests and campaigns. Although trivial, it is a simple solution used by many DSPs. The parameter is the

specific constant bid price.

Random bidding (RAND). Randomly choose a bid value in a given range. The parameter is the upper bound of the random bidding range.

Truth-telling bidding (TRUTH). If there is no budget constraint, one should bid the true value for each ad impression, which is CPA×CVR of the impression [4].

Linear bidding (LIN). In [19], the bid value is linearly proportional to the CVR with the bid scale parameter tuned to maximise the expected conversion number. This bidding strategy is widely used in industry.

Optimal real-time bidding (ORTB). This an optimal bidding strategy proposed in Chapter 3 to maximise clicks. Here the bidding strategy ORTB1 as in Eq. (3.16) is compared.

Statistical arbitrage mining (SAM_1, SAM_2). These are the two bidding strategies proposed in this chapter: SAM_1 is from Eq. (4.17) and SAM_2 is from Eq. (4.24), collectively denoted as SAMX.

SAM with competition modelling (SAM_{1C}, SAM_{2C}). In a real online environment, the advertisers will tune their bidding strategies according to their campaign performance. If many bidders adopt our SAMX bidding strategies, it is possible that this may change the market prices. In our offline empirical study, we follow [52] to adopt the OPT bidding strategy [133] to simulate the market price changes towards a locally envy-free equilibrium[8]. Note that this is not for comparing bidding strategies but for comparing auction environment where we would like to check whether our proposed SAMX algorithms would still make arbitrage profit when the market changes according to our actions. We only compare the performance of SAMX algorithms with those in the corresponding $SAMX_C$ settings.

---

[8]"Locally envy-free equilibrium" means no player in the game equilibrium would like to exchange the situation with others. The work [133] is on sponsored search with generalised second price auctions. By setting the slot number for each keyword auction as 1 and the CTR as 1.0, the OPT bidding strategy can be used for our display advertising scenario.

4.5.3.2   Campaign Selection Strategies

For campaign selection strategies, we compare the UNIFORM campaign selection, i.e., $\boldsymbol{v} = \boldsymbol{1}/M$, and the proposed PORTFOLIO-based campaign selection, i.e. the solution of Eq. (4.38), where PORTFOLIO will be denoted as GREEDY when α in Eq. (4.38) is set as 0. The conventional campaign selection scheme based on internal auctions [14], which always selects the campaign with the highest bid value, will be compared in the online A/B test in Section 4.7.

4.5.4   Evaluation Measures

We use the meta-bidder-level profit as the prime evaluation measure, which is calculated as

$$\mathtt{profit} = \#\mathtt{conversions} \times \mathtt{payoff}_{\mathtt{CPA}} - \mathtt{cost}. \qquad (4.39)$$

In addition to the profit, we also evaluate the profit margin, i.e., γ as discussed in Section 4.4 for each strategy, which is calculated by the profit divided by the cost. In addition, we report the number of impressions and conversions as well as the cost for each strategy.

4.6   Offline Empirical Study

4.6.1   Single Campaign Arbitrage

Table 4.2 reports the overall performance on the tested 9 campaigns from the iPinYou dataset. It can be observed that SAMX bidding strategies outperform all others regarding to the profit. SAM$_2$ further outperforms SAM$_1$ particularly in the hard payoff settings because of its more practical winning function. In addition, SAMX$_C$ strategies still make high arbitrage profit with the market competition modelling, which demonstrates the potential of SAMX strategies in a real market competition environment.

Furthermore, Figure 4.3 presents the performance change on the arbitrage profit and the margin of each algorithm w.r.t. the budget settings. The value on the x-axis means the proportion of the original total cost in the test data divided by the test budget. The higher the proportion is, the less the budget is. From Figure 4.3 we can have the following observations. (i) SAM$_1$ and SAM$_2$ outperform the rest in almost all the

Table 4.2: Single-campaign statistical arbitrage overall performance.

Easy payoff, 1/16 budget setting

Bid. Algo.	Profit (CNY)	Margin	Bids (M)	Imps. (K)	Cnvs.	Cost (CNY)
CONST	41.77	0.21	2.68	761.91	297	194.44
RAND	19.65	0.12	2.97	612.90	223	166.60
TRUTH	749.75	3.60	1.89	420.19	1,137	208.33
LIN	845.22	3.83	2.71	531.49	1,161	220.90
ORTB	869.43	4.03	2.87	632.38	1,172	215.78
SAM_1	1,141.72	6.02	3.26	471.46	1,504	189.55
SAM_2	**1,161.24**	5.97	3.42	606.97	1,534	194.40
SAM_{1C}	1,118.61	6.10	3.24	389.09	1,473	183.34
SAM_{2C}	1,141.01	5.87	3.41	563.74	1,513	194.38

Hard payoff, 1/16 budget setting

Bid. Algo.	Profit (CNY)	Margin	Bids (M)	Imps. (K)	Cnvs.	Cost (CNY)
CONST	-1.40	-0.25	4.10	81.55	10	5.53
RAND	1.08	10.47	4.10	8.36	4	0.10
TRUTH	214.08	2.13	4.03	373.66	1,430	100.30
LIN	45.63	0.21	2.71	531.49	1,161	220.90
ORTB	55.52	0.26	2.87	632.38	1,172	215.78
SAM_1	207.34	2.29	3.89	319.77	1,328	90.59
SAM_2	**227.76**	3.77	4.10	301.99	1,326	60.47
SAM_{1C}	204.73	2.25	3.88	308.44	1,322	90.98
SAM_{2C}	225.95	3.70	4.10	298.51	1,322	61.13

profit and margin comparisons with different budget settings. Such results indicate that the proposed statistical arbitrage mining bidding strategies are capable of making arbitrage profit and outperform all other existing bidding strategies. (ii) Under the higher budget setting, e.g., 1/2 or 1/4 of the total original spend, TRUTH produces comparable arbitrage profit as SAMX (Figures 4.3(a) and 4.3(b)). This is because when the budget is abundant, the tight budget constraint (i.e., the equality condition in Eq. (4.11)) is unnecessary to meet in order to maximise the profit. Under such a situation, the bidding problem will get back to the classic second price auction problem, where the truth-telling bidding strategy is optimal [10]. However, such abundant budget (win about half of the ad impressions in the market) is normally impossible even for large DSPs because the real daily RTB spot market is about billions of dollars. (iii) Under the lower budget setting, e.g., 1/64, 1/128 and 1/256 of the total original spend, the profit from TRUTH

(a) Profit (easy)

(b) Profit (hard)

(c) Profit margin (easy)

(d) Profit margin (hard)

(e) Competition profit drop (easy)

(f) Competition profit drop (hard)
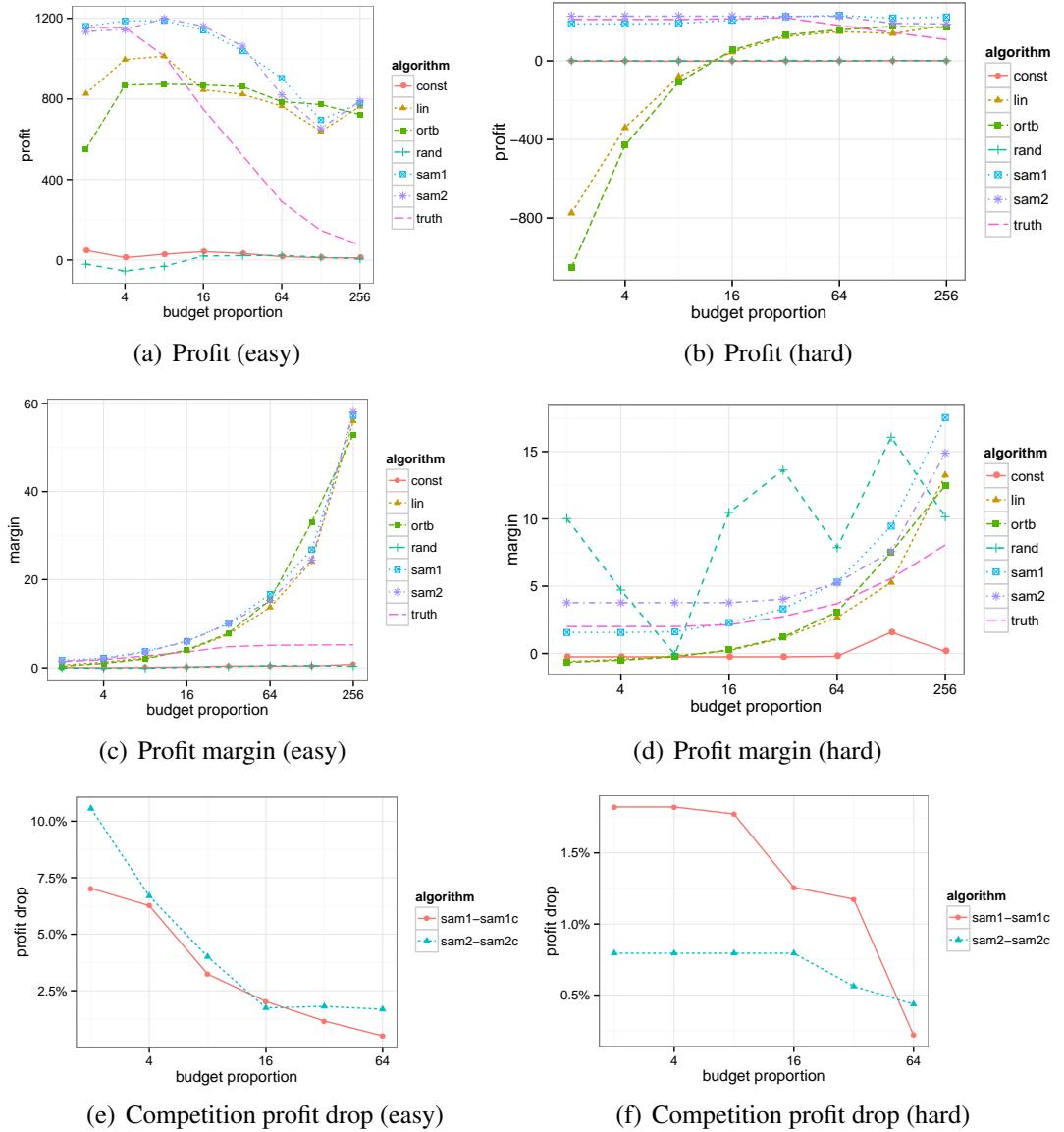
Figure 4.3: Single-campaign arbitrage performance against different settings.

drops significantly because of the budget constraint is quite important and the optimal bidding strategy is never truth-telling. On the contrary, LIN and ORTB act almost the same as SAMX. This is reasonable because under the lower budget settings, the budget is always exhausted. With the cost the same as the budget, the more conversions the more arbitrage profit. (iv) With the market competition modelling, SAMX$_C$ has dropped profit compared with SAMX but the drop is tolerable (less than 10%, Figures 4.3(e) and 4.3(f)). Specifically, when the budget gets lower, the profit drop percentage gets lower. The reason is that fewer auctions are won with lower budget so that the market does not change much.

(a) Profit (easy)

(b) Profit (hard)

(c) Competition profit drop (easy)

(d) Portfolio α tuning (easy)
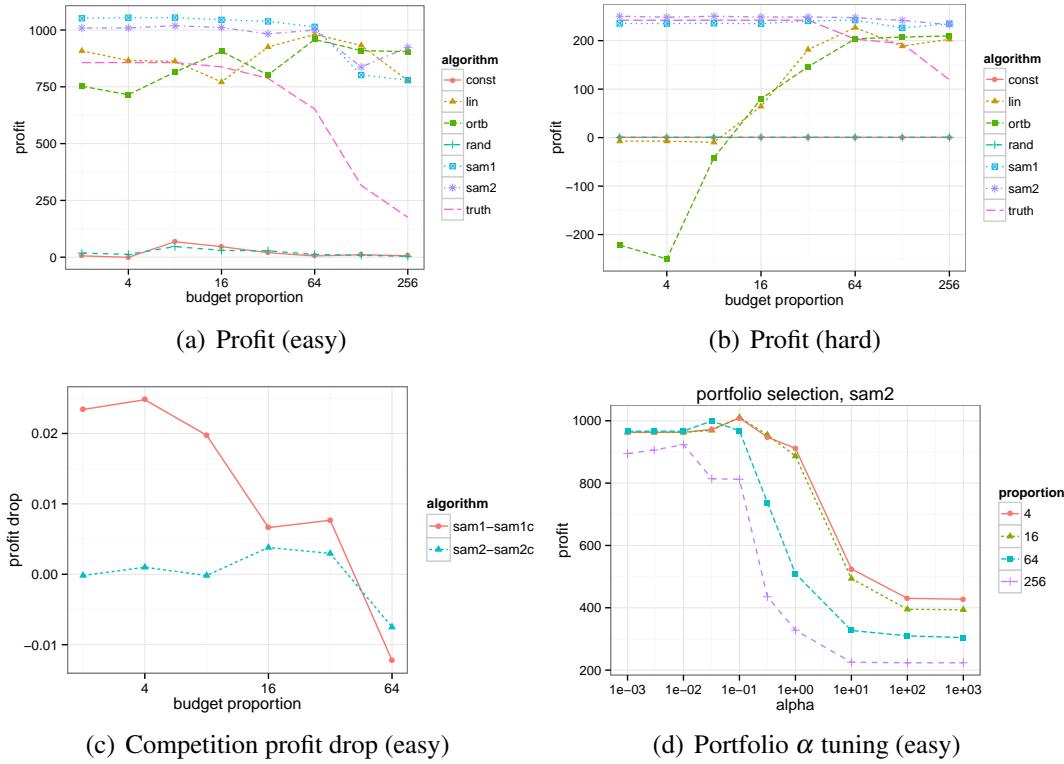
Figure 4.4: Multi-campaign arbitrage performance comparison and trend against different parameter and budget settings.

4.6.2   Multiple Campaign Arbitrage

Furthermore, the multi-campaign statistical arbitrage performance is studied. Specifically, we test 6 campaign portfolios from the iPinYou dataset. Each portfolio contains 4 campaigns with the data from the same period. For each portfolio, after the convergence of EM iterations, the empirically optimal $\boldsymbol{\nu}$ and bidding function $b(\boldsymbol{\theta}, r)$ are deployed in the campaign portfolio's test stage, where the auction volume and the budget are set as the same as in the training stage. Compared with the previous single campaign part, this part of experiment focuses more on the campaign portfolio selection, where the UNIFORM, GREEDY and PORTFOLIO selection methods are compared.

The overall results with 1/32 budget setting are reported in Table 4.3. For the comparison among the bidding strategies, SAMX overall outperforms others in both payoff settings. Figure 4.4 provides more detailed analysis. The profit trend against the budget setting, as shown in Figures 4.4(a) and 4.4(b), is consistent with the single campaign setting. The competitor model setting does not significantly drop the arbitrage profit as shown in Figure 4.4(c). Specifically, when the budget gets lower, the profit drop per-

Table 4.3: Multi-campaign statistical arbitrage overall performance.

Strategies		Easy payoff		Hard payoff	
Bidding algorithm	Campaign selection	Profit (CNY)	Margin	Profit (CNY)	Margin
LIN	GREEDY	501.12	6.63	68.59	0.91
LIN	PORTFOLIO	925.45	13.11	181.54	2.50
LIN	UNIFORM	747.00	9.53	127.14	1.62
ORTB	GREEDY	517.02	6.65	70.96	0.91
ORTB	PORTFOLIO	802.15	10.32	146.13	1.88
ORTB	UNIFORM	765.12	9.89	133.16	1.72
SAM$_1$	GREEDY	966.02	20.81	230.38	11.13
SAM$_1$	PORTFOLIO	**1,037.98**	15.84	240.63	7.96
SAM$_1$	UNIFORM	768.38	9.78	172.43	7.57
SAM$_2$	GREEDY	961.68	28.73	235.31	24.00
SAM$_2$	PORTFOLIO	983.01	17.21	**248.65**	13.61
SAM$_2$	UNIFORM	774.09	10.32	168.15	5.16
TRUTH	GREEDY	787.10	14.69	227.86	29.05
TRUTH	PORTFOLIO	787.10	14.69	242.07	18.34
TRUTH	UNIFORM	326.57	4.14	101.12	5.36

centage gets lower. The reason is that fewer auctions are won with lower budget so that the market does not change much. To compare campaign selection strategies, Table 4.3 shows that PORTFOLIO selection constantly outperforms UNIFORM and GREEDY selection. Compared with UNIFORM, GREEDY allocates all the auction volume and the budget onto the campaign evaluated as with the highest arbitrage profit margin, which theoretically maximises the expected profit. However, the result that PORTFOLIO outperforms GREEDY indicates there exists a risk-return tradeoff point which practically performs better than the maximum expectation solution. Furthermore, Figure 4.4(d) shows the change of total profit from the 6 tested campaign portfolios based on SAM$_2$ against the portfolio risk-averse parameter α in Eq. (4.38). Here setting α as a small enough value is equivalent to the greedy campaign selection. As we can see, as α increases from 10^{-3}, the profit first gets a rise to the peak value and then drops significantly. Among the different budget settings, we can observe a trend from Figure 4.4(d) that is the more budget, the higher the optimal α is. For 1/256 budget setting, the optimal α is 0.01, while 0.1 is optimal for 1/4 budget setting. This may be due to the fact that more budget brings more auction volume across a longer period, importing higher risk, which requires to be carefully hedged.

In addition, we present a case study on a campaign portfolio (3358, 3386, 3427 and

(a) Efficient frontier   (b) ν allocation w.r.t. risk
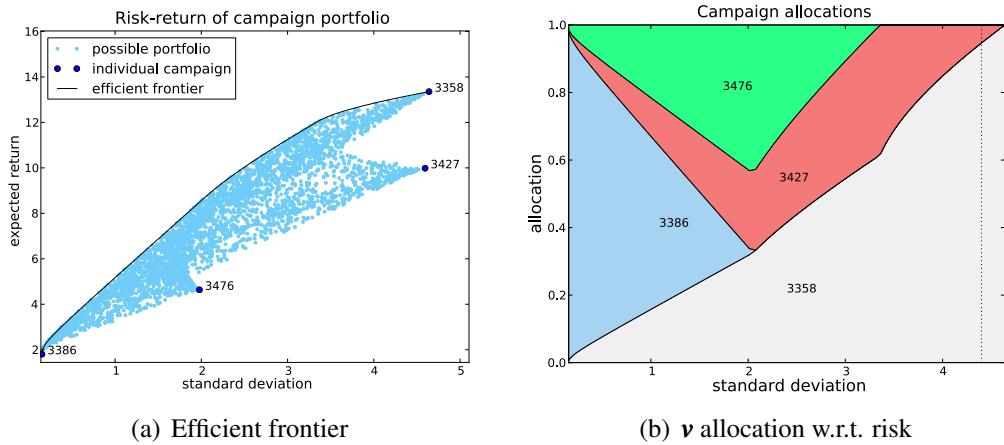
Figure 4.5: A case study of a 4-campaign portfolio.

3476 are four campaign IDs). Its risk-return analysis plot is shown in Figure 4.5(a) and the corresponding campaign selection probability allocation is shown in Figure 4.5(b). In Figure 4.5(a) the dark blue points stand for the expected profit margin and its standard deviation for 4 individual campaigns. As we can see, campaign 3358 has the highest expected margin as well as the highest risk while campaign 3386 is the most stable one but with the lowest expected margin. The best empirical portfolio selection is shown as the vertical dashed line in Figure 4.5(b), where 94.9% auction volume is allocated to campaign 3358 and 4.1% is allocated to campaign 3427. However, if the meta-bidder is more risk-averse, other two campaigns can be included in order to further reduce the standard deviation. The parameter α in Eq. (4.38) provides a flexible way to adjusting the risk and return trade-off.

4.6.3 Dynamic Multiple Campaign Arbitrage

In practice, as the market competition and the user behaviour change across the time, the meta-bidder should dynamically change its bidding strategy and campaign selection. In this part of experiment, we test the capability of our proposed SAM$_2$ bidding strategy with dynamic campaign portfolio selection over a 72 hour test period. The arbitrage bidding function and campaign selection probability are updated periodically, and the interval between two updates is referred as one round. Specifically, at the beginning of each round, the re-training of the arbitrage bidding function and campaign selection probability using Algorithm 4.1 is performed based on the bidding data collected from previous round. A problem here is that how frequent the update should be?

(a) Easy payoff setting (b) Hard payoff setting
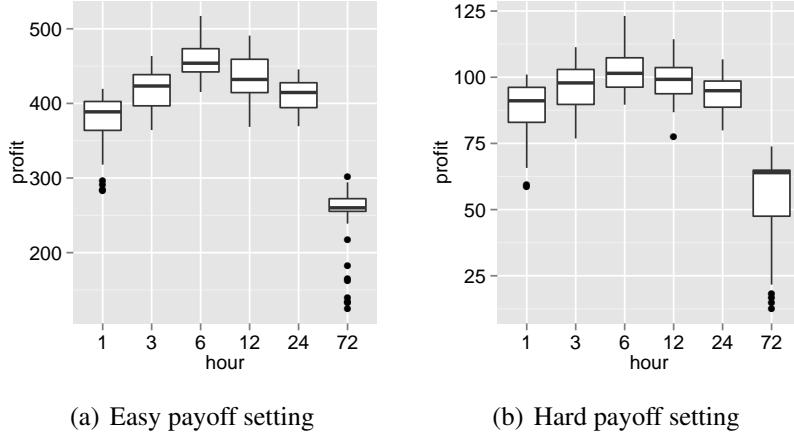
Figure 4.6: Dynamic multi-campaign arbitrage profit distribution with different update frequency.

It is apparent that if the round period is too long, it is difficult for the meta-bidder to catch the transient statistical arbitrage opportunities; if the round period is too short, the training data could be sparse and the model might overfit the data.

We test the dynamic multiple campaign arbitrage on 5 portfolios, each of which consists of 4 campaigns with the data logged within the same period. For each test campaign portfolio, we try the different update frequencies as well as different risk-averse α's. The box plots [116] of the arbitrage profit distribution with different update frequencies under two payoff settings are shown in Figure 4.6. From the results we can observe that (i) the positive profit values over all cases demonstrate the capability of SAM$_2$ to make dynamic arbitrages via periodic re-training. (ii) In both payoff settings, the dynamic SAMs (period no more than 24 hours) yield much better performance than the static SAM (period equals to 72 hours, i.e., only one update), which indicates the importance of dynamically re-training the models to catch the latest market situation. (ii) Among the different frequencies of dynamic updating, updating every 6 hours leads to the highest arbitrage profit. We believe this is a trade-off point between the abundance and recency of the training data. Note that the optimal update frequency may be different for other campaigns or different training settings.

In addition, Figure 4.7 presents a case study of the 72 hour dynamic 4-campaign arbitrage with the model update for every 6 hours. In each round, the calculated campaign selection probability (i.e., volume allocation) from portfolio optimisation, the estimated profit margin of each campaign, the empirical profit and cost are depicted. It
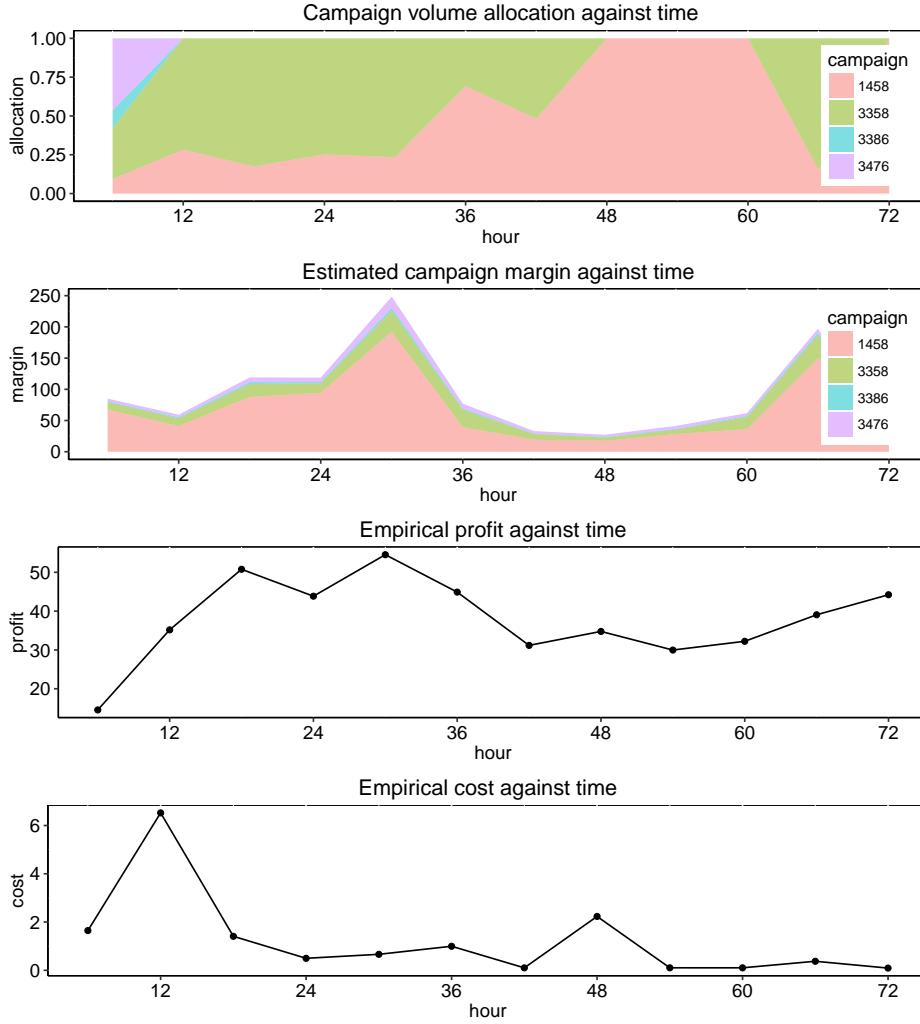
Figure 4.7: A case study of dynamic multi-campaign arbitrage performance and the corresponding margin estimation and volume allocation.

can be observed that the estimated margin for each campaign varies over time, which results in the change of campaign volume allocation across the time. The empirical profit shows the same trend with the estimated campaign margin, which to-some-extent highlights the effectiveness of the margin estimation in our model. Moreover, the cost in each round (i.e., 6 hours) is different, not necessarily be the average budget allocated for each round. It is possible that if the market is too competitive to make arbitrage profit, the resulting cost and profit could be both much low.

4.7   Online A/B Test

The proposed SAM algorithm has been deployed and tested in a live environment provided by BigTree DSP. The model training follows the scheme in Section 4.6.2. Specif-
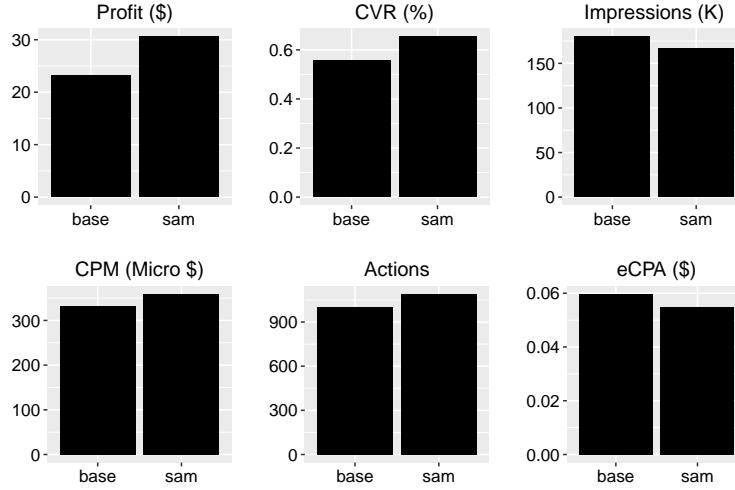
Figure 4.8: Online A/B test performance on BigTree DSP.

ically, with Algorithm 4.1, we obtain the empirically optimal SAM$_2$ bidding function $b(\theta, r)$ and campaign selection probability ν for the meta-bidder based on the 3-campaign training data described in Section 4.5.1, where the hyperparameter α in Eq. (4.38) is set as 0.1. As a control baseline, we deploy another meta-bidder with the basic linear bidding function [19, 4] and the internal auction-based campaign selection scheme [14]. During the online A/B test, every received bid request from the router of BigTree DSP will be randomly assigned to either of the two meta-bidders, which returns the bid response, including the selected campaign ad and the bid price, back to the ad exchange for auction. The online test was conducted during 23 hours between 13 and 14 Feb. 2015 with $60 budget for each meta-bidder.

Figure 4.8 presents the overall online performance of SAM and the baseline algorithm BASE. The online results on the commercial DSP verify the effectiveness of our algorithm in a real commercial setting: SAM leads to $30.6 arbitrage profit with $60 budget, which is a 51.1% margin and a 31.8% improvement over the BASE bidder setting. An interesting observation is that in spite of the higher CPM, SAM brings lower eCPA than BASE, which ultimately leads to higher arbitrage profit. This suggests that despite the market price and arbitrage margin are different across the campaigns, the SAM algorithm would be able to successfully identify and target to the cases that have higher arbitrage margin from those high-cost impressions (reflected by their high CPM).

4.8   Summary

In this chapter, we investigated a problem of strategy optimisation over multiple campaigns. Specifically, with a special case of the RTB business between the CPA ad campaigns and CPM ad inventories, the first study on statistical arbitrage mining in RTB display advertising has been conducted. We proposed a joint optimisation framework to maximise a multi-campaign meta-bidder's expected arbitrage profit with budget and risk constraints, which was then solved in an EM fashion. In the E-step the bid volume was reallocated according to the individual campaign's estimated risk and return, while in the M-step the arbitrage bidding function was optimised to maximise the expected arbitrage profit with the campaign volume allocation. Aside from the theoretical insights, the offline and online large-scale experiments with real-world data demonstrated the effectiveness of the proposed solution in exploiting arbitrage in various model settings and market environments. We believe this research would open up a whole new set of research questions that intersect between financial methods such as high-frequency trading [82], risk-management [65, 68] and data mining methodologies for display advertising and beyond.

Chapter 5

Unbiased Learning and Optimisation on Censored Auction Data

5.1 Background and Motivations

The rise of real-time bidding (RTB) based display advertising and behavioural targeting provides one of the most significant cases for machine learning applied to big data. The major supervised learning tasks range from predicting the market price distribution and volume of a given ad impression type [24], estimating the click-through rate (CTR) [134] and conversion rate [4], to the optimisation of bidding strategies [19, 128]. These data driven prediction and optimisation techniques enable ads to be more relevant and targeted to the underlying audience [128].

A challenging yet largely neglected problem in the aforementioned learning tasks is that the common supervised learning requires the training and prediction data to follow the same distribution, but in the online display advertising case, the training data is heavily censored by the ad auction selection, i.e., the process of auction selecting ads [135]. For advertisers, specifically, the above learning models, e.g., CTR estimation and bid optimisation, are operated over the full volume bid request stream in order to evaluate each potential impression and automatically generate the bid [29]. However, the auction selects the ad with the highest bid and displays it to the user, and only in this situation the corresponding user feedback, i.e., click and conversion, to this ad impression, along with the second price (or market price [28]) for this auction, are received by the advertisers as the labels of this data instance. Thus, as illustrated in Figure 5.1, the observation of a training instance is heavily influenced by its bid value; data instances
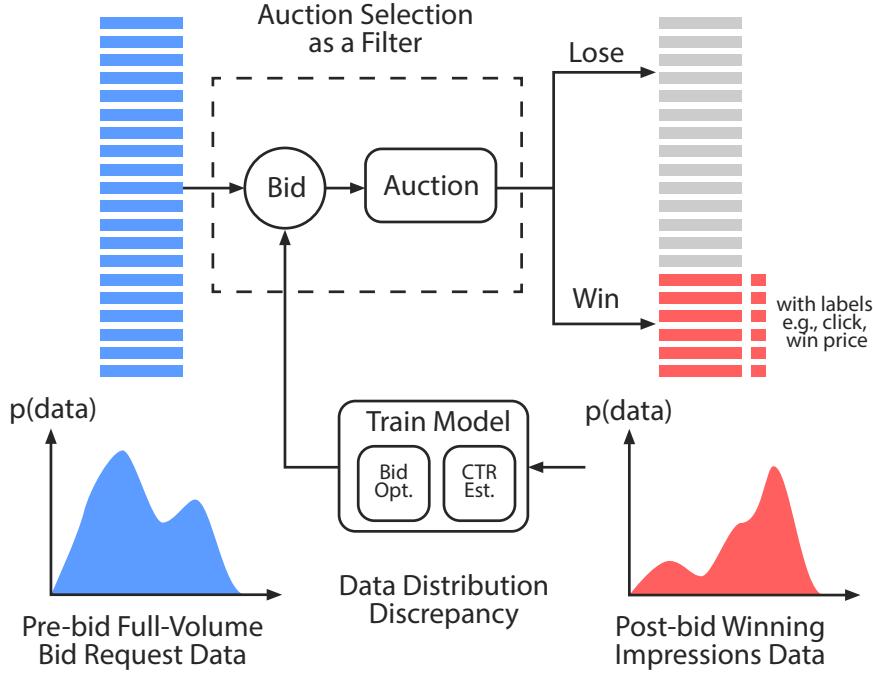
Figure 5.1: From an advertiser's perspective, the ad auction selection acts as a dynamic data filter based on bid value, which leads to distribution discrepancy between the post-bid training data (red) and pre-bid prediction data (blue).

with higher bid prices (than the expected market price) would generate a higher probability of winning and thus higher chance to be in the training data. A consequence is that the learning will be overly focused on the instances with a high winning probability (high bid), while neglecting the cases where the probability is small. Such a bias is problematic as intuitively conversions or clicks from those low market-valued impressions are more crucial than those from high market-valued impressions in order to obtain a more economic solution. Ultimately advertisers not only need to identify the impressions that have high chances to be clicks/converted, but also (and equally importantly) require the cost of winning those impressions to be relatively small. Thus, we need to have an unbiased learning framework that can take the final optimisation objective into account.

Typically, the bias problem is a missing data problem, which has been well-studied in the machine learning literature [90]. A direct solution would be to identify or assume the missing process and correct the discrepancy (e.g, [91, 92]) during the training. However, the data missing in RTB display advertising depends on both the advertiser's previous bidding strategy and the market competition, neither of them are known as a priori. There are some indirect solutions of alleviating the data bias such as by adding

random ad selection probability in the bidding strategy [94], but a better solution would be to decouple the solution with the previously employed bidding strategy (when acquiring the training data) and build a link to the final optimisation process.

In this chapter, we consider both CTR estimation [134, 4] and bid optimisation tasks [19, 128] and propose a flexible learning framework that eliminates such an auction-generated data bias towards a better learning and optimisation performance. According to the RTB auction mechanism, the labelled training data instance is observed only when the bid is higher than the market price. Inspired by the censored learning work [28], we explicitly model the auction winning probability with a bid landscape based on a non-parametric survival model [136], which is then estimated from the advertiser's historic bids. By importance sampling with the auction winning probabilities as propensity scores [137], we naturally incorporate it into the gradient derivation to produce a Bid-aware Gradient Descent (BGD) training scheme for both CTR prediction and bid optimisation tasks. Intuitively, our BGD shows that (i) the higher bid price the impression was won with, the lower valued gradient such data should generate; (ii) to generate a bid, historic bids will further adjust the gradient direction and provide a lower average budget for lower-bidden training instance when learning the bidding function. It is worth noticing that the proposed learning framework is generally applicable to various supervised learning and optimisation tasks mentioned above.

Besides the theoretical derivations, we also conduct empirical studies with the tasks of CTR estimation and bid optimisation on two large-scale real-world datasets. The results demonstrate large improvements brought from our solution over the start-of-the-art models. Moreover, the learning framework was also deployed on Yahoo! DSP in Sep. 2015 and brought 2.97% AUC lift for CTR estimation and 9.30% eCPC drop for bid optimisation over 9 campaigns in an online A/B test.

5.2 Unbiased Learning and Optimisation Framework

In online RTB display advertising, a bid request can be represented as a high dimensional feature vector [4]. Let us denote the vector as $\boldsymbol{x}$. Without loss of generality, we regard the bid requests as generated from an i.i.d. $\boldsymbol{x} \sim p_x(\boldsymbol{x})$ within a short period [128]. Based on the bid request $\boldsymbol{x}$, the ad agent (or demand-side platform, a.k.a. DSP)

will then provide a bid b_x following a bidding strategy. If such a bid wins the auction, the corresponding labels, i.e., user response y (either click or conversion) and market price z, are observed. Thus, the probability of a data instance (x, y, z) being observed relies on whether the bid b_x would win or not and we denote it as $P(\text{win}|x, b_x)$. Formally, this generative process of creating each observed instance of the training data $D = \{(x, y, z)\}$ is summarised as:

$$\underbrace{q_x(x)}_{\text{impression}} = \underbrace{P(\text{win}|x, b_x)}_{\text{auction selection}} \cdot \underbrace{p_x(x)}_{\text{bid request}}, \tag{5.1}$$

where the probability $q_x(x)$ describes how feature vector x is distributed within the training data.[1] The above equation indicates the relationship (bias) between the p.d.f. of the pre-bid full-volume bid request data (prediction) and the post-bid winning impression data (training); in other words, the predictive models would be trained on D, where $x \sim q_x(x)$, and be finally operated on the prediction data $x \sim p_x(x)$. In the following sections, we shall focus on the estimation of the winning probability $P(\text{win}|x, b_x)$ and then introduce our solutions of using it for creating bid-aware gradients to solve CTR estimation and bid optimisation problems.

5.2.1   Auction Winning by Survival Models

The RTB display advertising uses the second price auction [14]. In the auction, the market price z is defined as the second highest bid from the competitors for an auction. In other words, it is the lowest bid value one should have in order to win the auction. Following [28], we take a stochastic approach rather than a game theoretical one, and assume the market price z is a random variable generated from a fixed yet unknown p.d.f. $p_z^x(z)$; then the auction winning probability is the probability when the market price z is lower than the bid b_x:

$$w(b_x) \equiv P(\text{win}|x, b_x) = \int_0^{b_x} p_z^x(z) dz, \tag{5.2}$$

where to simplify the solution and reduce the sparsity of the estimation, the market price distribution is estimated on a campaign level rather than per impression x [24].

---

[1]The exact formula should be $q_x(x) \propto P(\text{win}|x, b_x) p_x(x)$. Here we omit the normaliser of $q_x(x)$ for formula simplicity.

Thus for each campaign, there is a $p_z(z)$ to estimate, resulting in the simplified winning function $w(b_{\boldsymbol{x}})$, similar to [28] and Chapters 3 and 4.

If we assume there is no data censorship, i.e., the ad agent wins all the bid requests and observes all the market prices, the winning probability $w_o(b_{\boldsymbol{x}})$ can directly come from the observation counts:

$$w_o(b_{\boldsymbol{x}}) = \frac{\sum_{(\boldsymbol{x}',y,z) \in D} \delta(z < b_{\boldsymbol{x}})}{|D|}, \tag{5.3}$$

where z is the historic market price of the bid request $\boldsymbol{x}'$, the indicator function $\delta(z < b_{\boldsymbol{x}}) = 1$ if $z < b_{\boldsymbol{x}}$ and 0 otherwise. We use it as a baseline of $w(b_{\boldsymbol{x}})$ modelling.

However, the above treatment is rather problematic as it does not take into account that in practice there are always a large portion of the auctions the advertiser loses ($z \geq b_{\boldsymbol{x}}$), in which the market price is not observed in the training data[2]. Thus, the observations of the market price are *right-censored*: when we lose, we only know that the market price is higher than our bid, but do not know its exact value. In fact, $w_o(b_{\boldsymbol{x}})$ is a biased model and over-estimates the winning probability. One way to look at this is that it ignores the counts for lost auctions where the historic bid price is higher than $b_{\boldsymbol{x}}$ (in this situation, the market price should have been higher than the historic bid price and thus higher than $b_{\boldsymbol{x}}$) in the denominator of Eq. (5.3). As we will show in our experiment, this estimator will consistently over-estimate the actual winning probability.

In this chapter, we use survival models [138] to handle the biased auction data. Survival models were originally proposed to predict patients' survival rate for given a time after certain treatment. As some patients might leave the investigation, researchers do not know their exact final survival period but only know the period is longer than the investigation period. Thus the data is right-censored. The auction scenario is quite similar, where the integer market price[3] is regarded as the patient's underlying survival period from low to high and the bid price as the investigation period from low to high. If the bid b wins the auction, the market price z is observed, which is analogous to the

---

[2]In the iPinYou dataset [29] we tested, the overall auction winning rate of 9 campaigns is 23.8%, which is already a very high rate. In practice, a common auction winning rate for a DSP is lower than 10%.

[3]The mainstream ad exchange auctions require integer bid prices. Without a fractional component, it is reasonable to analogise bid price to survival days.

observation of the patient's death on day z. If the bid b loses the auction, one only knows the market price z is higher than b, which is analogous to the patient's left from the investigation on day b.

Specifically, we follow [28] by leveraging the non-parametric Kaplan-Meier Product-Limit method [136] to estimate the market price distribution $p_z(z)$ based on the observed impressions and the lost bid requests.

Suppose there is a campaign that has participated in N RTB display ad auctions. Its bidding log is a list of N tuples $\langle b_i, w_i, z_i \rangle_{i=1...N}$, where b_i is the bid price of this campaign in the auction i, w_i is the boolean value of whether this campaign won the auction i, and z_i is the corresponding market price if $w_i = 1$. The problem is to model the probability of winning an ad auction $w(b_{\boldsymbol{x}})$ with the bid price $b_{\boldsymbol{x}}$.

If we transform our data into the form of $\langle b_j, d_j, n_j \rangle_{j=1...M}$, where the bid price $b_j < b_{j+1}$. d_j denotes the number of ad auction winning cases with the market price exactly valued $b_j - 1$ (in analogy to patients die on day b_j). n_j is the number of ad auction cases which cannot be won with bid price $b_j - 1$ (in analogy to patients survive to day b_j), i.e., the number of winning cases with the observed market price no lower than $b_j - 1$[4] plus the number of lost cases when the bid is no lower than $b_j - 1$. Then with bid price $b_{\boldsymbol{x}}$, the probability of losing an ad auction is

$$l(b_{\boldsymbol{x}}) = \prod_{b_j < b_{\boldsymbol{x}}} \frac{n_j - d_j}{n_j}, \tag{5.4}$$

which just corresponds to the probability a patient survives from day 1 to day $b_{\boldsymbol{x}}$. Thus the winning probability will be

$$w(b_{\boldsymbol{x}}) = 1 - \prod_{b_j < b_{\boldsymbol{x}}} \frac{n_j - d_j}{n_j}. \tag{5.5}$$

Table 5.1 gives an example of transforming the historic $\langle b_i, w_i, z_i \rangle$ data into the survival model data $\langle b_j, d_j, n_j \rangle$ and the corresponding winning probabilities calculated by Eqs. (5.5) and (5.3). We see that the Kaplan-Meier Product-Limit model, which is a non-parametric maximum likelihood estimator of the data [139], makes use of all winning and losing data to estimate the winning probability of each bid, whereas the

---

[4]We assume that if there is tie in the auction, the campaign will not win the auction.

Table 5.1: An example of data transformation of 8 instances with the bid price between 1 and 4. Left: tuples of bid, win and cost $\langle b_i, w_i, z_i \rangle_{i=1...8}$. Right: the transformed survival model tuples $\langle b_j, d_j, n_j \rangle_{j=1...4}$ and the calculated winning probabilities. Here we also provide a calculation example of $n_3 = 4$ shown as blue in the right table. The counted cases of n_3 in the left table are 2 winning cases with $z \geq 3 - 1$ and the 2 lost cases with $b \geq 3$, shown highlighted in blue color.

b_i	w_i	z_i
2	win	1
3	win	2
2	lose	$\times$
3	win	1
3	lose	$\times$
4	lose	$\times$
4	win	3
1	lose	$\times$

b_j	n_j	d_j	$\frac{n_j - d_j}{n_j}$	$w(b_j)$	$w_o(b_j)$
1	8	0	1	$1 - 1 = 0$	0
2	7	2	$\frac{5}{7}$	$1 - \frac{5}{7} = \frac{2}{7}$	$\frac{2}{4}$
3	4	1	$\frac{3}{4}$	$1 - \frac{5}{7}\frac{3}{4} = \frac{13}{28}$	$\frac{3}{4}$
4	2	1	$\frac{1}{2}$	$1 - \frac{5}{7}\frac{3}{4}\frac{1}{2} = \frac{41}{56}$	$\frac{4}{4}$

observation-only counting model $w_o(b_x)$ does not. As we can see in the table $w_o(b_x)$ is consistently higher than $w(b_x)$. Later in experiment, we will further demonstrate such comparisons with real-world data in Figure 5.5.

5.2.2 Task 1: CTR Estimation

Generally, given a training dataset $D = \{(x, y, z)\}$, where the data instance x follows the training data distribution $q_x(x)$, (the red data distribution in Figure 5.1), an *unbiased* supervised learning problem can be formalised into a loss-minimisation problem on prediction data distribution $p_x(x)$ (the blue data distribution in Figure 5.1):

$$\min_{\beta} \ \mathbb{E}_{x \sim p_x(x)}[\mathscr{L}(y, f_\beta(x))] + \lambda \Phi(\beta), \tag{5.6}$$

where $f_\beta(x)$ is β-parametrised prediction model to be learned; $\mathscr{L}(y, f_\beta(x))$ is the loss function based on the ground truth y and the prediction $f_\beta(x)$; $\Phi(\beta)$ is the regularisation term that penalises the model complexity; λ is the regularisation weight. With Eqs. (5.1) and (5.2), one can use importance sampling to reduce the bias of the training data:

$$\mathbb{E}_{x \sim p_x(x)}\left[\mathscr{L}(y, f_\beta(x))\right] = \int_x p_x(x)\mathscr{L}(y, f_\beta(x))dx = \int_x q_x(x)\frac{\mathscr{L}(y, f_\beta(x))}{w(b_x)}dx \tag{5.7}$$

$$= \mathbb{E}_{x \sim q_x(x)}\left[\frac{\mathscr{L}(y, f_\beta(x))}{w(b_x)}\right] = \frac{1}{|D|}\sum_{(x,y,z)\in D}\frac{\mathscr{L}(y, f_\beta(x))}{w(b_x)} = \frac{1}{|D|}\sum_{(x,y,z)\in D}\frac{\mathscr{L}(y, f_\beta(x))}{1 - \prod_{b_j < b_x}\frac{n_j - d_j}{n_j}}.$$

Based on this framework, if we obtain the auction winning probability $w(b_x)$, e.g., Eq. (5.5), we can eliminate the bias for each observed training data instance. Let us look at the case of CTR estimation with logistic regression [1]. With the logistic loss between the binary click label $\{-1, +1\}$ and the predicted probability and L2 regularisation, the framework of Eq. (5.7) is written as

$$\min_{\boldsymbol{\beta}} \frac{1}{|D|} \sum_{(\boldsymbol{x}, y, z) \in D} \frac{\log(1 + e^{-y\boldsymbol{\beta}^T \boldsymbol{x}})}{w(b_x)} + \frac{\lambda}{2} ||\boldsymbol{\beta}||_2^2, \qquad (5.8)$$

where the winning probability $w(b_x)$ is estimated for each observation instance, which is independent from the CTR estimation parameter $\boldsymbol{\beta}$; the update rule of $\boldsymbol{\beta}$ is routine using stochastic gradient descent. With the learning rate η, the derived *Bid-aware Gradient Descent* (BGD) calculation of Eq. (5.8) is

$$\boldsymbol{\beta} \leftarrow (1 - \eta \cdot \lambda)\boldsymbol{\beta} + \frac{\eta \cdot y \cdot e^{-y\boldsymbol{\beta}^T \boldsymbol{x}} \cdot \boldsymbol{x}}{(1 + e^{-y\boldsymbol{\beta}^T \boldsymbol{x}})(1 - \prod_{b_j < b_x} \frac{n_j - d_j}{n_j})}. \qquad (5.9)$$

Discussion. From the equation above, we observe that with a lower winning bid b_x, the probability $1 - \prod_{b_j < b_x} \frac{n_j - d_j}{n_j}$ of observing the instance in the training set is lower, and the corresponding gradient from the data instance is higher and vice versa as it is in the denominator.

This is intuitively correct as when a data instance $\boldsymbol{x}$ is observed with low probability, e.g., 10%, we can infer there are 9 more such a kind of data instances missed because of auction losing. Thus the training weight of $\boldsymbol{x}$ should be multiplied by 10 in order to recover statistics from the full-volume data. By contrast, if the winning bid is extremely high, which leads to 100% auction winning probability, then such data is observed from the true data distribution. Thus there will be no gradient reweighting on this data. Such a nonlinear relationship has been well captured in our model in the gradient updates, as illustrated in Figure 5.2.

5.2.3   Task 2: Bid Optimisation

Another important problem in online advertising is bid optimisation, i.e. to find the optimal bidding strategy to maximise a campaign KPI, restricted by the campaign budget. Essentially, the bidding function is abstracted as a function mapping from the estimated
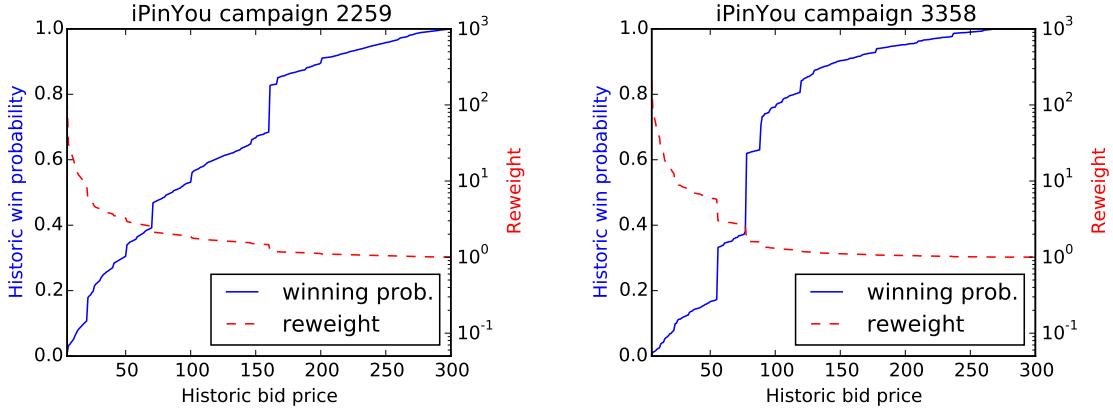
Figure 5.2: Winning probability and reweighting term in Eq. (5.9) against historic bid price.

CTR $f(\boldsymbol{x})$ to the bid price $b(f(\boldsymbol{x}))$.[5] According to Chapter 3, it is a functional optimisation problem:

$$\underset{b()}{\arg\max} \ T \int_{\boldsymbol{x}} f(\boldsymbol{x})w(b(f(\boldsymbol{x})))p_x(\boldsymbol{x})d\boldsymbol{x} \tag{5.10}$$

$$\text{subject to} \ T \int_{\boldsymbol{x}} b(f(\boldsymbol{x}))w(b(f(\boldsymbol{x})))p_x(\boldsymbol{x})d\boldsymbol{x} = B.$$

With the auction selection, the observed data distribution is actually $q_x(\boldsymbol{x})$. By Eq. (5.1), Eq. (5.10) is written as

$$\underset{b()}{\arg\max} \ T \int_{\boldsymbol{x}} f(\boldsymbol{x})w(b(f(\boldsymbol{x})))\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}d\boldsymbol{x} \tag{5.11}$$

$$\text{subject to} \ T \int_{\boldsymbol{x}} b(f(\boldsymbol{x}))w(b(f(\boldsymbol{x})))\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}d\boldsymbol{x} = B.$$

Note that $w(b_{\boldsymbol{x}})$ is different from $w(b(f(\boldsymbol{x})))$, where $b_{\boldsymbol{x}}$ is the historic bid price for the bid request $\boldsymbol{x}$ while $b(f(\boldsymbol{x}))$ is the bid price we want to optimise.

The Lagrangian is

$$\mathscr{L}(b(f),\lambda) = T \int_{\boldsymbol{x}} f(\boldsymbol{x})w(b(f(\boldsymbol{x})))\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}d\boldsymbol{x} \tag{5.12}$$

$$- \lambda T \int_{\boldsymbol{x}} b(f(\boldsymbol{x}))w(b(f(\boldsymbol{x})))\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}d\boldsymbol{x} + \lambda B,$$

According to the derivation in Chapter 3, the Euler-Lagrangian condition of

---

[5]We drop the CTR estimation parameter $\boldsymbol{\beta}$ here as it is not the parameter to optimise in this task.

Eq. (5.11) is

$$f(\boldsymbol{x})\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}\frac{\partial w(b(f(\boldsymbol{x})))}{\partial b(f(\boldsymbol{x}))} - \lambda\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}\left[w(b(f(\boldsymbol{x}))) + b(f(\boldsymbol{x}))\frac{\partial w(b(f(\boldsymbol{x})))}{\partial b(f(\boldsymbol{x}))}\right] = 0 \quad (5.13)$$

$$\Rightarrow \lambda w(b(f(\boldsymbol{x}))) = \left[f(\boldsymbol{x}) - \lambda b(f(\boldsymbol{x}))\right]\frac{\partial w(b(f(\boldsymbol{x})))}{\partial b(f(\boldsymbol{x}))}, \quad (5.14)$$

where we see that the optimal bidding function $b(f(\boldsymbol{x}))$ depends on the winning function $w(b)$. For example, if

$$w(b(f(\boldsymbol{x}))) = \frac{b(f(\boldsymbol{x}))}{c + b(f(\boldsymbol{x}))}, \quad (5.15)$$

where c is a constant, then the corresponding optimal bidding function is

$$b_{\text{ORTB}}(f(\boldsymbol{x})) = \sqrt{\frac{c}{\lambda}f(\boldsymbol{x}) + c^2} - c. \quad (5.16)$$

For the solution of λ, the Euler-Lagrangian condition w.r.t. λ is

$$\frac{\partial \mathcal{L}(b(f(\boldsymbol{x})), \lambda)}{\partial \lambda} = 0 \quad (5.17)$$

$$\Rightarrow \int_{\boldsymbol{x}} b(f(\boldsymbol{x}), \lambda) w(b(f(\boldsymbol{x}), \lambda))\frac{q_x(\boldsymbol{x})}{w(b_{\boldsymbol{x}})}d\boldsymbol{x} = \frac{B}{T} \quad (5.18)$$

$$\Rightarrow \frac{1}{|D|}\sum_{(\boldsymbol{x},y,z)\in D} b(f(\boldsymbol{x}), \lambda)\frac{w(b(f(\boldsymbol{x}), \lambda))}{w(b_{\boldsymbol{x}})} = \frac{B}{|D|}. \quad (5.19)$$

The numeric solution of λ is highly efficient. A feasible solution of Eq. (5.19) is to minimise the squared distance between the cost of each data instance and the averaged budget:

$$\min_{\lambda} \sum_{(\boldsymbol{x},y,z)\in D} \frac{1}{2}\left(\frac{b(f(\boldsymbol{x}), \lambda)w(b(f(\boldsymbol{x}), \lambda))}{1 - \prod_{b_j < b_{\boldsymbol{x}}}\frac{n_j - d_j}{n_j}} - \frac{B}{|D|}\right)^2. \quad (5.20)$$

As $b(f(\boldsymbol{x}), \lambda)$ always monotonically decreases w.r.t. λ and $w(b_{\boldsymbol{x}})$ monotonically increases w.r.t. $b(f(\boldsymbol{x}), \lambda)$, the objective of Eq. (5.20) is convex w.r.t. λ, which makes
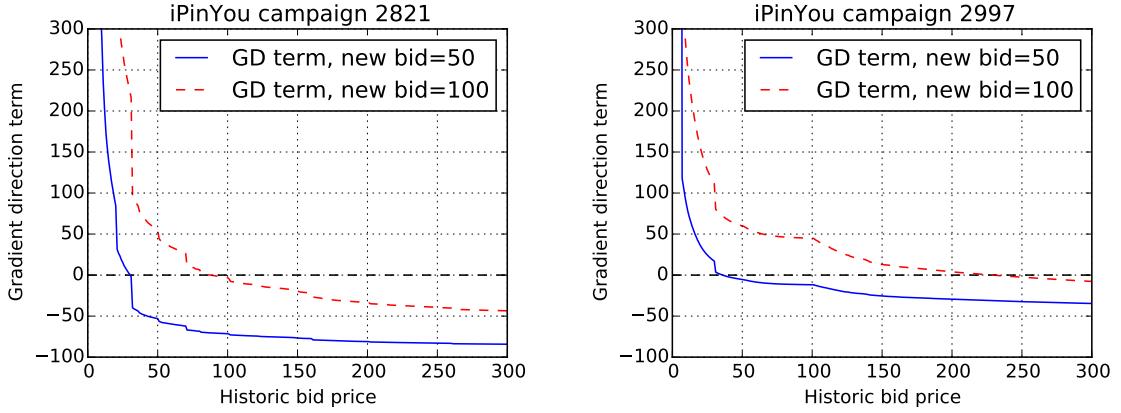
Figure 5.3: The gradient direction term in Eq. (5.21) against historic bid price $b_{\boldsymbol{x}}$ with two new bids $b(f(\boldsymbol{x}), \lambda)$.

the solution of λ easy to obtain. The BGD to solve λ is via updating

$$
\lambda \leftarrow \lambda - \eta \overbrace{\frac{1}{1 - \prod_{b_j < b_{\boldsymbol{x}}} \frac{n_j - d_j}{n_j}}}^{\text{instance reweighting}} \Bigg(\overbrace{\frac{b(f(\boldsymbol{x}), \lambda) w(b(f(\boldsymbol{x}), \lambda))}{1 - \prod_{b_j < b_{\boldsymbol{x}}} \frac{n_j - d_j}{n_j}} - \frac{B}{|D|}}^{\text{gradient direction}} \Bigg).
$$
$$
\Bigg(\underbrace{\frac{\partial b(f(\boldsymbol{x}), \lambda)}{\partial \lambda} w(b(f(\boldsymbol{x}), \lambda)) + b(f(\boldsymbol{x}), \lambda) \frac{\partial w(b(f(\boldsymbol{x}), \lambda))}{\partial \lambda}}_{\text{bidding function gradient}} \Bigg). \qquad (5.21)
$$

Discussion. Highlighted in Eq. (5.21), there are two factors related with the historic bid for updating λ. (i) The instance reweighting, similar with Eq. (5.9): a small historic bid $b_{\boldsymbol{x}}$ would generate a large weight, amplifying the importance of the training instance. (ii) The historic bid of the training instance also has an impact on the gradient direction, evidenced by the second factor of the update in Eq. (5.21).

The parameter λ converges when the second factor becomes zero. The ratio $B/|D|$ would ensure the budget to be allocated evenly across the new bids. The ratio between the winning rate of the new bid price $w(b(f(\boldsymbol{x}), \lambda))$ and that of the historic bid $1 - \prod_{b_j < b_{\boldsymbol{x}}} (1 - d_j/n_j)$ would adjust the discrepancy of the probability of seeing the impression in the training and that in the prediction.

To further understand this, Figure 5.3 illustrates the second factor (the gradient direction term) in Eq. (5.21) against historic bid price $b_{\boldsymbol{x}}$ on two sample campaigns with two new bids ($b(f(\boldsymbol{x}), \lambda) = 50$ and 100). We observe that when the historic bid is small, the gradient direction is more likely to stay positive and higher λ value in order to decrease the bid (as the bidding function gradient term in Eq. (5.21) is always

negative). For example, for a data instance with the a low historic bid $b_{\boldsymbol{x}}$, the probability of observing the data instance is low, which means that there are more similar or the same data instances that are missing in the training. Thus the bid should be lower to avoid overspending on such instances when running full-volume data. In addition, if the new bid price $b(f(\boldsymbol{x}), \lambda)$ is high, then the optimal bid price $b(f(\boldsymbol{x}), \lambda)$ should be lower to avoid budget overspending in full-volume data, which is reflected on the positive value of the gradient direction factor to make λ higher and $b(f(\boldsymbol{x}), \lambda)$ lower.

Note that with the pre-calculated reweighting factor $1/w(b_{\boldsymbol{x}}) = 1/(1 - \prod_{b_j < b_{\boldsymbol{x}}} \frac{n_j - d_j}{n_j})$, it is highly efficient to calculate the above BGD updating and solve λ.

5.3   Experimental Setup

5.3.1   Datasets

Two real-world datasets are used in our repeatable offline empirical study[6]: iPinYou and TukMob.

iPinYou runs the largest DSP in China. The publicly available[7] iPinYou dataset consists of 64.75M bid records, 19.50M impressions, 14.79K clicks and 16K CNY expense on 9 conventional display ad campaigns from different advertisers during 10 days in 2013. According to iPinYou [115], the last 3-day data for each campaign is set as test data while the rest is training data.

TukMob is a major DSP focusing on mobile game and video display ads in China. TukMob dataset is our proprietary dataset which consists of 3.00M impressions, 96.45K clicks and 2.51K CNY expense on 63 campaigns in a video display ad market from Feb. to Aug. 2015. The first 5/6 data in the time sequence is set as training data while the rest is test data.

Each data instance of both datasets can be represented as a triple $(\boldsymbol{x}, y, z)$, where y is the user click binary feedback, z is the historic winning price of the auction, and $\boldsymbol{x}$ is the bid request and ad features of that auction. The auction features contain the

---

[6]The repeatable experiment code has been published at `https://github.com/wnzhang/rtb-unbiased-learning`.

[7]Dataset link: `http://data.computational-advertising.org`.
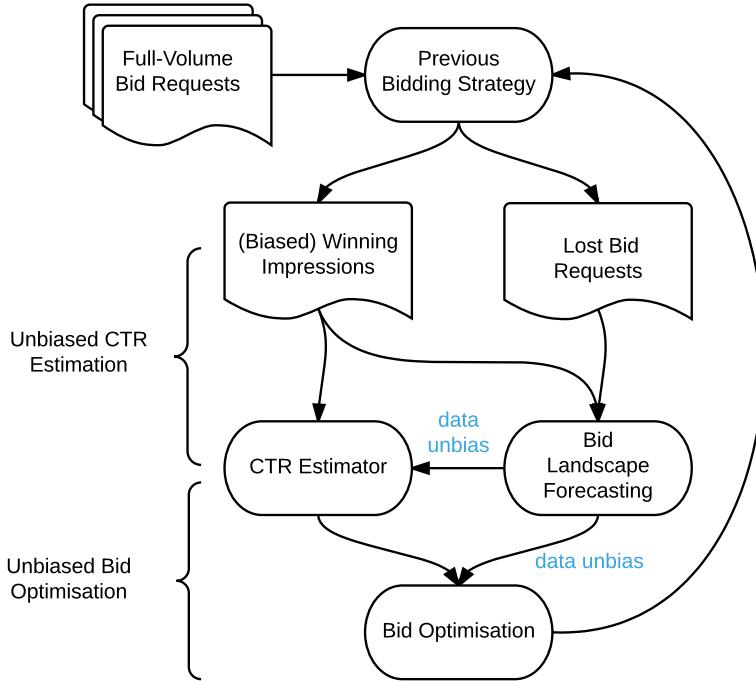
Figure 5.4: Experiment flow chart.

information of the user (e.g. the user interest segments, IP address, browser, operation system, location), advertiser (e.g. the creative format and size), publisher (e.g. the auction reserve price, ad slot size, page domain and URL).

We mainly report the experimental results on the iPinYou dataset for experiment reproducibility while the study on TukMob serves as an auxiliary part particularly for the high-CTR video ad marketplace to make our experiment more comprehensive.

The online A/B testing experiment is conducted based on Yahoo! DSP, a mainstream DSP in the United States ad market. The training dataset comes from its ad log in Aug. and Sep. 2015 while the online A/B testing is performed on 9 campaigns during 7 days of Sep. 2015, which involves 117.1M impressions, 95.4K clicks and 68.6K USD expense.

5.3.2 Experiment Flow

The experiment flow chart is shown in Figure 5.4. The original impression log data is reasonably assumed as full-volume bid request data in our experiment[8]. A truthtelling bidding strategy [4] is performed to simulate the historic bidding process and produce the winning (labelled but biased) impression data and lost (unlabelled) bid re-

---

[8]This assumption is reasonable as this dataset is collected with fixed large bid to reduce the auction-selection bias [115].

quest data. Based on these two datasets, the bid landscape forecasting module as in Eq. (5.5) estimates the market price distribution which acts as the winning function in Eq. (5.1). Thus the observation bias of each data instance from the impression log is estimated. With Eq. (5.8), the unbiased CTR estimation is performed. Furthermore, with the unbiased CTR estimator and the winning function, the unbiased bid optimisation is performed via Eq. (5.11) to get the new bidding function, which is in turn operated in the next prediction stage.

5.3.3   Compared Settings

CTR estimation and bid optimisation are the two tasks we investigate in this work. For each of these tasks, we compare the following four training schemes:

Traditional Biased Training (BIAS). The CTR estimation and bid optimisation are performed based on the impression data without considering any data bias, i.e., all $w(b_x)$ in Eqs. (5.8) and (5.11) are equal to 1. This is the routine training procedure used in most previous work [4, 19, 128].

Training using Observed Market Prices (UOMP). The bias of each training data instance is estimated by the bid landscape forecastor purely based on the observed market prices from impression log, without using the lost bid request data, i.e., all $w(b_x)$ in Eqs. (5.8) and (5.11) are estimated by Eq. (5.3).

Training via Kaplan-Meier Market Prices (KMMP). The bias of each training data instance is estimated by the bid landscape forecastor based on both observed market prices from impression log and the lost bid request data using Kaplan-Meier estimation, i.e., all $w(b_x)$ in Eqs. (5.8) and (5.11) are estimated by Eq. (5.5).

Training with Full-Volume Data (FULL). A progressive bidding strategy is performed to win all the bid requests via bidding extremely high. In such a case the full-volume bid requests are collected with labels to train the CTR estimator and optimise the bidding strategy. In such a setting, the data has no bias and is of full volume, and thus it is regarded as the (unrealistic) upper bound setting of the training.
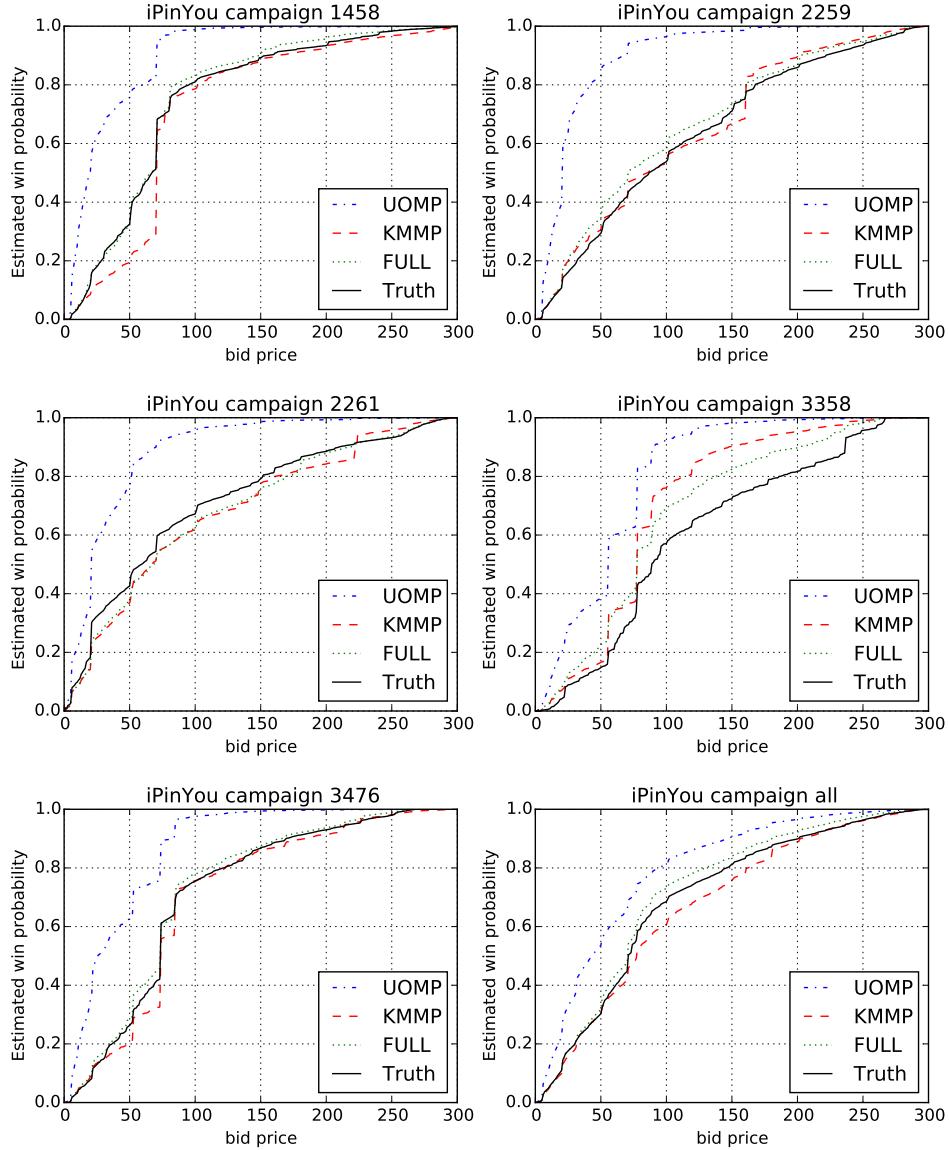
Figure 5.5: Winning probability against bid price (iPinYou).

5.4 Offline Empirical Study

5.4.1 Winning Probability Estimation

Before evaluating the practical CTR estimation and bid optimisation tasks, let us first take an analysis of the compared models' performance on winning probability estimation, i.e., $w(b_{\boldsymbol{x}})$ in Eq. (5.2).

First, Table 5.2 demonstrates the statistics of the full-volume data and the winning impression data by the 'historic' truth-telling bidding strategy as described in Section 5.3.2. As can be observed, for both datasets the winning impression data which is fed into BIAS, UOMP and KMMP training schemes is much smaller than the full-volume

Table 5.2: Winning data statistics: the full-volume data is used in FULL training scheme, while the winning data is used in BIAS, UOMP and KMMP training schemes (both datasets).

iPinYou Camp.	Full Volume	Win Volume	Win Rate
1458	2,055,371	257,077	12.51%
2259	557,038	239,328	42.96%
2261	458,412	213,930	46.67%
2821	881,708	305,134	34.61%
2997	208,292	60,556	29.07%
3358	1,161,403	336,769	29.00%
3386	1,898,535	332,223	17.50%
3427	1,729,177	563,592	32.59%
3476	1,313,574	303,341	23.09%
all	10,263,506	3,973,989	38.72%
TukMob Camp.	Full Volume	Win Volume	Win Rate
all	2,500,000	962,690	38.51%

Table 5.3: Winning probability estimation performance comparison (iPinYou).

Camp.	Pearson Correlation			KL-Divergence		
	UOMP	KMMP	FULL	UOMP	KMMP	FULL
1458	0.9067	0.9903	0.9995	0.4053	0.1204	0.0407
2259	0.7811	0.9959	0.9980	0.7163	0.1870	0.0713
2261	0.9018	0.9947	0.9972	0.3483	0.1057	0.0346
2821	0.8234	0.9947	0.9931	0.5659	0.1421	0.0697
2997	0.8535	0.9285	0.9955	0.3862	0.1761	0.0210
3358	0.9269	0.9772	0.9926	0.5243	0.2652	0.1521
3386	0.9116	0.9821	0.9995	0.3232	0.1391	0.0444
3427	0.9743	0.9977	0.9996	0.1838	0.0762	0.0525
3476	0.9303	0.9979	0.9993	0.3807	0.1147	0.0451
all	0.9795	0.9958	0.9988	0.0893	0.0385	0.0237

data which is fed into FULL training scheme.

Figure 5.5 shows the curves of the winning probability w.r.t. the bid price with three compared settings, i.e., UOMP, KMMP and FULL, on iPinYou dataset. As expected, all the curves start from 0 given the bid 0 and then increase as the bid price increases and finally converge to 1 when the bid price surpasses a threshold (300 for iPinYou dataset). The TRUTH curve is built from all the market price observations from the full-volume prediction data, regarded as the ground truth here. We observe that FULL curve is the closest one to TRUTH curve since FULL makes use of the full-volume training data and is naturally unbiased. The only reason of the slight difference between FULL and TRUTH is the data distribution shift between the training and prediction periods. UOMP always over-estimates the winning probability, as pointed out in Section 5.2.1. Compared to UOMP, KMMP curve is much closer to TRUTH, which

Table 5.4: CTR performance on iPinYou dataset.

Camp.	AUC (%)				Cross Entropy (‰)			
	BIAS	UOMP	KMMP	FULL	BIAS	UOMP	KMMP	FULL
1458	98.26	98.56	99.13	98.57	2.42	2.39	2.39	2.32
2259	60.27	60.94	62.00	67.37	4.04	4.03	4.02	4.00
2261	57.49	58.86	59.05	60.91	3.75	3.74	3.74	3.72
2821	59.25	59.69	60.28	62.36	7.07	7.06	7.04	6.92
2997	59.35	60.50	60.79	59.28	32.89	32.84	32.81	32.38
3358	96.59	96.78	97.01	97.32	4.48	4.47	4.38	4.36
3386	73.74	74.01	74.16	78.23	8.84	8.83	8.83	8.64
3427	96.04	96.42	96.78	97.02	3.37	3.37	3.33	3.31
3476	93.66	93.55	92.19	95.93	4.35	4.34	4.34	4.08
all	71.76	73.84	74.80	78.38	7.71	7.61	7.55	7.31

Table 5.5: CTR performance on TukMob dataset.

Camp.	AUC (%)			
	BIAS	UOMP	KMMP	FULL
all	60.49	60.51	60.67	60.96

shows its advantage of making use of the lost bid request data to improve the winning probability estimation.

Table 5.3 presents the detailed Pearson correlation and KL-divergence between each of the three compared settings and TRUTH on iPinYou dataset. We observe that for all investigated campaigns, KMMP provides a much better estimation, i.e., higher Pearson correlation and lower KL-divergence, than UOMP, and it is even highly comparable with FULL on Pearson correlation. These results demonstrate the surprisingly large improvement that the lost and free bid request data brings to the estimation of winning probability (market price distribution).

5.4.2 CTR Estimation Results

With different biased or unbiased settings, we train the logistic regression model and evaluate its performance. Table 5.4 presents the detailed AUC and cross entropy performance of these 4 compared training schemes for each campaign in iPinYou dataset. Table 5.5 presents the AUC performance comparison on TukMob dataset. We can observe that (i) the proposed unbiased training schemes UOMP and KMMP always outperform the biased but widely adopted BIAS training scheme on all the test campaigns (except for 3476). Specifically, KMMP yields a 3.04% AUC improvement over BIAS on the whole iPinYou dataset, which is a large AUC improvement in ad CTR estima-
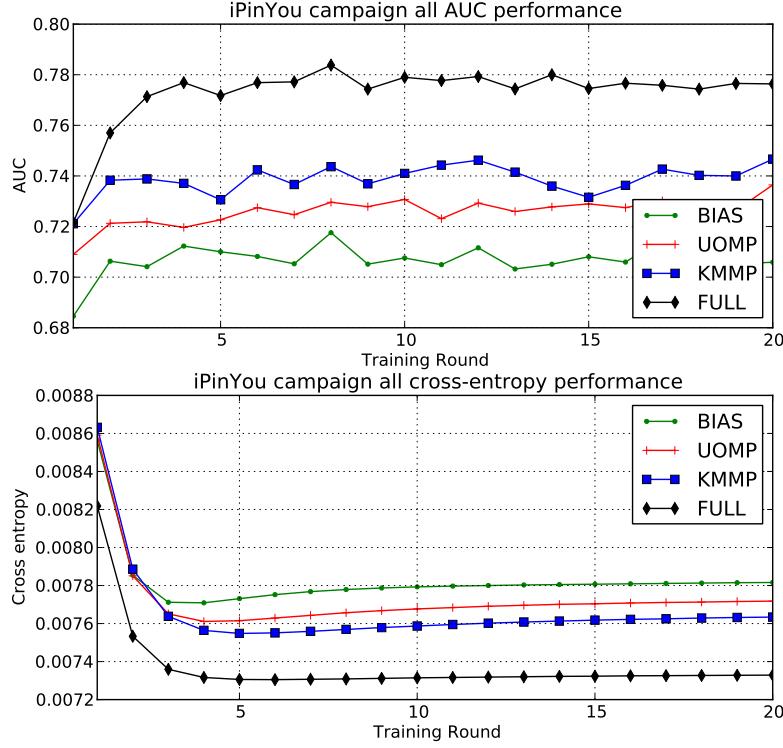
Figure 5.6: CTR performance training convergence (iPinYou).

tion tasks. Such outperformance shows the effectiveness of our models in eliminating the training data instance bias which makes the prediction model generalise better on prediction data. (ii) Comparing the unbiased settings UOMP and KMMP and the upper bound oracle setting FULL, we can see KMMP outperforms UOMP for all the campaigns (except for 3476). For some campaigns, e.g., 1458 and 2997, KMMP even slightly outperforms FULL[9] which again shows the advantages of making use of the lost auction information for better estimating the instance bias.

Figure 5.6 shows the AUC and the cross entropy on prediction data of all iPinYou campaigns for each training round. We can observe the unbiased UOMP and KMMP models learn stably and consistently outperform BIAS. FULL substantially outperforms other compared training schemes, which is not surprising as FULL obtains much more training data instances (as shown in Table 5.2) and the data distribution is unbiased.

5.4.3 Bid Optimisation Results

For bid optimisation experiment, we mainly focus on the click performance improvement from bidding strategy parameter optimisation via Eqs. (5.16) and (5.19) instead of the difference of CTR estimation. Thus in our training/prediction environment, the

---

[9]This is mainly caused by the local data distribution, which is not significant.

Table 5.6: Bid optimisation click performance (iPinYou).

Camp.	1/64 budget setting				1/4 budget setting			
	BIAS	UOMP	KMMP	FULL	BIAS	UOMP	KMMP	FULL
1458	363	400	460	468	469	470	471	482
2259	5	5	5	7	49	51	51	50
2261	5	5	7	7	35	36	35	45
2821	18	18	23	27	65	91	106	134
2997	37	39	42	44	156	188	222	226
3358	86	117	137	140	183	198	220	221
3386	9	9	22	38	69	78	144	165
3427	103	119	154	169	242	262	286	314
3476	6	8	11	13	59	99	108	106
all	268	372	462	521	1,584	1,740	1,871	2,087

Table 5.7: Bid optimisation click performance (TukMob).

Budget Setting	Click Number				eCPC			
	BIAS	UOMP	KMMP	FULL	BIAS	UOMP	KMMP	FULL
1/32	846	848	866	871	9.25	9.23	9.04	8.99
1/16	1,829	1,831	1,863	1,838	8.56	8.55	8.40	8.52
1/8	3,721	3,721	3,774	3,775	8.42	8.42	8.30	8.29
1/4	7,181	7,178	7,226	7,257	8.72	8.72	8.67	8.63
1/2	13,127	13,132	13,163	13,019	9.54	9.54	9.52	9.62

logistic regression CTR estimator is trained based on a separate unbiased training data and is shared in all 4 compared training schemes of bid optimisation. For each training scheme, we train the optimal parameter λ in Eq. (5.19) via the biased or unbiased training data, then apply the corresponding bidding strategy Eq. (5.16) on prediction data to observe its performance.

We follow Chapters 3 and 4 to set the budget proportions to perform offline bid optimisation, where the train/test budget is set as 1/64, 1/32, 1/16, 1/8, 1/4 and 1/2 of the total expense of the train/test dataset. We cannot set the proportion as 1 because in such a case one may simply bid infinity to win all the impressions and clicks in the data and just spend all the budget.

Table 5.6 shows the click performance of the 4 compared training schemes with 1/64 and 1/4 budget settings respectively for each iPinYou campaign. Table 5.7 shows the overall click and eCPC performance comparison against different budget settings on TukMob dataset. We can observe that the unbiased UOMP and KMMP consistently outperform the traditional BIAS which were used in the most of the previous bid opti-
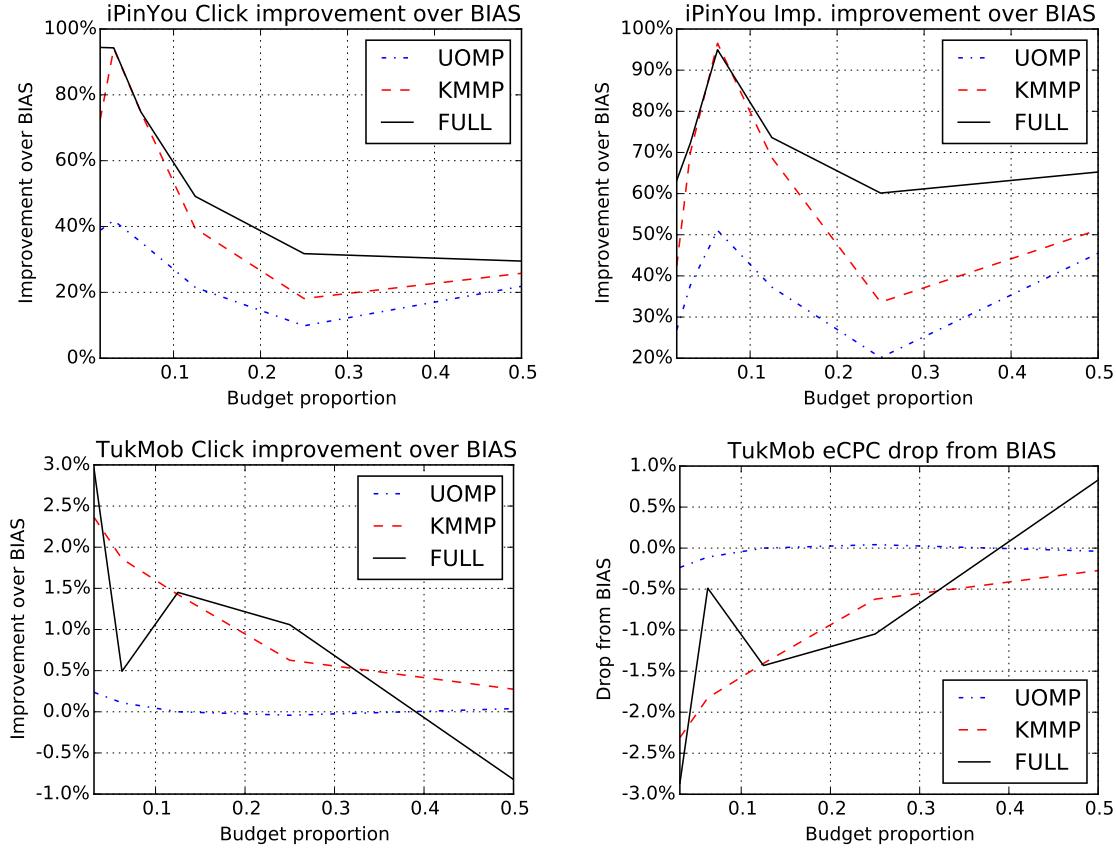
Figure 5.7: Improvement over BIAS w.r.t. budget proportions.

misation work [36, 19, 128]. This shows the great potential of our proposed unbiased training schemes in bid optimisation. Furthermore, KMMP outperforms UOMP and it is very close to the theoretic upper bound from FULL, in 17 out of 20 test cases, suggesting it is generally much better to leverage the winning probability obtained from the censored observations of both winning impressions and lost bid requests.

Figure 5.7 further provides the click, impression improvement percentages and eCPC drop percentage of the unbiased training schemes against BIAS with different budget settings. The improvements for clicks and impressions are positive for all budget settings and the eCPC drops are negative for all budget settings (except FULL on 1/2), which show the robustness of the unbiased training schemes. Also we can observe that KMMP dominates UOMP and heavily approaches the upper bound FULL.

5.5   Online A/B Test

We have deployed the unbiased KMMP training scheme on Yahoo! DSP in the United States ad market and performed the online A/B testing for 9 campaigns during 7 days

Table 5.8: Online A/B testing of CTR estimation (Yahoo!).

Camp.	BIAS AUC	KMMP AUC	AUC Lift
C1	63.78%	64.12%	0.34%
C2	87.45%	88.58%	1.13%
C3	69.73%	75.52%	5.79%
C4	88.82%	89.55%	0.73%
C5	69.71%	72.29%	2.58%
C6	89.33%	90.70%	1.37%
C7	77.76%	78.92%	1.16%
C8	74.57%	76.98%	2.41%
C9	71.04%	73.12%	2.08%
all	73.48%	76.45%	2.97%

in Sep. 2015. For each campaign, we create two experiment buckets: control and treatment. Each is allocated 50% of the bid request traffic (based on user ID to avoid attribution conflicts), and 50% of the campaign's budget. The control bucket uses gradient boosting decision tree click predictor [140] trained with BIAS, while the click model used in the treatment bucket is trained with KMMP. The deployed bidding strategy is the conventional truth-telling bidding [4].

In order to perform an unbiased evaluation of the CTR estimation, we deployed a bidding agent performing very high constant bidding in Sep. 2015 to collect an ad impression dataset which can be regarded as full-volume unbiased test data. The training data is still the traditional biased ad impression dataset collected during Aug. and early Sep. 2015. Table 5.8 provides the detailed CTR estimation performance for each campaign and the overall performance. As can be observed, KMMP provides a consistent AUC improvement over BIAS across all investigated campaigns. The overall AUC is 73.48% for BIAS and 76.45% for KMMP, i.e. a 2.97% AUC lift, which is a very large improvement for CTR estimation task in practice.

Table 5.9 further presents the detailed performance of A/B testing of bid optimisation on the 9 campaigns. Figure 5.8 depicts the relative difference comparing the performance of KMMP against BIAS. We found that with the same campaign budget the KMMP-trained model acquires more clicks (most of the time) but fewer impressions than the BIAS-trained one, which makes its CTR much higher than BIAS. The reason is that there is less over-prediction on many cheap cases. In the biased training data, the over-predicted CTR and cheap cases are more likely to be sampled because the historic

Table 5.9: Online A/B testing of bid optimisation (Yahoo!).

Camp.	Impressions (M)		Clicks (K)		CTR (%)		eCPC ($/click)	
	BIAS	KMMP	BIAS	KMMP	BIAS	KMMP	BIAS	KMMP
C1	1.07	0.89	0.62	0.67	0.06	0.08	4.54	4.16
C2	7.73	6.02	0.94	1.19	0.01	0.02	7.49	5.89
C3	22.18	16.96	27.18	30.06	0.12	0.18	0.26	0.23
C4	0.37	0.12	0.61	0.61	0.16	0.49	2.46	2.48
C5	9.57	7.51	6.42	6.93	0.07	0.09	0.45	0.42
C6	0.32	0.22	0.46	0.46	0.14	0.21	2.17	2.18
C7	10.13	7.31	2.99	3.28	0.03	0.04	0.37	0.34
C8	1.04	0.52	1.04	1.13	0.10	0.22	1.92	1.78
C9	13.67	11.46	5.12	5.71	0.04	0.05	1.76	1.58
all	66.07	51.01	45.37	50.03	0.07	0.10	0.76	0.69

bidding strategy over bade on these cheap cases, vice versa on less-prediction over expensive cases. With the KMMP training scheme, the bidding strategy to-some-extent gets rid of such a bias to avoid over-prediction on cheap cases, which provides fewer impressions but more clicks.

Overall, with the same budget, the bidding strategy trained with KMMP achieves much better eCPC (9.30% drop) and CTR (42.8% rise) than the conventional one trained with BIAS. The KMMP-trained click model effectively alleviates over-prediction especially in the low-CTR region and thus becomes more efficient in acquiring clicks. Therefore, with the bidding strategy with unbiased KMMP-trained click model, campaigns could acquire clicks in a more cost-effective way.

5.6 Summary

In this chapter, we have studied the data observation bias problem in display advertising generated from the auction selection that would hurt the performance of various supervised learning models. To address this problem, we proposed a model-free learning framework that eliminates the model bias generated from censored auction data. The derived Bid-aware Gradient Descent (BGD) learning scheme naturally incorporates the historic auction and bid information. We found that the historic bid for each instance could influence both BGD learning weight and update direction. Comprehensive empirical study based on iPinYou and TukMob datasets demonstrated the large improvement of our learning framework over strong baselines in both CTR estimation and bid optimisation tasks. With light engineering work, the learning framework was
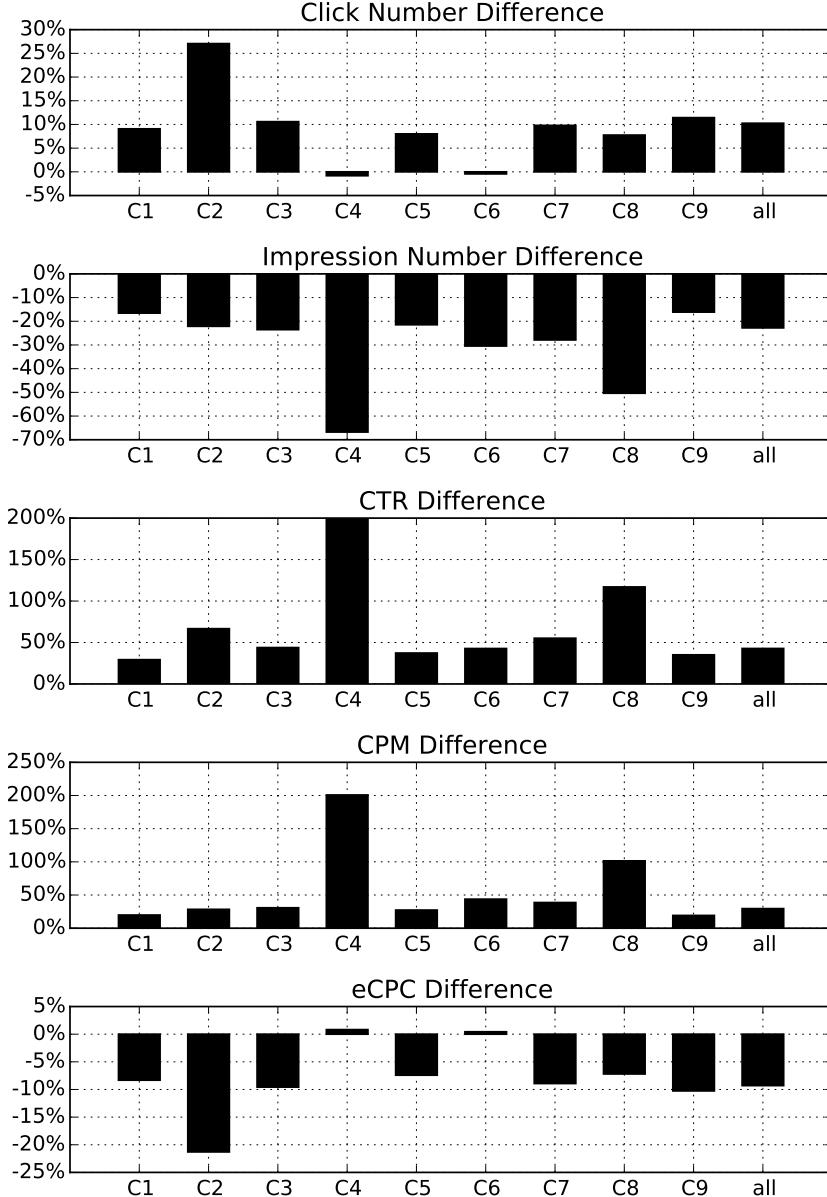
Figure 5.8: Relative performance difference between KMMP and BIAS in Yahoo! online A/B testing: (KMMP-BIAS)/BIAS.

deployed on Yahoo! DSP and brought 2.97% AUC lift in CTR estimation and 9.30% eCPC drop in bid optimisation over 9 campaigns.

It is important to point out that such a learning framework is flexible with other supervised learning tasks than the investigated ones in this work, such as budget pacing and frequency capping in online advertising as well as other data science problems, such as interactive recommender systems [141], off-policy reinforcement learning [142], which are our planned future work.

Chapter 6

Feedback Control Mechanism

6.1 Background and Motivations

As has been pointed out in Figure 1.9 and in Chapters 3 and 4, despite the ability of delivering performance-driven advertising, RTB, unfortunately, results in high volatilities, measured by major Key Performance Indicators (KPIs), such as CPM (cost per mille), AWR (auction winning ratio), eCPC (effective cost per click) and CTR (click-through rate). Such instability causes advertisers ample difficulty in optimising and controlling the KPIs against their cost.

In this chapter, we propose to employ feedback control theory [96] to solve the instability problem in RTB. Feedback controllers are widely used in various applications for maintaining dynamically changing variables at the predefined reference values. The application scenarios range from the plane direction control [143] to the robot artificial intelligence [144]. In our RTB scenario, the specific KPI value, depending on the requirements from the advertisers, is regarded as the variable we want to control with a pre-specified reference value. Our study focuses on two use cases. (i) For performance-driven advertising, we are concerned with the feedback control of the average cost on acquiring a click, measured by effective cost per click (eCPC). (ii) For branding based advertising, to ensure a certain high exposure of a campaign, we focus on the control of the ratio of winning the auctions for the targeted impressions, measured by auction winning ratio (AWR). More specifically, we take each of them as the control input signal and consider the gain (the adjustment value) of bid price as the control output signal for each incoming ad display opportunity (the bid request). We develop two controllers to test: the widely used proportional-integral-derivative (PID) controller [145] and the

waterlevel-based (WL) controller [146]. We conduct large-scale experiments to test the feedback control performance with different settings of reference value and reference dynamics. Through the empirical study, we find that PID and WL controllers are capable of controlling eCPC and AWR, while PID further provides a better control accuracy and robustness than WL.

Furthermore, we investigate whether the proposed feedback control can be employed for controllable bid optimisation. It is common that the performance of an ad campaign (e.g., eCPC) varies from different channels (e.g., ad exchanges, user geographic regions and PC/mobile devices) [29]. If one can reallocate some budget from less cost-effective channels to more cost-effective ones, the campaign-level performance would improve [52]. In this chapter, we formulate the multi-channel bid optimisation problem and propose a model to calculate the optimal reference eCPC for each channel. Our experiments show that the campaign-level click number and eCPC achieve significant improvements with the same budget.

Moreover, the proposed feedback control mechanism has been implemented and integrated in a commercial DSP. The conducted live test shows that in a real and noisy setting the proposed feedback mechanism has the ability to produce controllable advertising performance.

To sum up, the contributions of this chapter are as follows. (i) We study the instability problem in RTB and investigate its solution by leveraging the feedback control mechanism. (ii) Extensive offline and online experiments show that PID controller is better than other alternatives and finds the optimal way to settle the variable in almost all studied cases. (iii) We further discover that feedback controllers are of great potential to perform bid optimisation through settling the eCPC at the reference value calculated by our proposed mathematical click maximisation framework.

6.2   RTB Feedback Control System

Figure 6.1 presents the diagram of the proposed RTB feedback control system. The traditional bidding strategy is represented as the *bid calculator* module in the DSP bidding agent. As studied in Chapter 3, the bid decision depends on two factors for each ad impression: the utility (e.g., CTR, expected revenue) and cost (i.e., expected charged price) [128]. In a widely adopted bidding strategy [19], the utility is evaluated
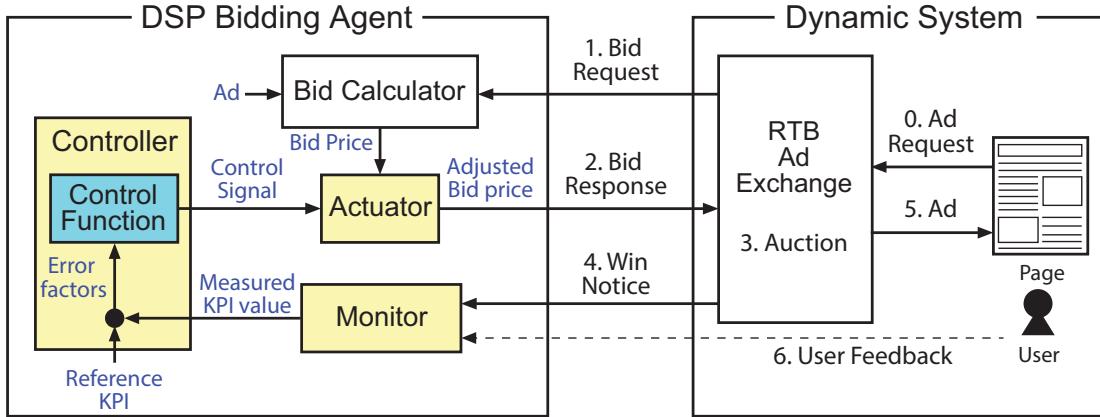
Figure 6.1: Feedback controller integrated in the RTB system.

by CTR estimation while the base bid price is tuned based on the bid landscape [24] for the cost evaluation. The generalised bidding strategy in [19] is

$$b(t) = b_0 \frac{\theta_t}{\theta_0}, \tag{6.1}$$

where θ_t is the estimated CTR for the bid request at moment t; θ_0 is the average CTR under a target condition (e.g., a user interest segment); and b_0 is the tuned base bid price for the target condition. In this work, to make the work with more generality, we adopt this widely used bidding strategy with a logistic CTR estimator [1].

The controller plays as a role which adjusts the bid price from the bid calculator. Specifically, the *monitor* receives the auction win notice from the ad exchange and the user click feedback from the ad tracking system, which as a whole we regard as the dynamic system. Then the current KPI values, such as AWR and eCPC can be calculated. If the task is to control the eCPC with the reference value, the error factor between the reference eCPC and the measured eCPC is calculated then sent into the control function. The output control signal is sent to the *actuator*, which uses the control signal to adjust the original bid price from the bid calculator. The adjusted bid price is packaged with the qualified ad into the bid response and sent back to the ad exchange for auction.

6.2.1 Actuator

For the bid request at the moment t, the actuator takes into the current control signal $\phi(t)$ to adjust the bid price from $b(t)$ (Eq. (6.1)) to a new value $b_a(t)$. In our model, the

control signal, which will be mathematically defined in the next subsections, is a gain on the bid price. Generally, when the control signal $\phi(t)$ is zero, there should be no bid adjustment. There could be different actuator models, and in our work we choose to use

$$b_a(t) = b(t)\exp\{\phi(t)\}, \tag{6.2}$$

where the model satisfies $b_a(t) = b(t)$ when $\phi(t) = 0$. Other models such as the linear model $b_a(t) \equiv b(t)(1 + \phi(t))$ are also investigated in our study but it performs poorly in the situations when a big negative control signal is sent to the actuator, where the linear actuator will usually respond a negative or a zero bid, which is meaningless in our scenario. By contrast, the exponential model is a suitable solution to addressing the above drawback because it naturally avoids generating a negative bid. In the later empirical study we mainly report the analysis based on the exponential-form actuator model.

6.2.2 PID Controller

The first controller we investigate is the classic PID controller [145]. As its name implies, a PID controller produces the control signal from a linear combination of the proportional factor, the integral factor and the derivative factor based on the error factor:

$$e(t_k) = x_r - x(t_k), \tag{6.3}$$

$$\phi(t_{k+1}) \leftarrow \lambda_P \underbrace{e(t_k)}_{\text{proportional}} + \lambda_I \underbrace{\sum_{j=1}^{k} e(t_j)\triangle t_j}_{\text{integral}} + \lambda_D \underbrace{\frac{\triangle e(t_k)}{\triangle t_k}}_{\text{derivative}}, \tag{6.4}$$

where the error factor $e(t_k)$ is the reference value x_r minus the current controlled variable value $x(t_k)$, the update time interval is given as $\triangle t_j = t_j - t_{j-1}$, the change of error factors is $\triangle e(t_k) = e(t_k) - e(t_{k-1})$, and λ_P, λ_I, λ_D are the weight parameters for each control factor. Note that here the control factors are all in discrete time $(t_1, t_2, \ldots)$ because bidding events are discrete and it is practical to periodically update the control factors. All control factors $(\phi(t), e(t_k), \lambda_P, \lambda_I, \lambda_D)$ remain the same between two updates. Thus for all time t between t_k and t_{k+1}, the control signal $\phi(t)$ in Eq. (6.2) equals $\phi(t_k)$. We see that P factor tends to push the current variable value to the refer-
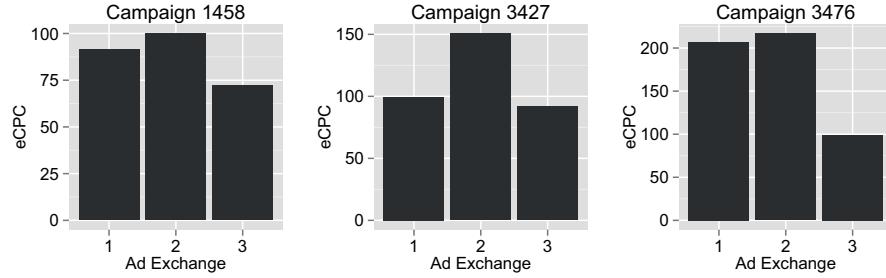
Figure 6.2: Different eCPCs across different ad exchanges. Dataset: iPinYou.

ence value; *I* factor reduces the accumulative error from the beginning to the current time; *D* factor controls the fluctuation of the variable.

6.2.3 Waterlevel-based Controller

The Waterlevel-based (WL) controller is another feedback control model which was originally used for switching devices controlled by water level [146]:

$$\phi(t_{k+1}) \leftarrow \phi(t_k) + \gamma(x_r - x(t_k)), \tag{6.5}$$

where γ is the step size parameter for $\phi(t_k)$ update in exponential scale.

Compared to PID, the WL controller only takes the difference between the variable value and the reference value into consideration. Moreover, it provides a sequential control signal. That is, the next control signal is an adjustment based on the previous one.

6.2.4 Setting References for Click Maximisation

In this subsection, we investigate how the proposed feedback control mechanism can be used for controllable bid optimisation. In addition to adopting specific models of bidding strategies [19, 128] or budget allocation [36], we demonstrate that the feedback control mechanism can be leveraged as a model-free click maximisation framework embedded with any bidding strategies and performs automatic budget allocation across different *channels*[1] via setting smart reference values.

When an advertiser specifies the targeted audience (usually also combined with ad impression contextual categories) for their specific campaign, the impressions that fit the target rules may come from separate channels such as different ad exchanges,

---

[1]A game theoretic analysis of bidders' selection among channels (exchanges) are discussed in [147], but this is out of the scope of this thesis.

user regions, users' PC/mobile devices etc. It is common that the DSP integrates with several ad exchanges and delivers the required ad impressions from all those ad exchanges (as long as the impressions fit the target rule), although the market prices [28] may be significantly different. Figure 6.2 illustrates that, for the same campaign, there is a difference in terms of eCPC across different ad exchanges. As pointed out in [29], the differences are also found in other channels such as user regions and devices.

The cost differences provide advertisers a further opportunity to optimise their campaign performance based on eCPCs. To see this, suppose a DSP is integrated to two ad exchanges A and B. For a campaign in this DSP, if its eCPC from exchange A is higher than that from exchange B, which means the inventories from exchange B are more cost effective than those from exchange A, then by reallocating some budget from exchange A to B will *potentially* reduce the overall eCPC of this campaign. Practically the budget reallocation can be done by reducing the bids for exchange A while increasing the bids for exchange B. Here we formally propose a model of calculating the equilibrium eCPC of each ad exchange, which will be used as the optimal reference eCPC for the feedback control that leads to a maximum number of clicks given the budget constraint.

Mathematically, suppose for a given ad campaign, there are n ad exchanges (could be other channels), i.e., $1, 2, \ldots, n$, that have the ad volume for a target rule. In our formulation we focus on optimising clicks, while the formulation of conversions can be obtained similarly. Let ξ_i be the eCPC on ad exchange i, and $c_i(\xi_i)$ be the click number that the campaign acquires in the campaign's lifetime if we tune the bid price to make its eCPC be ξ_i for ad exchange i. For advertisers, they want to maximise the campaign-level click number given the campaign budget B [128]:

$$\max_{\xi_1,\ldots,\xi_n} \sum_i c_i(\xi_i) \tag{6.6}$$

$$s.t. \ \sum_i c_i(\xi_i)\xi_i = B. \tag{6.7}$$

Its Lagrangian is

$$\mathscr{L}(\xi_1,\ldots,\xi_n,\alpha) = \sum_i c_i(\xi_i) - \alpha(\sum_i c_i(\xi_i)\xi_i - B), \tag{6.8}$$
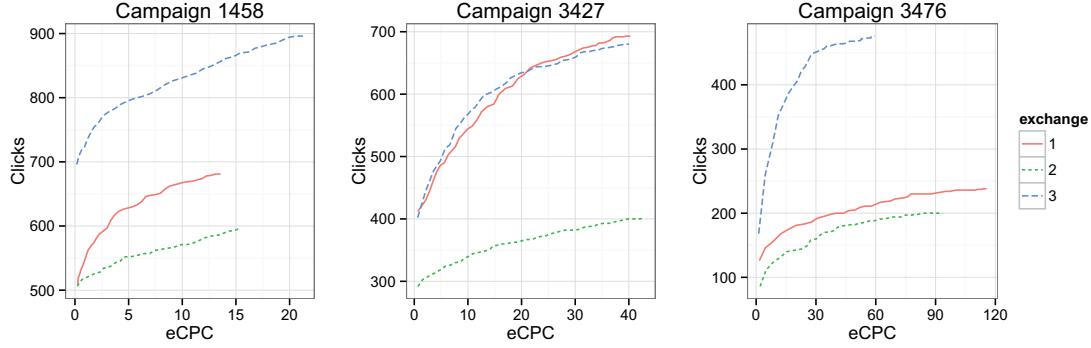
Figure 6.3: #Clicks against eCPC on different ad exchanges.

where α is the Lagrangian multiplier. Then we take its gradient on ξ_i and let it be 0:

$$\frac{\partial \mathcal{L}(\xi_1, \ldots, \xi_n, \alpha)}{\partial \xi_i} = c_i'(\xi_i) - \alpha(c_i'(\xi_i)\xi_i + c_i(\xi_i)) = 0, \tag{6.9}$$

$$\frac{1}{\alpha} = \frac{c_i'(\xi_i)\xi_i + c_i(\xi_i)}{c_i'(\xi_i)} = \xi_i + \frac{c_i(\xi_i)}{c_i'(\xi_i)}, \tag{6.10}$$

where the equation holds for each ad exchange i. As such, we can use α to bridge the equations for any two ad exchanges i and j:

$$\frac{1}{\alpha} = \xi_i + \frac{c_i(\xi_i)}{c_i'(\xi_i)} = \xi_j + \frac{c_j(\xi_j)}{c_j'(\xi_j)}. \tag{6.11}$$

So the optimal solution condition is given as follows:

$$\frac{1}{\alpha} = \xi_1 + \frac{c_1(\xi_1)}{c_1'(\xi_1)} = \xi_2 + \frac{c_2(\xi_2)}{c_2'(\xi_2)} = \cdots = \xi_n + \frac{c_n(\xi_n)}{c_n'(\xi_n)}, \tag{6.12}$$

$$\sum_i c_i(\xi_i)\xi_i = B. \tag{6.13}$$

With sufficient data instances, we find that $c_i(\xi_i)$ is always a concave and smooth function. Some examples are given in Figure 6.3. Based on the observation, it is reasonable to define a general polynomial form of the $c_i(\xi_i)$ functions:

$$c_i(\xi_i) = c_i^* a_i \left(\frac{\xi_i}{\xi_i^*} \right)^{b_i}, \tag{6.14}$$

where ξ_i^* is the campaign's historic average eCPC on the ad inventories from ad exchange i during the training data period, and c_i^* is the corresponding click number.

These two factors are directly obtained from the training data. Parameters a_i and b_i are to be tuned to fit the training data (e.g., in Figure 6.3).

Substituting Eq. (6.14) into Eq. (6.12) gives

$$\frac{1}{\alpha} = \xi_i + \frac{c_i(\xi_i)}{c_i'(\xi_i)} = \xi_i + \frac{\frac{c_i^* a_i}{\xi_i^{*b_i}} \xi_i^{b_i}}{\frac{c_i^* a_i}{\xi_i^{*b_i}} b_i \xi_i^{b_i-1}} = \left(1 + \frac{1}{b_i}\right)\xi_i. \tag{6.15}$$

We can then rewrite Eq. (6.12) as

$$\frac{1}{\alpha} = \left(1 + \frac{1}{b_1}\right)\xi_1 = \left(1 + \frac{1}{b_2}\right)\xi_2 = \cdots = \left(1 + \frac{1}{b_n}\right)\xi_n. \tag{6.16}$$

$$\text{Thus } \xi_i = \frac{b_i}{\alpha(b_i+1)}. \tag{6.17}$$

Interestingly, from Eq. (6.17) we find that the equilibrium is *not* in the state that the eCPCs from the exchanges are the same. Instead, it is when any amount of budget reallocated among the exchanges does not make more total clicks; for instance, in a two-exchange case, the equilibrium is reached when the increase of the clicks from one exchange equals the decrease from the other (Eq. (6.9)). More specifically, from Eq. (6.17) we observe that for ad exchange i, if its click function $c_i(\xi_i)$ is quite flat, i.e., the click number increases more slowly as its eCPC increases in a certain area, then its learned b_i should be small. This means the factor $\frac{b_i}{b_i+1}$ is small as well; then from Eq. (6.17) we can see the optimal eCPC in ad exchange i should be relatively small.

Substituting Eqs. (6.14) and (6.17) into Eq. (6.7) gives

$$\sum_i \frac{c_i^* a_i}{\xi_i^{*b_i}} \left(\frac{b_i}{b_i+1}\right)^{b_i+1} \left(\frac{1}{\alpha}\right)^{b_i+1} = B, \tag{6.18}$$

where for simplicity, we denote for each ad exchange i, its parameter $\frac{c_i^* a_i}{\xi_i^{*b_i}} \left(\frac{b_i}{b_i+1}\right)^{b_i+1}$ as δ_i. This give us a simpler form as:

$$\sum_i \delta_i \left(\frac{1}{\alpha}\right)^{b_i+1} = B. \tag{6.19}$$

In most settings of b_i values, there is no closed form to solve Eq. (6.19) for α. However, as b_i is non-negative and $\sum_i \delta_i (\frac{1}{\alpha})^{b_i+1}$ monotonically increases against $\frac{1}{\alpha}$, one

can easily obtain the solution for α by using a numeric solution such as the stochastic gradient decent or the Newton method [148]. Finally, based on the solved α, we can find the optimal eCPC ξ_i for each ad exchange i using Eq. (6.17). In fact, these eCPCs are the reference value we want the campaign to achieve for the corresponding ad exchanges. By setting x_r in Eq. (6.3) as ξ_i for each ad exchange i, we can use PID controllers to achieve these reference eCPCs so as to achieve the maximum number of clicks on the campaign level.

As a special case, if we regard the whole volume of the campaign as one channel, this method can be directly used as a general bid optimisation tool. It makes use of the campaign's historic data to decide the optimal eCPC and then the click optimisation is performed by controlling the eCPC to settle at the optimal eCPC as reference. Note that this multi-channel click maximisation framework is flexible to incorporate any bidding strategies.

6.3 Experimental Setup

We conduct comprehensive experiments to study the proposed RTB feedback control mechanism. Our focus in this section is on offline evaluation using a publicly-available real-world dataset. To make our experiment repeatable, we have published the experiment code[2]. The online deployment and test on a commercial DSP will be reported in Section 6.5.

6.3.1 Dataset

In consistency with previous chapters, we test our system on the publicly available dataset collected from iPinYou DSP [115], where 9 campaigns from this datasets are investigated. More statistics and analysis of the dataset is available in Section 3.4.1. Note that for each campaign, there are normally 10-day ad log from the dataset, and according to the data publisher [115], the last three-day data of each campaign is split as the test data and the rest as the training data.

6.3.2 Evaluation Protocol

We follow the evaluation protocol from previous chapters on bid optimisation and an RTB contest [115] to run our experiment. Specifically, for each data record, we pass

---

[2]`https://github.com/wnzhang/rtbcontrol`

the feature information to our bidding agent. In our bidding agent, the bid calculator generates a new bid based on the CTR prediction and other parameters in Eq. (6.1), and then the actuator adjusts the bid based on the control signal as in Eq. (6.2). We then compare the adjusted bid with the logged actual auction winning price. If the bid is higher than the auction winning price, we know the bidding agent has won this auction, paid the winning price, and obtained the ad impression. If from the ad impression record there is a click, then the placement has generated a positive outcome (one click) with a cost equal to the winning price. If there is no click, the placement has resulted in a negative outcome and wasted the money. The control parameters are updated every 2 hours (as one round).

Again, it is worth mentioning that historical user feedback has been widely used for evaluating information retrieval systems [149] and recommender systems [150]. All of them used historic clicks as a proxy for relevancy to train the prediction model as well as to form the ground truth. Similarly, our evaluation protocol keeps the user contexts, displayed ads (creatives etc.), bid requests, and auction environment unchanged. We intend to answer that under the same context if the advertiser were given a different or better bidding strategy or employed a feedback loop, whether they would be able to get more clicks with the budget limitation. The click would stay the same as nothing has been changed for the users. This methodology works well for evaluating bid optimisation [28] (Chapters 3, 4 and 5) and has been adopted in the display advertising industry [115].

6.3.3   Evaluation Measures

We adopt several commonly used measures in feedback control systems [95]. We define the *error band* as the $\pm 10\%$ interval around the reference value. If the controlled variable settles within this area, we consider that the variable is successfully controlled. The **speed** of convergence (to the reference value) is also important. Specifically, we evaluate the *rise time* to check how fast the controlled variable will get into the error band. We also use the *settling time* to evaluate how fast the controlled variable will be successfully restricted into the error band. However, the fast convergence may cause the problem of inaccurate control. Thus, two control **accuracy** measures are introduced. We use the *overshoot* to measure the percentage of value that the controlled variable

passes over the reference value. After the settling (called the steady state), we use the *RMSE-SS* to evaluate the root mean square error between the controlled variable value and the reference value. At last, we measure the control **stability** by calculating the standard deviation of the variable value after settling, named as *SD-SS*.

For bid optimisation performance, we use the campaign's total achieved click number and eCPC as the prime evaluation measures. We also monitor the impression related performance such as impression number, AWR and CPM.

6.4 Offline Empirical Study

Our empirical study consists of five parts with the focus on controlling two KPIs: eCPC and AWR. (i) In Section 6.4.1, we answer whether the proposed feedback control systems are practically capable of controlling the KPIs. (ii) In Section 6.4.2, we study the control difficulty with different reference value settings. (iii) In Section 6.4.3, we focus on the PID controller and investigate its attributes on settling the target variable. (iv) In Section 6.4.4, we leverage the PID controllers as a bid optimisation tool and study their performance on optimising the campaign's clicks and eCPC across multiple ad exchanges. (v) Finally, more discussions about PID parameter tuning and online updating will be given in Section 6.4.5.

6.4.1 Control Capability

For each campaign, we check the performance of the two controllers on two KPIs. We first tune the control parameters on the training data to minimise the settling time. Then we adopt the controllers over the test data and observe the performance. The detailed control performance on each campaign is provided in Table 6.1 for eCPC[3] and Table 6.2 for AWR. Figure 6.4 shows the controlled KPI curves against the timesteps (i.e., round). The dashed horizontal line represents the reference.

We see from the results that (i) all the PID controllers can settle both KPIs within the error band (with the settling time less than 40 rounds), which indicates that the PID control is capable of settling both KPIs at the given reference value. (ii) The WL controller on eCPC does not work that well on test data, even though we could find good parameters on training data. This is due to the fact that WL controller tries to

---

[3] "-" cells mean invalid because of the failure to rise or settle.

Table 6.1: Overall control performance on eCPC.

Campaign	Controller	Rise	Settling	Overshoot	RMSE-SS	SD-SS
1458	PID	1	5	7.73	0.0325	0.0313
	WL	6	36	0	0.0845	0.0103
2259	PID	7	7	8.03	0.0449	0.0411
	WL	6	-	0	-	-
2261	PID	3	23	17.66	0.0299	0.0294
	WL	5	-	0	-	-
2821	PID	17	22	14.47	0.0242	0.0216
	WL	-	-	0	-	-
2997	PID	17	17	0.75	0.0361	0.026
	WL	-	-	0	-	-
3358	PID	3	7	23.89	0.0337	0.0287
	WL	-	-	0	-	-
3386	PID	9	13	7.90	0.0341	0.0341
	WL	-	-	0	-	-
3427	PID	1	12	29.03	0.0396	0.0332
	WL	-	-	0	-	-
3476	PID	1	5	7.64	0.0327	0.031
	WL	1	-	17.11	-	-

Table 6.2: Overall control performance on AWR.

Campaign	Controller	Rise	Settling	Overshoot	RMSE-SS	SD-SS
1458	PID	4	10	16.86	0.0153	0.0093
	WL	3	7	0.00	0.0448	0.0231
2259	PID	4	6	17.08	0.0076	0.0072
	WL	1	13	3.91	0.0833	0.0113
2261	PID	1	4	16.39	0.0205	0.0203
	WL	1	-	2.02	-	-
2821	PID	6	8	16.44	0.0086	0.0086
	WL	1	3	5.77	0.0501	0.0332
2997	PID	1	8	13.68	0.0151	0.0151
	WL	1	-	0.00	-	-
3358	PID	2	8	22.08	0.0250	0.0213
	WL	1	7	0.13	0.0332	0.0211
3386	PID	4	8	18.85	0.0133	0.0118
	WL	1	5	2.95	0.0300	0.0291
3427	PID	2	6	26.63	0.0200	0.0179
	WL	3	13	0.24	0.0482	0.0257
3476	PID	2	6	27.15	0.0175	0.0161
	WL	1	7	1.49	0.0308	0.0271

affect the average system behaviour through transient performance feedbacks while facing the huge dynamics of RTB. (iii) We can also observe that for WL on AWR, most campaigns are controllable while there are still two campaigns that fail to settle at the reference value. (iv) Compared to PID on AWR, WL always results in higher RMSE-
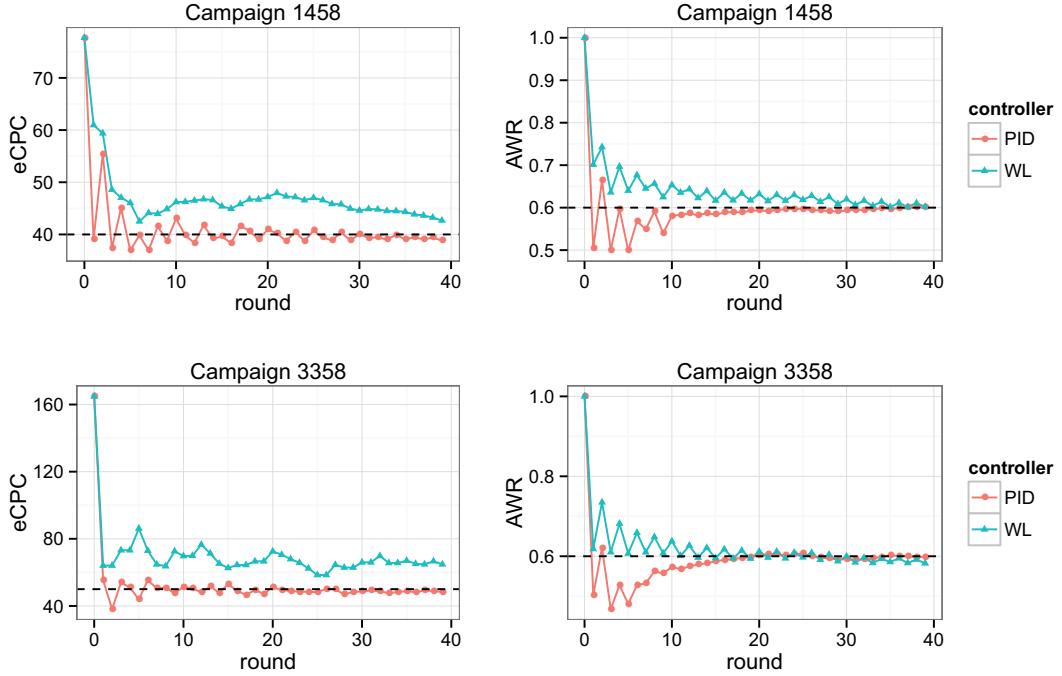
Figure 6.4: Control performance on AWR and eCPC.

SS and SD-SS values but lower percentage overshoot. Those control settings with a fairly short rise time usually face a higher overshoot.

According to above results, PID controller outperforms the WL controller in the tested RTB cases. We believe this is due to the fact that the integral factor in PID controller helps reduce the accumulative error (i.e., RMSE-SS) and the derivative factor helps reduce the variable fluctuation (i.e., SD-SS). And it is easier to settle the AWR than the eCPC. This is mainly because AWR only depends on the market price distribution while eCPC additionally involves the user feedback, i.e., CTR, where the prediction is associated with significant uncertainty.

6.4.2 Control Difficulty

In this section, we extend our control capability experiments further by adding higher and lower reference values in comparison. Our goal is to investigate the impact of different levels of reference values on control difficulty. We follow the same scheme to train and test the controllers as Section 6.4.1. However, instead of showing the exact performance value, our focus here is on the performance comparison with different reference settings.

The distribution of achieved settling time, RMSE-SS and SD-SS, with the setting of three reference levels, i.e., low, middle and high, are shown in the form of box plot

(a) PID on eCPC

(b) PID on AWR
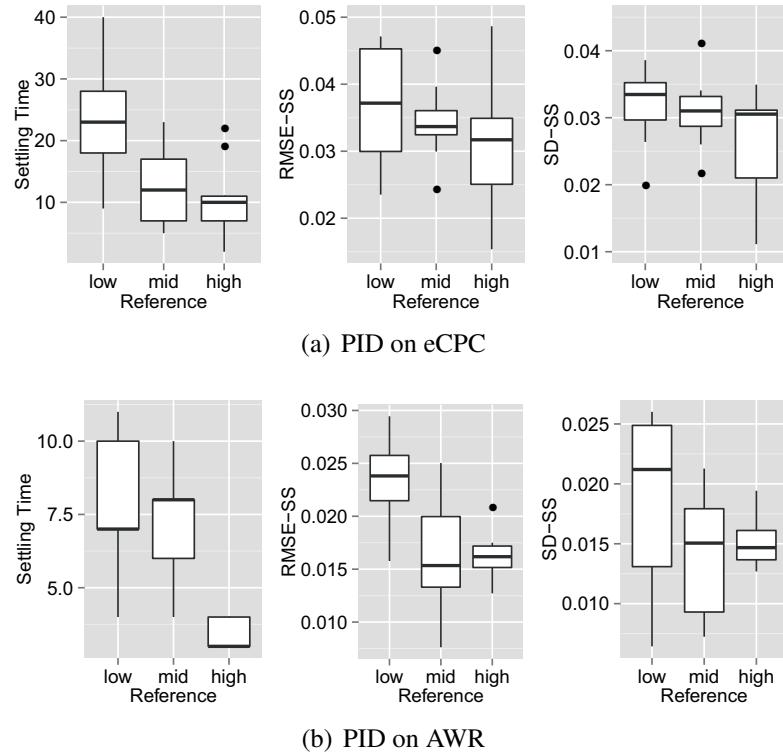
Figure 6.5: Control difficulty comparison with PID.
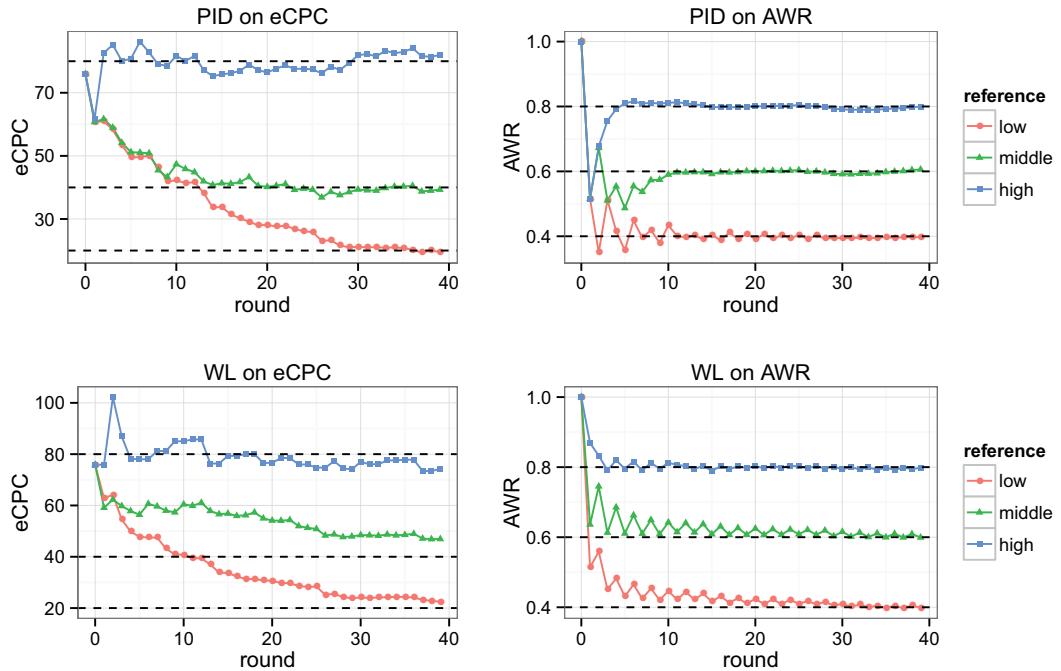
Figure 6.6: Control performance for campaign 3386 on AWR and eCPC with different refer-
ence values.

[116] in Figures 6.5(a) and 6.5(b) for the eCPC and AWR control with PID. We observe

that the average settling time, RMSE-SS and SD-SS, are reduced as the reference values

get higher. This shows that generally the control tasks with higher reference eCPC and

AWR are easier to achieve because one can simply bid higher to win more and spend more. Also as the higher reference is closer to the initial performance value, the control signal does not bring serious bias or volatility, which leads to the lower RMSE-SS and SD-SS.

Figure 6.6 gives the specific control curves of the two controllers with three reference levels on a sample campaign 3386. We find that the reference value which is farthest away from the initial value of the controlled variable brings the largest difficulty for settling, both on eCPC and AWR. This suggests that advertisers setting an ambitious control target will introduce the risk of unsettling or large volatility. The advertisers should try to find a best trade-off between the target value and the practical control performance.

6.4.3 PID Settling: Static vs. Dynamic References

The combination of proportional, integral and derivative factors enables the PID feedback to automatically adjust the settling progress during the control lifetime with high efficiency [151]. Alternatively, one can empirically adjust the reference value in order to achieve the desired reference value. For the example of eCPC control, if the campaign's achieved eCPC is higher than the initial reference value right after exhausting the first half budget, the advertiser might want to lower the reference value in order to accelerate the downward adjustment and finally reach its initial eCPC target before running out of the budget. The PID feedback controller implicitly handles such a problem via its integration factor [151, 152]. In this section, we investigate with our RTB feedback control mechanism whether it is still necessary for advertisers to intentionally adjust the reference value according to the campaign's real-time performance.

Dynamic Reference Adjustment Model. To simulate the advertisers' strategies to adaptively change the reference value of eCPC and AWR under the budget constraint, we propose a dynamic reference adjustment model to calculate the new reference $x_r(t_{k+1})$ after t_k:

$$x_r(t_{k+1}) = \frac{(B - s(t_k))x_r x(t_k)}{Bx(t_k) - s(t_k)x_r}, \tag{6.20}$$

where x_r is the initial reference value, $x(t_k)$ is the achieved KPI (eCPC or AWR) at
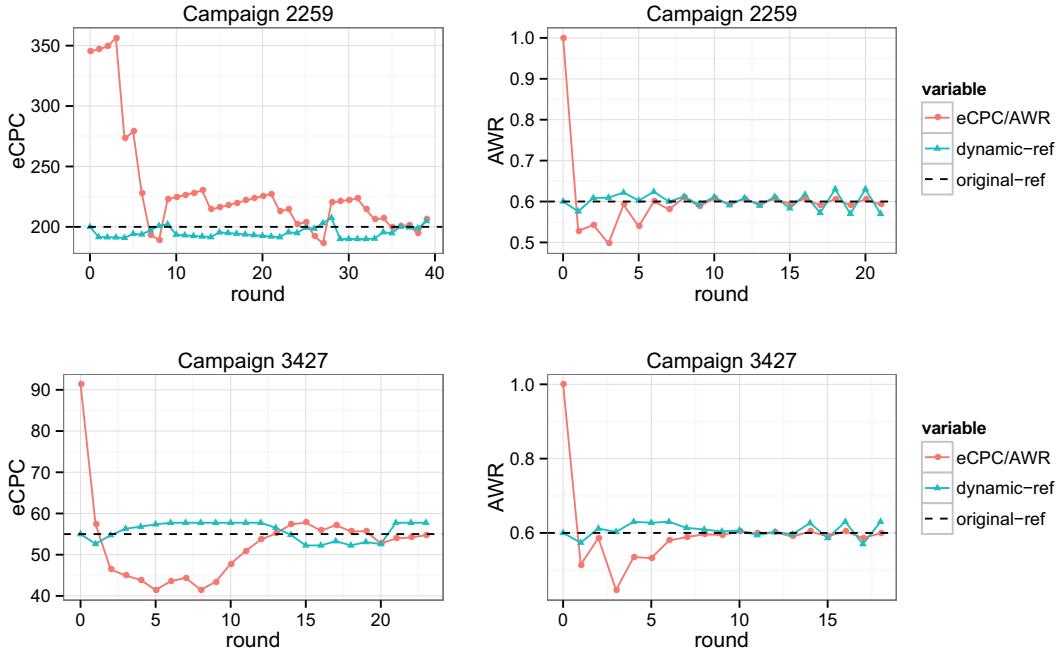
Figure 6.7: Dynamic reference control with PID.

timestep t_k, B is the campaign budget, $s(t_k)$ is the cost so far. We can see from Eq. (6.20) that when $x(t_k) = x_r$, $x_r(t_{k+1})$ will be set the same as x_r; when $x(t_k) > x_r$, $x_r(t_{k+1})$ will be set lower than x_r and vice versa. For readability, we leave the detailed derivation in Section 6.7. Using Eq. (6.20) we calculate the new reference eCPC/AWR $x_r(t_{k+1})$ and use it to substitute x_r in Eq. (6.3) to calculated the error factor so as to make the dynamic-reference control.

Results and Discussions. Figure 6.7 shows the PID control performance with dynamic reference calculated based on Eq. (6.20). The campaign performance gets stopped at the point where the budget is exhausted. From the figure, we see that for both eCPC and AWR control, the dynamic reference takes an aggressive approach and pushes the eCPC or AWR across the original reference value (dashed line). This actually simulates some advertisers' strategy: when the performance is lower than the reference, then higher the dynamic reference to push the total performance to the initial reference more quickly, vice versa. Furthermore, for AWR control, we can see the dynamic reference fluctuates seriously when the budget is to be exhausted soon. This is because when there is insufficient budget left, the reference value will be set much high or low by Eq. (6.20) in order to push the performance back to the initial target. Apparently this is an ineffective solution.

(a) PID on eCPC

(b) PID on AWR
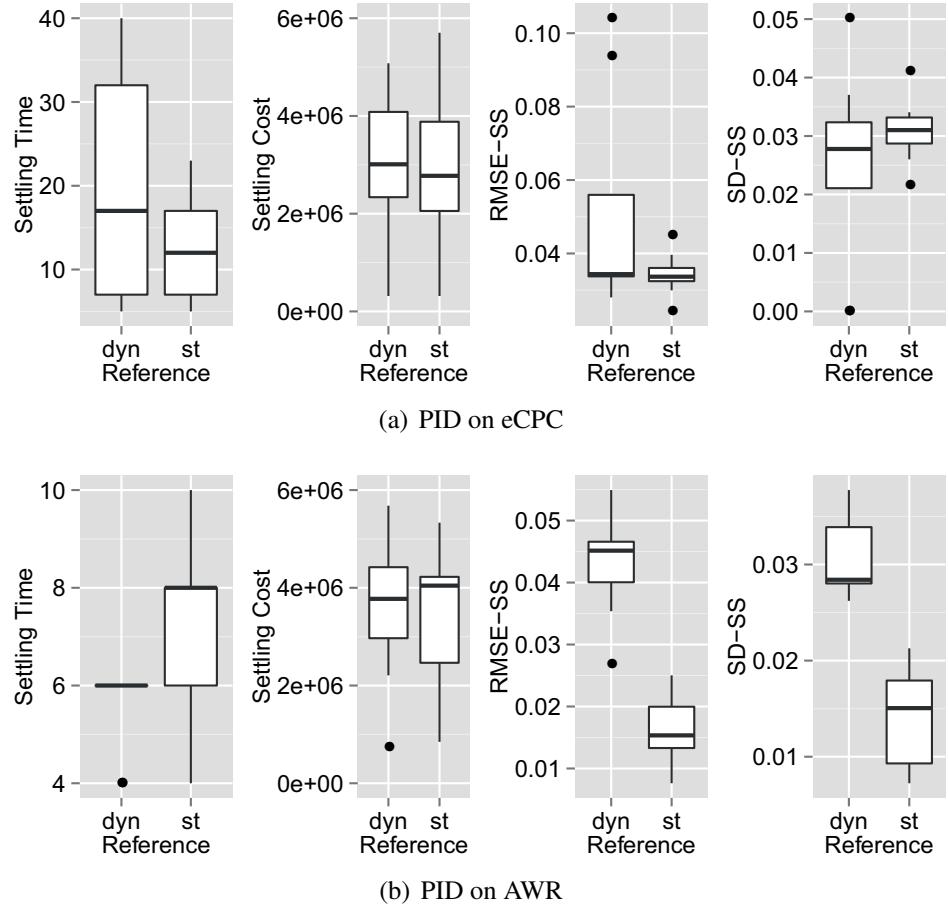
Figure 6.8: Dynamic vs. static reference with PID.

Furthermore, we directly compare the quantitative control performance between dynamic-reference controllers (DYN) with the standard static-reference ones (ST) using PID. Besides the settling time, we also compare the settling cost, which is the spent budget before settling. The overall performance across all the campaigns is shown in Figure 6.8(a) for eCPC control and Figure 6.8(b) for AWR control, respectively. The results show that (i) for eCPC control, the dynamic-reference controllers do *not* perform better than the static-reference ones; (ii) for AWR control, the dynamic-reference controllers could reduce the settling time and cost, but the accuracy (RMSE-SS) and stability (SD-SS) is much worse than the static-reference controllers. This is because the dynamic reference itself brings volatility (see Figure 6.7). These results demonstrate that PID controller does perform a good enough way to setting the variable towards the pre-specified reference without the need of dynamically adjust the reference to accelerate using our methods. Other dynamic reference models might be somewhat effective but this is not the focus of this research.

Table 6.3: Control performance on multi-exchanges with the reference eCPC set for click maximisation.

Campaign	Ad Exchange	Rise	Settling	Campaign	Ad Exchange	Rise	Settling
	1	13	26		1	9	20
1458	2	15	18	3358	2	14	39
	3	13	13		3	26	26
	1	10	38		1	6	18
2259	2	3	14	3386	2	12	12
	3	3	29		3	1	1
	1	3	30		1	16	16
2261	2	7	38	3427	2	35	35
	3	0	35		3	23	23
	1	6	17		1	18	29
2821	2	3	10	3476	2	22	28
	3	15	15		3	19	22
	4	4	38				

6.4.4   Reference Setting for Click Maximisation

We now study how the proposed feedback control could be used for click optimisation purpose. As we have discussed in Section 6.2.4, bid requests usually come from different ad exchanges where the market competitions and thus the CPM prices are disparate. We have shown that given a budget constraint, the number of clicks is maximised if one can control the eCPC in each ad exchange by settling it at an optimal eCPC reference for each of them respectively.

In this experiment, we build a PID feedback controller for each of its integrated ad exchanges, where their reference eCPCs are calculated via Eqs. (6.17) and (6.19). We train the PID parameters on the training data of each campaign, and then test the bidding performance on the test data. As shown in Table 6.3, the eCPC on all the ad exchanges for all tested campaigns get settled at the reference values[4] (settling time less than 40). We denote our multi-exchange eCPC feedback control method as MULTIPLE. Besides MULTIPLE, we also test a baseline method which assigns a single optimal uniform eCPC reference across all the ad exchanges, denoted as UNIFORM. We also use the linear bidding strategy without feedback control [19] as a baseline, denoted as NONE[5].

The comparisons over various evaluation measures are reported in Figure 6.9. We

---

[4]Campaign 2997 is only integrated with one ad exchange, thus not compared here.

[5]Other bidding strategies ([36] and ORTB in Chapter 3) are also investigated. Producing similar results, they are omitted here for clarity.
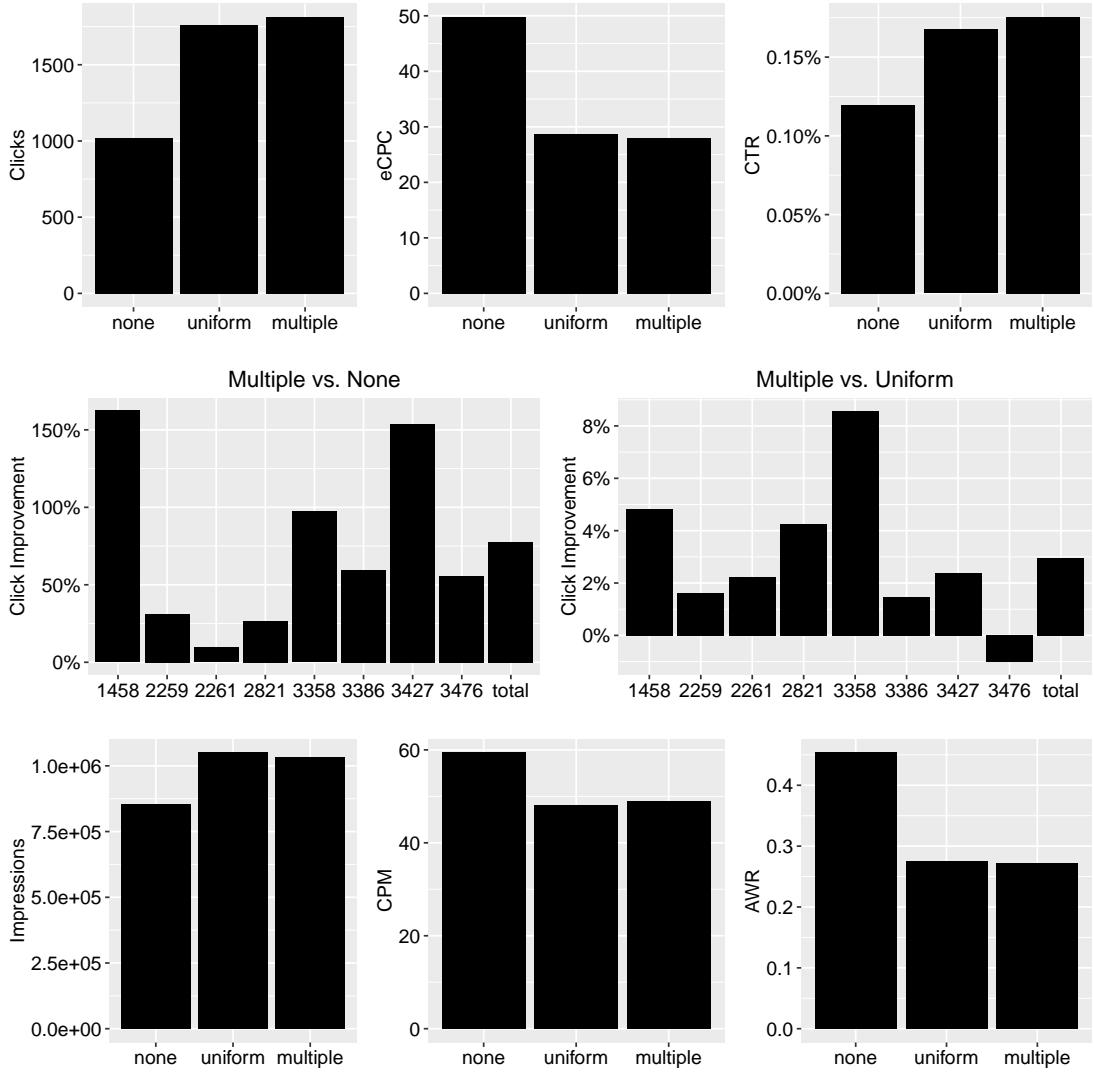
Figure 6.9: Bid optimisation performance.

observe that (i) the feedback-control-enabled bidding strategies UNIFORM and MULTI-PLE significantly outperform the non-controlled bidding strategy NONE in terms of the number of achieved clicks and eCPC. This suggests that properly controlling eCPCs would lead to an optimal solution for maximising clicks. (ii) By reallocating the budget via setting different reference eCPCs on different ad exchanges, MULTIPLE further outperforms UNIFORM on 7 out of 8 tested campaigns. (iii) On the impression related measures, the feedback-control-enabled bidding strategies earn more impressions than the non-controlled bidding strategy by actively lowering their bids (CPM) and thus AWR, but achieving more bid volumes. This suggests that by allocating more budget to the lower valued impressions, one could potentially generate more clicks. As a byproduct, this confirms the theoretical finding reported in Chapter 3.
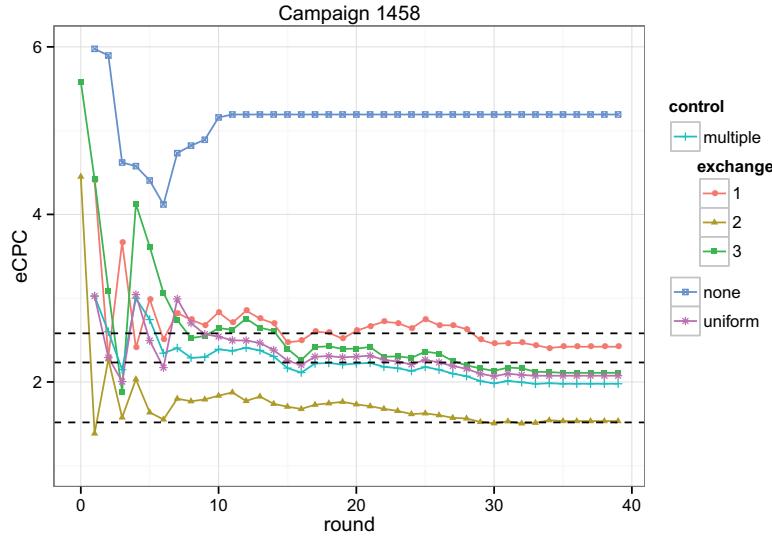
Figure 6.10: Settlement of multi-exchange feedback control.

As a case study, Figure 6.10 plots the settling performance of the three methods on campaign 1458. The three dashed horizontal lines are the reference eCPCs on three ad exchanges. We see that the eCPCs on the three ad exchanges successfully settle at the reference eCPCs. At the same time, the campaign-level eCPC (MULTIPLE) settles at a lower value than UNIFORM and NONE.

6.4.5   PID Parameter Tuning

In this subsection, we share some lessons learned about PID controller parameter tuning and online updating.

Parameter Search. Empirically, λ_D does not change the control performance significantly. Just a small valued λ_D, e.g., 1×10^{-5}, will reduce the overshoot and slightly shorten the settling time. Thus the parameter search is focused on λ_P and λ_I. Instead of using the computationally expensive grid search, we perform an adaptive coordinate search. For every update, we fix one parameter and shoot another one to seek for the optimal value leading shortest settling time, and the line searching step length shrinks exponentially for each shooting. Normally after 3 or 4 iterations, the local optima is reached and we find such a solution is highly comparable with the expensive grid search.

Setting $\phi(t)$ Bounds. We also find that setting up upper/lower bounds of control signal $\phi(t)$ is important to make KPIs controllable. Due to the dynamics in RTB, it is common that user CTR drops during a period, which makes eCPC much higher. The
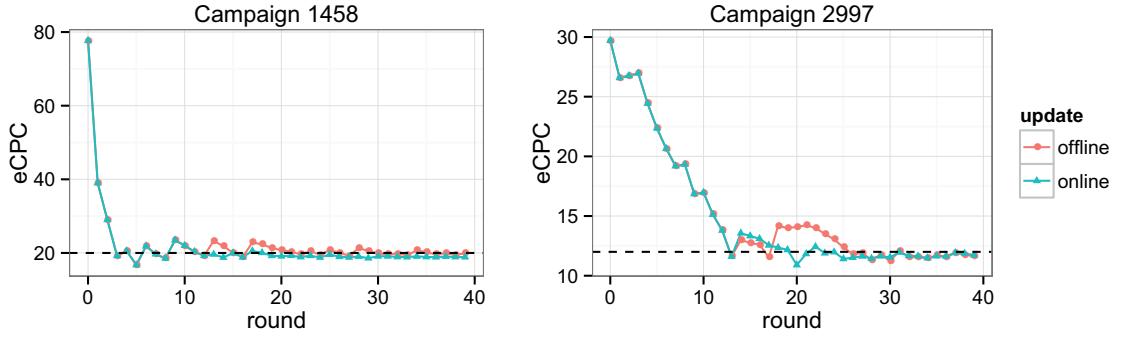
Figure 6.11: Control with online/offline parameter updating.

corresponding feedback would probably result in a large negative gain on the bids, leading extremely low bid prices and thus no win, no click and no additional cost at all for remaining rounds. In such a case, a proper lower bound (-2) of $\phi(t)$ aims to eliminate above extreme effects by preventing from a seriously negative control signal. In addition, an upper bound (5) is used in order to avoid excessive variable growth beyond the reference value.

Online Parameter Updating. As the DSP running with feedback control, the collected data can be immediately utilised for training a new PID controller and updating the older one. We investigate the possibility of the online updating of PID parameters with the recent data. Specifically, after initialising the PID parameters using training data, we re-train the controller for every 10 rounds (i.e., before round 10, 20 and 30) in the test stage using all previous data with the same parameter searching method as in the training stage. The parameter searching in re-training takes about 10 minutes for each controller, which is far shorter than the round period (2 hours). Figure 6.11 shows the control performance with PID parameters tuned online and offline respectively. As we can see after the 10th round (i.e., the first online tuning point), the online-tuned PIDs manage to control the eCPC around the reference value more effectively than the offline-tuned one, resulting shorter settling time and lower overshoot. In addition, no obvious disturbance or instability occurs when we switch parameters. With the online parameter updating, we can start to train the controllers based on several-hour training data and adaptively update the parameters from the new data to improve the control performance.
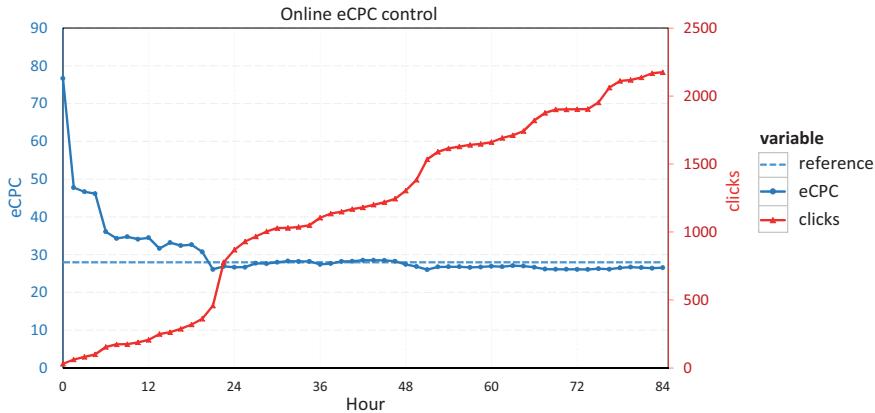
Figure 6.12: The online eCPC control performance and the accumulative click numbers of a
mobile game campaign on BigTree DSP.

6.5   Online A/B Test

The proposed RTB feedback control system has been deployed and tested in live on
BigTree DSP[6], a performance-driven mobile advertising DSP in China. BigTree DSP
focuses on the programmatic buying for optimal advertising performance on mobile
devices, which makes it an ideal place to test our proposed solution.

The deployment environment is based on Aliyun elastic cloud computing servers.
A three-node cluster is deployed for the DSP bidding agent, where each node is in
Ubuntu 12.04, with 8 core Intel Xeon CPU E5-2630 (2.30GHz) and 8GB RAM. The
controller module is implemented in Python with uWSGI and Nginx.

For BigTree DSP controller module, we deploy the PID control function and tune
its parameters. Specifically, we use the last 6-week bidding log data in 2014 as the
training data for tuning PID parameters. A three-fold validation process is performed
to evaluate the generalisation of the PID control performance, where the previous week
data is used as the training data while the later week data is used for validation. The
control factors ($\phi(t), e(t_k)$ in Eq. (6.4)) are updated for every 90 minutes. After ac-
quiring a set of robust and effective PID parameters, we launch the controller module,
including the monitor and actuator submodules, on BigTree DSP.

Figure 6.12 shows the online eCPC control performance on one of the iOS mobile
game campaigns during 84 hours from 7 Jan. 2015 to 10 Jan. 2015. The reference
eCPC is set as 28 CNF (0.28 CNY) by the advertiser, which is about 0.8 times the
average eCPC value of the previous week where there was no control. Following the

---

[6]http://www.bigtree.mobi/
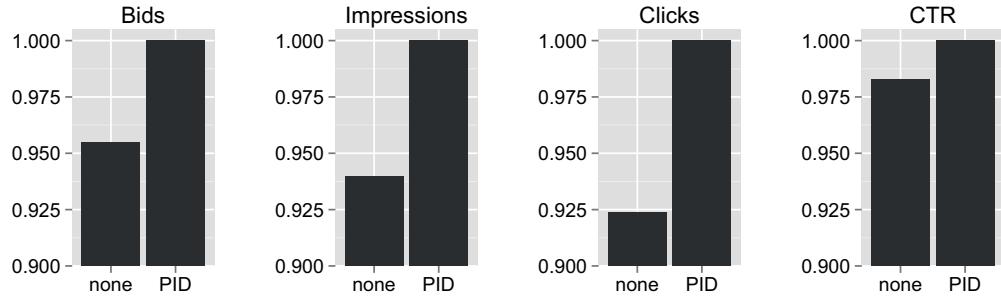
Figure 6.13: Relative performance for online test.

same training process described in the previous section, we update the online control factors for every 90 minutes. From the result we can see the eCPC value dropped from the beginning 79 CNF to 30 CNF during the first day and then settled closed to the reference afterwards.

In the meantime, A/B testing is used to compare with the non-controlled bidding agent (with the same sampling rate but disjoint bid requests). Figure 6.13 shows the corresponding advertising performance comparison between a non-controlled bidding agent and the PID-control bidding agent during the test period with the same budget. As we can see, by settling the eCPC value around the lower reference eCPC, the PID-control bidding agent acquires more bid volume and win more (higher-CTR) impressions and clicks, which demonstrates its ability of optimising the performance.

Compared with the offline empirical study, the online running is more challenging: (i) all pipeline steps including the update of the CTR estimator, the KPI monitor linked to the database and the PID controller should operate smoothly against the market turbulence; (ii) the real market competition is highly dynamic during the new year period when we launched our test; (iii) other competitors might tune their bidding strategies independently or according to any changes of their performance after we employ the controlled bidding strategy. In sum, the successful eCPC control on an online commercial DSP demonstrates the effectiveness of our proposed feedback control RTB system.

6.6 Summary

In this chapter, we have proposed a feedback control mechanism for RTB display advertising, with the aim of improving its robustness of achieving the advertiser's KPI target. We mainly studied PID and WL controllers for controlling the eCPC and AWR KPIs. Through our comprehensive empirical study, we have the following discover-

ies. (i) Despite the high dynamics in RTB, the KPI variables are controllable using our feedback control mechanism. (ii) Different reference values bring different control difficulties, which are reflected in the metrics of control speed, accuracy and stability. (iii) The PID controller naturally finds its best way to settle the variable, and there is no necessity to adjust the reference value for accelerating the PID settling. (iv) By settling the eCPCs to the optimised reference values, the feedback controller is capable of making bid optimisation. Deployed on a commercial DSP, the online test demonstrates the effectiveness of the feedback control mechanism in generating controllable advertising performance. In the future work, we will further study the applications based on feedback controllers in RTB, such as budget pacing and retargeting frequency capping.

6.7 Chapter Appendix: Reference Adjust Models

Here we provide the detailed derivation of the proposed dynamic-reference model Eq. (6.20) in Section 6.4.3.

Reference eCPC Adjustment. Let ξ_r be the initial eCPC target, $\xi(t_k)$ be the achieved eCPC before the moment t_k, $s(t_k)$ be the total cost so far, and B be the campaign budget. In such a setting, the current achieved click number is $s(t_k)/\xi(t_k)$ and the target click number is B/ξ_r. In order to achieve the overall eCPC to ξ_r, i.e., to achieve the total click number B/ξ_r with the budget B, the reference eCPC for the remaining time $\xi_r(t_{k+1})$ should satisfy

$$\frac{s(t_k)}{\xi(t_k)} + \frac{B - s(t_k)}{\xi_r(t_{k+1})} = \frac{B}{\xi_r}. \tag{6.21}$$

Solving the equation we have

$$\xi_r(t_{k+1}) = \frac{(B - s(t_k))\xi_r\xi(t_k)}{B\xi(t_k) - s(t_k)\xi_r}. \tag{6.22}$$

Reference AWR Adjustment. Let ρ_r be the initial AWR target, $\rho(t_k)$ be the achieved AWR before the moment t_k, $n(t_k)$ be the participated auction number so far. If we know the expected total auction volume N during the campaign's lifetime[7], the reference

---

[7] Typically the expected total auction volume during the campaign's lifetime can be estimated, which is a calculation part for bid landscape forecasting problem. See [24] for details.

AWR for the remaining time period $\rho_r(t_{k+1})$ should satisfy

$$n(t_k)\rho(t_k) + (N - n(t_k))\rho_r(t_{k+1}) = N\rho_r. \tag{6.23}$$

However, in our scenario the auction volume is sufficient to run out the campaign budget, which means changing the bid price will directly influence the auction volume that the campaign participates. In such a setting, we consider the budget B and current expense $s(t_k)$. Let $w(z)$ be the expected auction winning ratio for a certain campaign with the bid price z. Here z can also be regarded as the market price, which is the highest bid distribution for the campaign's participated auctions. It is clear that $w(z)$ monotonically increases w.r.t. z. If we want to achieve a winning ratio ρ on the campaign, the expected cost for each impression is $w^{-1}(\rho)$. Furthermore, the achieve impression volume with the expense s is $s/w^{-1}(\rho)$. Then Eq. (6.23) will be rewritten as

$$\frac{s(t_k)}{w^{-1}(\rho(t_k))} + \frac{B - s(t_k)}{w^{-1}(\rho_r(t_{k+1}))} = \frac{B}{w^{-1}(\rho_r)}. \tag{6.24}$$

Different campaigns could have different winning functions $w(z)$, which can be modelled according to their auction data. In Chapter 3, we proposed a concise winning function form $w(z) = z/(z + \varepsilon)$, where ε is the tuning parameter used to fit the data. In such a case, the inverse winning function is $w^{-1}(\rho) = \rho\varepsilon/(1 - \rho)$. Substituting $w^{-1}(\rho)$ into Eq. (6.24), the solution of $\rho_r(t_{k+1})$ is

$$\rho_r(t_{k+1}) = \frac{(B - s(t_k))\rho_r\rho(t_k)}{B\rho(t_k) - s(t_k)\rho_r}, \tag{6.25}$$

which happens to be the same form as eCPC update Eq. (6.22). Using x as a general notation for eCPC and AWR variables leads to Eq. (6.20) in Section 6.4.3.

Chapter 7

Conclusions and Future Work

In this thesis, we have studied the essential research problems of bidding strategy optimisation from the perspective of an advertiser or a DSP in RTB display advertising.

For the fundamental single-campaign bidding function optimisation problem, we have proposed a novel functional optimisation framework which maximises the specific target KPI with constraints of the campaign's lifetime auction volume and budget. This framework is flexible to take in various settings of market environment and user behaviour to derive the corresponding optimal bidding functions, including the widely used linear bidding functions and some novel non-linear bidding functions.

Furthermore, for the advanced multi-campaign statistical arbitrage mining problem, we have proposed a joint optimisation framework to maximise the overall expected arbitrage profit with budget and risk constraints, which is then solved in an EM fashion. In the E-step the auction volume is reallocated according to the individual campaign's estimated risk and return, while in the M-step the arbitrage bidding function is optimised to maximise the expected arbitrage profit with the campaign volume allocation.

On the other hand, to address the problems of biased model learning and optimisation on censored labelled data generated from the auction selection, we have proposed a model-free learning framework which explicitly estimates the probability of observing each training data instance and incorporates such a probability into the learning or optimisation stage. The derived bid-aware gradient descent learning scheme is capable of helping eliminate model bias and yield improved performance in various supervised learning and optimisation tasks.

Last but not least, to deal with the KPI instability problem caused by dynamic RTB ad market competition and user behaviour, we have designed a feedback control

mechanism embedded into the bidding agent to dynamically adjust the bid price so as to control the target KPI variable close to a predefined reference value to achieve the robust and controllable advertising performance.

Besides the scientific innovations, extensive repeatable experiments on large-scale real-world datasets have been performed to verify the effectiveness of the proposed bidding strategy optimisation solutions, unbiased learning schemes and feedback controllers. More importantly, these optimised solutions are quite flexible to any estimation models of the bid request utility and cost, which makes them of light engineering cost to be deployed in production. All the proposed solutions have been (once) deployed in commercial DSP systems[1] and received significant performance improvement in the online A/B tests. In sum, the scientific and empirical contributions of this research are significant in terms of moving towards optimal RTB display advertising performance.

We have also figured out some promising future work. First, it is valuable to further formulate a *learning to bid* problem, which takes the user response prediction, bid landscape modelling and bid optimisation as a single learning system. Each data instance fed into the system is a request-bid-feedback event loop, containing the bid request features, the historic bid price, the auction result and the corresponding user feedback if winning. Second, the value of exploring the uncertainty of user behaviour, market competition and bidding is rare in the existing literature. However, the risk of deficit for performance-driven RTB ad campaigns is a key problem for advertisers and their DSPs. By building Bayesian models of user response rates and bidding functions, risk management can be performed on impression-level bidding process to yield risk-return balance for performance-driven RTB ad campaigns. Moreover, the action of bidding based on bid request features and the later observed auction results and user feedback can be naturally modelled as a reinforcement learning problem. Unlike the keyword-level bidding in sponsored search [28], the bid decision in RTB environment is made based on the real-time bid request features, the associated historic user information and campaign parameters, which make the bidding policy learning unique and challenging.

---

[1]The optimal RTB strategy (from Chapter 3) was deployed on iPinYou Optimus DSP in Dec. 2013. The statistical arbitrage strategy (from Chapter 4) was deployed in Feb. 2015 on BigTree Mobile DSP. Unbiased learning and optimisation techqinues (from Chapter 5) was deployed in Sep. 2015 on Yahoo! DSP. The feedback control mechanism (from Chapter 6) was deployed in Dec. 2014 on BigTree Mobile DSP.

Appendix A

Full List of Publications

Weinan Zhang has published the following conference papers and journal articles during his Ph.D.

1) Weinan Zhang, Tianxiong Zhou, Jun Wang, and Jian Xu. Bid-aware gradient descent for unbiased learning with censored data in display advertising. In *KDD*, 2016. **(Chapter 5)**

2) Weinan Zhang, Tianming Du, Jun Wang. Deep learning over multi-field categorical data: a case study on user response prediction. In *ECIR*, 2016.

3) Weinan Zhang, Lingxi Chen, Jun Wang. Implicit look-alike modelling in display ads: transfer collaborative filtering to CTR estimation . In *ECIR*, 2016.

4) Weinan Zhang, Ulrich Paquet, Katja Hofmann. Collective noise contrastive estimation for policy transfer learning. In *AAAI*, 2016.

5) Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. Feedback control of real-time display advertising. In *WSDM*, 2016. **(Chapter 6)**

6) Weinan Zhang and Jun Wang. Statistical abitrage mining for display advertising. In *KDD*, 2015. **(Chapter 4)**

7) Weinan Zhang, Amr Ahmed, Jie Yang, Vanja Josifovski, and Alex J. Smola. Annotating needles in the haystack without looking: product information extraction from emails. In *KDD*, 2015.

8) Xiaoxue Zhao*, Weinan Zhang*, and Jun Wang. Risk-hedged venture capital investment recommendation. In *RecSys*, 2015.

9) Weinan Zhang, Shuai Yuan, and Jun Wang. Optimal real-time bidding for display advertising. In *KDD*, 2014. **(Chapter 3)**

10) Ying Zhang, Weinan Zhang, Bin Gao, Xiaojie Yuan, and Tie-Yan Liu. Bid keyword suggestion in sponsored search based on competitiveness and relevance. *IPM* 50(4), 2014.

11) Weinan Zhang, Shuai Yuan, Jun Wang, and Xuehua Shen. Real-time bidding benchmarking with ipinyou dataset. *arXiv:1407.7073*, 2014.

12) Weinan Zhang, Jun Wang, Bowei Chen, and Xiaoxue Zhao. To personalize or not: a risk management perspective. In *RecSys*, 2013.

13) Weinan Zhang, Tianqi Chen, Jun Wang, and Yong Yu. Optimizing top-n collaborative filtering via dynamic negative item sampling. In *SIGIR*, 2013 (short paper).

14) Thomas Stone, Weinan Zhang, and Xiaoxue Zhao. An empirical study of top-n recommendation for venture finance. In *CIKM*, 2013 (short paper).

15) Xiaoxue Zhao*, Weinan Zhang*, and Jun Wang. Interactive collaborative filtering. In *CIKM*, 2013.

16) Li Tian, Weinan Zhang, Antonis Bikakis, Haofen Wang, Yong Yu, Yuan Ni, Feng Cao. MeDetect: A LOD-based system for collective entity annotation in biomedicine. In *Web Intelligence*, 2013.

"*" means joint first authorship with equal contribution to the paper.

Appendix B

Glossary of Technical Terms

Here is a summary of the explanations of the mentioned technique terms in the thesis.

Ad Exchange A marketplace which connects the media sellers (publishers) and buyers (advertisers) via network message parsing with a predefined protocol, and select the buyers for each sold media inventory (ad impression) by auctions.

Ad Inventory A notion of the advertising volume regarded as the virtual assets owned by the publisher. The unit of ad inventory is an ad display opportunity, i.e., an ad impression.

Ad Slot A region of the page to place the ad creative.

AWR, Auction Winning Ratio From the micro perspective, AWR means the probability of winning a specific ad auction with a specific bid value; from the macro perspective, AWR means the impression number divided by the participated auction number from a certain volume during a certain period.

Bid, Bid Value, Bid Price The amount of the money the advertiser wants to pay for the ad display opportunity being auctioned.

Bid Optimisation The designing of the bidding function such that the consequent advertising performance, measured by some KPIs, is optimised as mush as possible.

Bidding Agent A functional module of performing bid calculation for each received bid request and a qualified ad in DSP.

Bidding Function The function abstracted from the bidding strategy inputs a bid request and possibly some environment information and outputs the bid price.

Bidding Strategy The bidding logic which inputs a bid request and possibly some environment information and outputs the bid price.

Budget The total amount of money available for advertising cost during a campaign lifetime.

Campaign A series of ads sharing the same advertising target and making up an integrated marketing communication.

Click A click on the ad creative from a page, which directs the user to the landing page of the ad.

CNF, Chinese Fen The unit of Chinese currency, which is 0.01 CNY. On 3 Mar 2016, 653CNF=1USD.

CNY, Chinese Yuan The unit of Chinese currency. On 3 Mar 2016, 6.53CNY=1USD.

Conversion An event showing a user has become a customer of the advertiser. The conversion event can be defined by various of actions, such as a successful page landing, a registration on the advertiser's website, an email subscription, making a deposit, a product purchase etc.

CPA, Cost per Action or Cost per Acquisition A predefined amount of money the advertiser pays the ad agent (DSP in RTB display advertiser, search engine in sponsored search) when a specified action has been observed on the delivered ad impression. The action can be defined by various of actions, such as a successful page landing, a registration on the advertiser's website, an email subscription, making a deposit, a product purchase etc.

CPC, Cost per Click A predefined amount of money the advertiser pays the ad agent (DSP in RTB display advertiser, search engine in sponsored search) when a user click has been observed on the delivered ad impression.

CPM, Cost per Mille A predefined amount of money the advertiser pays the ad agent (DSP in RTB display advertiser, search engine in sponsored search) for each delivered ad impression, often counted by one thousand of the same cases of ad impressions.

CPS, Cost per Sale A predefined amount of money the advertiser pays the ad agent (DSP in RTB display advertiser, search engine in sponsored search) when a specified sale has been made after the user sees the ad impression.

Creative The content of a specific ad, often in the format of images for display advertising and text for sponsored search. Javascript based creatives are also allowed in some ad exchanges to enable interactive creatives. The hyperlink on the creative points to the landing page that the advertiser wants the user to browse.

CTR, Click-Through Rate From the micro perspective, CTR means the probability of a specific user in a specific context clicking a specific ad; from the macro perspective, CTR means the click number divided by the impression number from a certain volume during a certain period.

CVR, Conversion Rate From the micro perspective, CVR means the probability of the user conversion is observed after showing the ad impression; from the macro perspective, CVR means the conversion number divided by the impression number from a certain volume during a certain period.

DMP, Data Management Platform The platform which collects, analyses and trades user behaviour information. DSPs are its major clients.

DSP, Demand-Side Platform The platform which serves advertisers to manage their campaigns and submits real-time bidding responses for each bid request to the ad exchange via computer algorithms.

eCPA, Effective Cost per Action (or Acquisition) The average cost for acquiring an action, also called efficient cost per action or expected cost per action in some references.

eCPC, Effective Cost per Click The average cost for acquiring a click, also called efficient cost per click or expected cost per click in some references.

First-Price Auction The auction where the winner, i.e., the participator with the highest bid value, pays her bid value.

Impression An ad display in front of the user.

KPI, Key Performance Indicator A certain quantitative measurement of advertising performance, such as impression number, click number, conversion number, CPM, eCPC, eCPA, AWR, CTR etc.

Market Price A different name of winning price, defined on a specific bid request, which means the lowest bid value to win the auction of this bid request, i.e., the highest bid value from other competitors of this auction.

ROI, Return on Investment The ratio of the profit (revenue minus cost) gained from advertising over the advertising cost.

RTB, Real-Time Bidding A display ads trading mechanism where the ad inventory is traded on impression level via an instant ad auction with the bid values returned from the advertisers calculated in real time, e.g., less than 100ms.

Second-Price Auction The auction where the winner, i.e., the participant with the highest bid value, pays the second highest bid.

SSP, Supply-Side Platform The platform which serves publishers to manage the ad inventory of the sites. Upon each page loading, the SSP sends the ad request for each of the RTB ad slot to the ad exchange. Once the ad exchange returned the ID or code of the winning ad, SSP calls the corresponding ad server for the ad creative.

User Segmentation The subsets of users divided by users' demographical information, e.g., age, gender, location and occupation, or interest categories or tags. Normally, user segmentation is provided by DMP or ad exchange to help advertisers perform demographical or behavioural targeting. The bidding strategy can also highly leverage such information to perform effective bidding.

Winning Price A different name of market price, defined on a specific bid request, which means the lowest bid value to win the auction of this bid request.

Bibliography

[1] Matthew Richardson, Ewa Dominowska, and Robert Ragno. Predicting clicks: estimating the click-through rate for new ads. In *Proceedings of the 16th international conference on World Wide Web*, pages 521–530. ACM, 2007.

[2] Brian Dalessandro, Rod Hook, Claudia Perlich, and Foster Provost. Evaluating and optimizing online advertising: Forget the click, but there are good proxies. *Big Data*, 3(2):90–102, 2015.

[3] Deepak Agarwal, Bo Long, Jonathan Traupman, Doris Xin, and Liang Zhang. Laser: A scalable response prediction platform for online advertising. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 173–182. ACM, 2014.

[4] Kuang-chih Lee, Burkay Orten, Ali Dasdan, and Wentong Li. Estimating conversion rate in display advertising from past erformance data. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 768–776. ACM, 2012.

[5] PwC and Interactive Advertising Bureau. IAB internet advertising revenue report: 2014 full year results. 2015.

[6] eMarketer. UK ad spending: Q1 2014 forecast and comparative estimates, 2015.

[7] eMarketer. Total media ad spending growth slows worldwide. Online report published at `http://www.emarketer.com/Article/Total-Media-Ad-Spending-Growth-Slows-Worldwide/1012981`, 2015. Accessed: 2015-11-04.

[8] PwC and Interactive Advertising Bureau. IAB internet advertising revenue report: 2015 first six months results. 2015.

[9] 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, 2006.

[10] Benjamin Edelman, Michael Ostrovsky, and Michael Schwarz. Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. Technical report, National Bureau of Economic Research, 2005.

[11] Avi Goldfarb and Catherine Tucker. Online display advertising: Targeting and obtrusiveness. *Marketing Science*, 30(3):389–404, 2011.

[12] Amr Ahmed, Yucheng Low, Mohamed Aly, Vanja Josifovski, and Alexander J Smola. Scalable distributed inference of dynamic user interests for behavioral targeting. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 114–122. ACM, 2011.

[13] PwC and Interactive Advertising Bureau. IAB internet advertising revenue report: 2010 full year results. 2011.

[14] Shuai Yuan, Jun Wang, and Xiaoxue Zhao. Real-time bidding for online advertising: measurement and analysis. In *Proceedings of the Seventh International Workshop on Data Mining for Online Advertising*, page 3. ACM, 2013.

[15] Rómer Rosales, Haibin Cheng, and Eren Manavoglu. Post-click conversion modeling and analysis for non-guaranteed delivery display advertising. In *Proceedings of the fifth ACM international conference on Web search and data mining*, pages 293–302. ACM, 2012.

[16] S Muthukrishnan. Ad exchanges: Research issues. In *Internet and network economics*, pages 1–12. Springer, 2009.

[17] Google. The arrival of real-time bidding. 2011.

[18] Arpita Ghosh and Aaron Roth. Selling privacy at auction. *Games and Economic Behavior*, 2013.

[19] Claudia Perlich, Brian Dalessandro, Rod Hook, Ori Stitelman, Troy Raeder, and Foster Provost. Bid optimizing and inventory scoring in targeted online advertising. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 804–812. ACM, 2012.

[20] Lampros C Stavrogiannis, Enrico H Gerding, and Maria Polukarov. Auction mechanisms for demand-side intermediaries in online advertising exchanges. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, pages 1037–1044. International Foundation for Autonomous Agents and Multiagent Systems, 2014.

[21] Ye Chen, Pavel Berkhin, Bo Anderson, and Nikhil 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. ACM, 2011.

[22] Richard J Oentaryo, Ee-Peng Lim, Jia-Wei Low, David Lo, and Michael Finegold. Predicting response in mobile advertising with hierarchical importance-aware factorization machine. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 123–132. ACM, 2014.

[23] Amr Ahmed, Abhimanyu Das, and Alexander J Smola. Scalable hierarchical multitask learning algorithms for conversion optimization in display advertising. In *Proceedings of the 7th ACM international conference on Web search and data mining*, pages 153–162. ACM, 2014.

[24] Ying Cui, Ruofei Zhang, Wei Li, and Jianchang Mao. Bid landscape forecasting in online ad exchange marketplace. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 265–273. ACM, 2011.

[25] Michael Durbin. *All about high-frequency trading*. McGraw Hill Professional, 2010.

[26] Morton I Kamien and Nancy L Schwartz. *Dynamic optimization: the calculus of variations and optimal control in economics and management*. Courier Corporation, 2012.

[27] Weinan Zhang, Jun Wang, Bowei Chen, and Xiaoxue Zhao. To personalize or not: a risk management perspective. In *Proceedings of the 7th ACM Conference on Recommender Systems*, pages 229–236. ACM, 2013.

[28] Kareem Amin, Michael Kearns, Peter Key, and Anton Schwaighofer. Budget optimization for sponsored search: Censored learning in MDPs. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence*, 2012.

[29] Weinan Zhang, Shuai Yuan, Jun Wang, and Xuehua Shen. Real-time bidding benchmarking with iPinYou dataset. *arXiv preprint arXiv:1407.7073*, 2014.

[30] Guillaume Roels and Kristin Fridgeirsdottir. Dynamic revenue management for online display advertising. *Journal of Revenue & Pricing Management*, 8(5):452–466, 2009.

[31] Vijay Bharadwaj, Wenjing Ma, Michael Schwarz, Jayavel Shanmugasundaram, Erik Vee, Jack Xie, and Jian Yang. Pricing guaranteed contracts in online display advertising. In *Proceedings of the 19th ACM international conference on Information and knowledge management*, pages 399–408. ACM, 2010.

[32] Anand Bhalgat, Jon Feldman, and Vahab Mirrokni. Online allocation of display ads with smooth delivery. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1213–1221. ACM, 2012.

[33] Jon Feldman, Monika Henzinger, Nitish Korula, Vahab S Mirrokni, and Cliff Stein. Online stochastic packing applied to display ad allocation. In *Algorithms– ESA 2010*, pages 182–194. Springer, 2010.

[34] Jian Yang, Erik Vee, Sergei Vassilvitskii, John Tomlin, Jayavel Shanmugasundaram, Tasos Anastasakos, and Oliver Kennedy. Inventory allocation for online graphical display advertising. *arXiv preprint arXiv:1008.3551*, 2010.

[35] Vijay Bharadwaj, Peiji Chen, Wenjing Ma, Chandrashekhar Nagarajan, John Tomlin, Sergei Vassilvitskii, Erik Vee, and Jian Yang. Shale: an efficient algorithm for allocation of guaranteed display advertising. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1195–1203. ACM, 2012.

[36] Kuang-Chih Lee, Ali Jalali, and Ali 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*, page 1. ACM, 2013.

[37] Shuai Yuan, Jun Wang, Bowei Chen, Peter Mason, and Sam Seljan. An empirical study of reserve price optimisation in real-time bidding. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1897–1906. ACM, 2014.

[38] Arpita Ghosh, Preston McAfee, Kishore Papineni, and Sergei Vassilvitskii. Bidding for representative allocations for display advertising. In *Internet and Network Economics*, pages 208–219. Springer, 2009.

[39] Bowei Chen, Shuai Yuan, and Jun Wang. A dynamic pricing model for unifying programmatic guarantee and real-time bidding in display advertising. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, pages 1–9. ACM, 2014.

[40] Jon Feldman, S Muthukrishnan, Martin Pal, and Cliff Stein. Budget optimization in search-based advertising auctions. In *Proceedings of the 8th ACM conference on Electronic Commerce*, pages 40–49. ACM, 2007.

[41] Kartik Hosanagar and Vadim Cherepanov. Optimal bidding in stochastic budget constrained slot auctions. In *Proceedings of the 9th ACM conference on Electronic Commerce*, pages 20–20. ACM, 2008.

[42] Arpita Ghosh, Benjamin IP Rubinstein, Sergei Vassilvitskii, and Martin Zinkevich. Adaptive bidding for display advertising. In *Proceedings of the 18th international conference on World Wide Web*, pages 251–260. ACM, 2009.

[43] Animesh Animesh, Vandana Ramachandran, and Siva Viswanathan. Online advertisers bidding strategies for search, experience, and credence goods: An empirical investigation. In *Second Workshop on Sponsored Search Auctions. EC*. Citeseer, 2005.

[44] Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5):22, 2007.

[45] Christian Borgs, Jennifer Chayes, Nicole Immorlica, Kamal Jain, Omid Etesami, and Mohammad Mahdian. Dynamics of bid optimization in online advertisement auctions. In *Proceedings of the 16th international conference on World Wide Web*, pages 531–540. ACM, 2007.

[46] Yunhong Zhou, Deeparnab Chakrabarty, and Rajan Lukose. Budget constrained bidding in keyword auctions and online knapsack problems. In *Internet and Network Economics*, pages 566–576. 2008.

[47] Brendan Kitts and Benjamin Leblanc. Optimal bidding on keyword auctions. *Electronic Markets*, 14(3):186–201, 2004.

[48] Ying Zhang, Weinan Zhang, Bin Gao, Xiaojie Yuan, and Tie-Yan Liu. Bid keyword suggestion in sponsored search based on competitiveness and relevance. *Information Processing & Management*, 50(4):508–523, 2014.

[49] S Muthukrishnan, Martin Pál, and Zoya Svitkina. Stochastic models for budget optimization in search-based advertising. In *Internet and Network Economics*. 2007.

[50] Andrei Broder, Evgeniy Gabrilovich, Vanja Josifovski, George Mavromatis, and Alex Smola. Bid generation for advanced match in sponsored search. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 515–524. ACM, 2011.

[51] Eyal Even Dar, Vahab S Mirrokni, S Muthukrishnan, Yishay Mansour, and Uri Nadav. Bid optimization for broad match ad auctions. In *Proceedings of the 18th international conference on World wide web*, pages 231–240. ACM, 2009.

[52] Weinan Zhang, Ying Zhang, Bin Gao, Yong Yu, Xiaojie Yuan, and Tie-Yan Liu. Joint optimization of bid and budget allocation in sponsored search. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1177–1185. ACM, 2012.

[53] R Gummadi, Peter B Key, and Alexandre Proutiere. Optimal bidding strategies in dynamic auctions with budget constraints. In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pages 588–588. IEEE, 2011.

[54] Zoë Abrams. Revenue maximization when bidders have budgets. In *Proceedings of the seventeenth annual ACM-SIAM symposium on Discrete algorithm*, pages 1074–1082. Society for Industrial and Applied Mathematics, 2006.

[55] Filip Radlinski, Andrei Broder, Peter Ciccolo, Evgeniy Gabrilovich, Vanja Josifovski, and Lance Riedel. Optimizing relevance and revenue in ad search: a query substitution approach. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 403–410. ACM, 2008.

[56] Nikhil R Devanur and Thomas P Hayes. The adwords problem: online keyword matching with budgeted bidders under random permutations. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 71–78. ACM, 2009.

[57] Yunzhang Zhu, Gang Wang, Junli Yang, Dakan Wang, Jun Yan, Jian Hu, and Zheng Chen. Optimizing search engine revenue in sponsored search. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 588–595. ACM, 2009.

[58] Jian Xu, Xuhui Shao, Jianjie Ma, Kuang-chih Lee, Hang Qi, and Quan Lu. Lift-based bidding in ad selection. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, 2016.

[59] Ron Berman. Beyond the last touch: Attribution in online advertising. *Available at SSRN 2384211*, 2013.

[60] Tanmoy Chakraborty, Eyal Even-Dar, Sudipto Guha, Yishay Mansour, and S Muthukrishnan. Selective call out and real time bidding. In *Internet and Network Economics*, pages 145–157. 2010.

[61] Johnathan Mun. *Modeling risk: Applying Monte Carlo simulation, real options analysis, forecasting, and optimization techniques*, volume 347. John Wiley & Sons, 2006.

[62] Michel Crouhy, Dan Galai, and Robert Mark. *The essentials of risk management*, volume 1. McGraw-Hill New York, 2006.

[63] Ricardo Antunes and Vicente Gonzalez. A production model for construction: A theoretical framework. *Buildings*, 5(1):209–228, 2015.

[64] Jun Wang and Bowei Chen. Selling futures online advertising slots via option contracts. In *Proceedings of the 21st international conference companion on World Wide Web*, pages 627–628. ACM, 2012.

[65] Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1):77–91, 1952.

[66] John Hull. *Risk Management and Financial Institutions*, volume 733. Wiley, 2012.

[67] William F Sharpe. The sharpe ratio. *Streetwise–the Best of the Journal of Portfolio Management*, 1998.

[68] Edwin J Elton, Martin J Gruber, Stephen J Brown, and William N Goetzmann. *Modern portfolio theory and investment analysis*. John Wiley & Sons, 2009.

[69] Jianhan Zhu, Jun Wang, Ingemar J Cox, and Michael J Taylor. Risky business: modeling and exploiting uncertainty in information retrieval. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 99–106. ACM, 2009.

[70] Lidan Wang, Paul N Bennett, and Kevyn Collins-Thompson. Robust ranking models via risk-sensitive optimization. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 761–770. ACM, 2012.

[71] Kevyn Collins-Thompson. Reducing the risk of query expansion via robust constrained optimization. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 837–846. ACM, 2009.

[72] Jun Wang and Jianhan Zhu. Portfolio theory of information retrieval. In *Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval*, pages 115–122. ACM, 2009.

[73] Yue Shi, Xiaoxue Zhao, Jun Wang, Martha Larson, and Alan Hanjalic. Adaptive diversification of recommendation results via latent factor portfolio. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 175–184. ACM, 2012.

[74] Hengshu Zhu, Hui Xiong, Yong Ge, and Enhong Chen. Mobile app recommendations with security and privacy awareness. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 951–960. ACM, 2014.

[75] Ruslan Salakhutdinov and Andriy Mnih. Bayesian probabilistic matrix factorization using markov chain monte carlo. In *Proceedings of the 25th international conference on Machine learning*, pages 880–887. ACM, 2008.

[76] Hongke Zhao, Le Wu, Qi Liu, Yong Ge, and Enhong Chen. Investment recommendation in p2p lending: A portfolio perspective with risk management. In *Data Mining (ICDM), 2014 IEEE International Conference on*, pages 1109–1114. IEEE, 2014.

[77] Xiaoxue Zhao, Weinan Zhang, and Jun Wang. Risk-hedged venture capital investment recommendation. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 75–82. ACM, 2015.

[78] Dell Zhang and J Lu. Batch-mode computational advertising based on modern portfolio theory. *Advances in Information Retrieval Theory*, 386(5766):380–383, 2009.

[79] Steve Hogan, Robert Jarrow, Melvyn Teo, and Mitch Warachka. Testing market efficiency using statistical arbitrage with applications to momentum and value strategies. *Journal of Financial economics*, 73(3):525–565, 2004.

[80] Oleg Bondarenko. Statistical arbitrage and securities prices. *Review of Financial Studies*, 16(3):875–919, 2003.

[81] William K Bertram. Analytic solutions for optimal statistical arbitrage trading. *Physica A*, 389(11):2234–2243, 2010.

[82] Evan Gatev, William N Goetzmann, and K Geert Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *Review of Financial Studies*, 19(3):797–827, 2006.

[83] Ruggiero Cavallo, R Preston Mcafee, and Sergei Vassilvitskii. Display advertising auctions with arbitrage. *ACM Transactions on Economics and Computation*, 3(3):15, 2015.

[84] Lihong Li. Offline evaluation and optimization for interactive systems. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, pages 413–414. ACM, 2015.

[85] Léon Bottou, Jonas Peters, Joaquin Quinonero-Candela, Denis X Charles, D Max Chickering, Elon Portugaly, Dipankar Ray, Patrice Simard, and Ed Snelson. Counterfactual reasoning and learning systems: The example of computational advertising. *The Journal of Machine Learning Research*, 14(1):3207–3260, 2013.

[86] Lihong Li, Wei Chu, John Langford, and Xuanhui Wang. Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms. In *Proceedings of the fourth ACM international conference on Web search and data mining*, pages 297–306. ACM, 2011.

[87] John Langford, Alexander Strehl, and Jennifer Wortman. Exploration scavenging. In *Proceedings of the 25th international conference on Machine learning*, pages 528–535. ACM, 2008.

[88] Miroslav Dudík, Dumitru Erhan, John Langford, and Lihong Li. Sample-efficient nonstationary policy evaluation for contextual bandits. In *Proceedings of the 28th Conference on Uncertainty in Artificial Intelligence*, 2012.

[89] Alex Strehl, John Langford, Lihong Li, and Sham M Kakade. Learning from logged implicit exploration data. In *Advances in Neural Information Processing Systems*, pages 2217–2225, 2010.

[90] Pedro J García-Laencina, José-Luis Sancho-Gómez, and Aníbal R Figueiras-Vidal. Pattern classification with missing data: a review. *Neural Computing and Applications*, 19(2):263–282, 2010.

[91] Rong Pan, Yunhong Zhou, Bin Cao, Nathan N Liu, Rajan Lukose, Martin Scholz, and Qiang Yang. One-class collaborative filtering. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pages 502–511. IEEE, 2008.

[92] Benjamin M Marlin and Richard S Zemel. Collaborative prediction and ranking with non-random missing data. In *Proceedings of the third ACM conference on Recommender systems*, pages 5–12. ACM, 2009.

[93] Wush Chi-Hsuan Wu, Mi-Yen Yeh, and Ming-Syan Chen. Predicting winning price in real time bidding with censored data. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1305–1314. ACM, 2015.

[94] Thore Graepel, Joaquin Q Candela, Thomas Borchert, and Ralf Herbrich. Web-scale bayesian click-through rate prediction for sponsored search advertising in microsoft's bing search engine. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 13–20, 2010.

[95] Karl Johan Astrom and Richard M Murray. *Feedback systems: an introduction for scientists and engineers*. Princeton university press, 2010.

[96] Karl Johan Åström and PR Kumar. Control: A perspective. *Automatica*, 50(1):3–43, 2014.

[97] Otto Mayr. The origins of feedback control. *Scientific American*, 223(4):110–118, 1970.

[98] Otto Mayr and Lynwood Bryant. The origins of feedback control. *Systems, Man and Cybernetics, IEEE Transactions on*, 1(4), 1971.

[99] Ronald Kline. Harold black and the negative-feedback amplifier. *Control Systems, IEEE*, 13(4):82–85, 1993.

[100] Dennis S Bernstein. Feedback control: an invisible thread in the history of technology. *Control Systems, IEEE*, 22(2):53–68, 2002.

[101] Thomas L Dean, Kenneth Basye, Robert Chekaluk, Seungseok Hyun, Moises Lejter, and Margaret Randazza. Coping with uncertainty in a control system for navigation and exploration. In *Proceedings of 15th AAAI Conference on Artificial Intelligence*, pages 1010–1015, 1990.

[102] Peter Lute and Dolf van Paassen. Optimal indoor temperature control using a predictor. *Control Systems, IEEE*, 15(4):4–10, 1995.

[103] Kui Meng, Yue Wang, Xu Zhang, Xiao-chun Xiao, and Geng-du Zhang. Control theory based rating recommendation for reputation systems. In *Networking, Sensing and Control, 2006. ICNSC'06. Proceedings of the 2006 IEEE International Conference on*, pages 162–167. IEEE, 2006.

[104] Valentina Zanardi and Licia Capra. Dynamic updating of online recommender systems via feed-forward controllers. In *Proceedings of the 6th International Symposium on Software Engineering for Adaptive and Self-Managing Systems*, pages 11–19. ACM, 2011.

[105] Tamas Jambor, Jun Wang, and Neal Lathia. Using control theory for stable and efficient recommender systems. In *Proceedings of the 21st international conference on World Wide Web*, pages 11–20. ACM, 2012.

[106] Niklas Karlsson and Jianlong Zhang. Applications of feedback control in online advertising. In *American Control Conference (ACC), 2013*, pages 6008–6013. IEEE, 2013.

[107] Andrei Broder, Marcus Fontoura, Vanja Josifovski, and Lance Riedel. A semantic approach to contextual advertising. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 559–566. ACM, 2007.

[108] Joanna Jaworska and Marcin Sydow. Behavioural targeting in on-line advertising: An empirical study. In *Web Information Systems Engineering-WISE 2008*, pages 62–76. Springer, 2008.

[109] Vijay Krishna. *Auction theory*. Academic press, 2009.

[110] Suleyman Cetintas, Datong Chen, and Luo Si. Forecasting user visits for online display advertising. *Information Retrieval*, 16(3):369–390, 2013.

[111] Kilian Weinberger, Anirban Dasgupta, John Langford, Alex Smola, and Josh Attenberg. Feature hashing for large scale multitask learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 1113–1120. ACM, 2009.

[112] Olivier Chapelle. Modeling delayed feedback in display advertising. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1097–1105. ACM, 2014.

[113] Shuai Yuan and Jun Wang. Sequential selection of correlated ads by pomdps. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, pages 515–524. ACM, 2012.

[114] Flavio M Menezes and Paulo Klinger Monteiro. *An introduction to auction theory*. Oxford University Press, USA, 2005.

[115] Hairen Liao, Lingxiao Peng, Zhenchuan Liu, and Xuehua Shen. iPinYou global rtb bidding algorithm competition dataset. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, pages 1–6. ACM, 2014.

[116] Robert McGill, John W Tukey, and Wayne A Larsen. Variations of box plots. *The American Statistician*, 1978.

[117] Weinan Zhang, Tianqi Chen, Jun Wang, and Yong Yu. Optimizing top-n collaborative filtering via dynamic negative item sampling. In *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pages 785–788. ACM, 2013.

[118] Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. An experimental comparison of click position-bias models. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*, pages 87–94. ACM, 2008.

[119] DSP bidding optimization task. `http://contest.ipinyou.com/rule.shtml`. Accessed: 09-Apr-2016.

[120] Steven Kuhn. Prisoner's dilemma. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, chapter 4. Fall 2014 edition, 2014.

[121] Anatol Rapoport and Albert M Chammah. *Prisoner's dilemma: A study in conflict and cooperation*, volume 165. University of Michigan press, 1965.

[122] Burton G Malkiel and Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

[123] Doni Bloomfield. High-speed ad traders profit by arbitraging your eyeballs. `http://www.bloomberg.com/news/articles/2014-11-07/high-speed-ad-traders-profit-by-arbitraging-your-eyeballs`, 2014. Accessed: 09-Apr-2016.

[124] R Preston McAfee and Sergei Vassilvitskii. An overview of practical exchange design. *Current Science*, 2012.

[125] Yu Jeffrey Hu. Performance-based pricing models in online advertising. *Available at SSRN 501082*, 2004.

[126] Benjamin Edelman. The design of online advertising markets. *Handbook of Market Design*, 2010.

[127] Brian Dalessandro, Claudia Perlich, Ori Stitelman, and Foster Provost. Causally motivated attribution for online advertising. In *Proceedings of the Sixth Interna-*

tional Workshop on Data Mining for Online Advertising and Internet Economy, page 7. ACM, 2012.

[128] Weinan Zhang, Shuai Yuan, and Jun Wang. Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1077–1086. ACM, 2014.

[129] Walter R Gilks. *Markov chain monte carlo*. Wiley Online Library, 2005.

[130] Alan E Gelfand and Adrian FM Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American statistical association*, 85(410):398–409, 1990.

[131] Walter R Gilks and Pascal Wild. Adaptive rejection sampling for gibbs sampling. *Applied Statistics*, pages 337–348, 1992.

[132] Arnaud Doucet, Simon Godsill, and Christophe Andrieu. On sequential monte carlo sampling methods for bayesian filtering. *Statistics and computing*, 10(3):197–208, 2000.

[133] Nittala Chaitanya and Yadati Narahari. Optimal equilibrium bidding strategies for budget constrained bidders in sponsored search auctions. *Operational Research*, 12(3):317–343, 2012.

[134] H Brendan McMahan, Gary Holt, David Sculley, Michael Young, Dietmar Ebner, Julian Grady, Lan Nie, Todd Phillips, Eugene Davydov, Daniel Golovin, et al. Ad click prediction: a view from the trenches. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1222–1230. ACM, 2013.

[135] Hal R Varian. Online ad auctions. *The American Economic Review*, pages 430–434, 2009.

[136] Edward L Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American statistical association*, 53(282):457–481, 1958.

[137] Keisuke Hirano, Guido W Imbens, and Geert Ridder. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4):1161–1189, 2003.

[138] Norman L Johnson. *Survival models and data analysis*, volume 74. John Wiley & Sons, 1999.

[139] Dorota M Dabrowska. Non-parametric regression with censored survival time data. *Scandinavian Journal of Statistics*, pages 181–197, 1987.

[140] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, et al. Practical lessons from predicting clicks on ads at facebook. In *Proceedings of the Eighth International Workshop on Data Mining for Online Advertising*, pages 1–9. ACM, 2014.

[141] Xiaoxue Zhao, Weinan Zhang, and Jun Wang. Interactive collaborative filtering. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management*, pages 1411–1420. ACM, 2013.

[142] Hirotaka Hachiya, Takayuki Akiyama, Masashi Sugiayma, and Jan Peters. Adaptive importance sampling for value function approximation in off-policy reinforcement learning. *Neural Networks*, 22(10):1399–1410, 2009.

[143] Robert C Nelson. *Flight stability and automatic control*, volume 2. WCB/McGraw Hill, 1998.

[144] Gilbert Peterson and Diane J Cook. Decision-theoretic layered robotic control architecture. In *Proceedings of the sixteenth national conference on Artificial intelligence and the eleventh Innovative applications of artificial intelligence conference innovative applications of artificial intelligence*, page 976. American Association for Artificial Intelligence, 1999.

[145] Stuart Bennett. Development of the pid controller. *Control Systems*, 13(6):58–62, 1993.

[146] Bernard W Dezotell. Water level controller, June 9 1936. US Patent 2,043,530.

[147] Lampros C Stavrogiannis, Enrico H Gerding, and Maria Polukarov. Competing intermediary auctions. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 667–674. International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[148] Roberto Battiti. First-and second-order methods for learning: between steepest descent and newton's method. *Neural computation*, 4(2):141–166, 1992.

[149] Jingfang Xu, Chuanliang Chen, Gu Xu, Hang Li, and Elbio Renato Torres Abib. Improving quality of training data for learning to rank using click-through data. In *Proceedings of the third ACM international conference on Web search and data mining*, pages 171–180. ACM, 2010.

[150] Yifan Hu, Yehuda Koren, and Chris Volinsky. Collaborative filtering for implicit feedback datasets. In *Data Mining, 2008. ICDM'08. Eighth IEEE International Conference on*, pages 263–272. Ieee, 2008.

[151] Shankar P Bhattacharyya, H Chapellat, and Lee H Keel. Robust control. *The Parametric Approach, by Prentice Hall PTR*, 1995.

[152] Daniel E Rivera, Manfred Morari, and Sigurd Skogestad. Internal model control: Pid controller design. *Industrial & engineering chemistry process design and development*, 25(1):252–265, 1986.