Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting

Jun Wang
University College London
MediaGamma Ltd
junwang@cs.ucl.ac.uk

Weinan Zhang
Shanghai Jiao Tong University
wnzhang@sjtu.edu.cn

Shuai Yuan
MediaGamma Ltd
shuai.yuan@mediagamma.com

the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

J. Wang, W. Zhang and S. Yuan. *Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting*. Foundations and Trends® in Information Retrieval, vol. 11, no. 4-5, pp. 297–435, 2017.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-68083-310-2

Foundations and Trends® in Information Retrieval

Volume 11, Issue 4-5, 2017

Editorial Scope

Topics

Foundations and Trends® in Information Retrieval publishes survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation, and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians

Foundations and Trends® in Information Retrieval, 2017, Volume 11, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Foundations and Trends® in Information Retrieval
Vol. 11, No. 4-5 (2017) 297–435

DOI: 10.1561/1500000049

Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting

Jun Wang
University College London
MediaGamma Ltd
junwang@cs.ucl.ac.uk

Weinan Zhang
Shanghai Jiao Tong University
wnzhang@sjtu.edu.cn

Shuai Yuan
MediaGamma Ltd
shuai.yuan@mediagamma.com

Contents

Abstract

The most significant progress in recent years in online display advertising is what is known as the Real-Time Bidding (RTB) mechanism to buy and sell ads. RTB essentially facilitates buying an individual ad impression in real time while it is still being generated from a user's visit. RTB not only scales up the buying process by aggregating a large amount of available inventories across publishers but, most importantly, enables direct targeting of individual users. As such, RTB has fundamentally changed the landscape of digital marketing. Scientifically, the demand for automation, integration and optimisation in RTB also brings new research opportunities in information retrieval, data mining, machine learning and other related fields. In this monograph, an overview is given of the fundamental infrastructure, algorithms, and technical solutions of this new frontier of computational advertising. The covered topics include user response prediction, bid landscape forecasting, bidding algorithms, revenue optimisation, statistical arbitrage, dynamic pricing, and ad fraud detection.

J. Wang, W. Zhang and S. Yuan. *Display Advertising with Real-Time Bidding (RTB) and Behavioural Targeting*. Foundations and Trends® in Information Retrieval, vol. 11, no. 4-5, pp. 297–435, 2017.

DOI: 10.1561/1500000049.

1

Introduction

An advertisement is a marketing message intended to encourage potential customers to purchase a product or to subscribe to a service. Advertising is also a way to establish a brand image through the repeated presence of an advertisement (ad) associated with the brand in the media. Television, radio, newspaper, magazines, and billboards are among the major channels that traditionally place ads, however, the advancement of the Internet enables users to seek information online. Using the Internet, users are able to express their information requests, navigate specific websites and perform e-commerce transactions. Major search engines have continued to improve their retrieval services and users' browsing experience by providing relevant results. Since many more businesses and services are transitioning into the online space, the Internet is a natural choice for advertisers to widen their strategy, reaching potential customers among Web users [Yuan et al., 2012].

As a result, online advertising is now one of the fastest advancing areas in the IT industry. In display and mobile advertising, the most significant technical development in recent years is the growth of Real-Time Bidding (RTB), which facilitates a real-time auction for a display opportunity. Real-time means the auction is per impression and the process usually occurs less than 100 milliseconds before the ad is placed. RTB has fundamentally changed the landscape of the digital media market by scaling the buying process across a large number of available inventories among publishers in an automatic fashion. It also encourages *user behaviour targeting*, a significant shift towards buying focused on user data rather than contextual data [Yuan et al., 2013].

Scientifically, the further demand for automation, integration and optimisation in RTB opens new research opportunities in the fields such as Information Retrieval (IR), Data Mining (DM), Machine Learning (ML), and Economics. IR researchers, for example, are facing the challenge of defining the relevancy of underlying audiences given a campaign goal, and consequently, developing techniques to find and filter them out in the real-time bid request data stream [Zhang et al., 2016a, Perlich et al., 2012]. For data miners, a fundamental task is identifying repeated patterns over the large-scale streaming data of bid requests, winning bids and ad impressions [Cui et al., 2011]. For machine learners, an emerging problem is telling a machine to react to a data stream, i.e., learning to bid cleverly on behalf of advertisers and brands to maximise conversions while keeping costs to a minimum [Xu et al., 2016, Kan et al., 2016, Cai et al., 2017].

It is also of great interest to study learning over multi-agent systems and consider the incentives and interactions of each individual learner (bidding agent). For economics researchers, RTB provides a new playground for micro impression-level auctions with various bidding strategies and macro multiple marketplace competitions with different pricing schemes, auction types and floor price settings, etc.

More interestingly, per impression optimisation allows advertisers and agencies to maximise effectiveness based on their own, or the 3rd party, user data across multiple sources. Advertisers buy impressions from multiple publishers to maximise certain Key Performance Indicators (KPIs) such as clicks or conversions, while publishers sell their impressions through multiple advertisers to optimise their revenue [Yuan and Wang, 2012]. This brings the online advertising market a step closer to the financial markets, where marketplace unity is strongly promoted. A common objective, such as optimising clicks or conversions across webpages, advertisers, and users, calls for significant multi-disciplinary research that combines statistical Machine Learning, Data Mining, Information Retrieval, and behavioural targeting with game theory, economics and optimisation.

Despite its rapid growth and huge potential, many aspects of RTB remain unknown to the research community for a variety of reasons. In this monograph, we aim to offer insightful knowledge of real-world systems, to bridge the gaps between industry and academia, and to provide an overview of the fundamental infrastructure, algorithms, and technical and research challenges of the new frontier of computational advertising.

1.1 A short history of online advertising

The first online ad appeared in 1994 when there were only around 30 million people on the Web. The Web version of the Oct. 27, 1994 issue of *HotWired* was the first to run a true banner ad for AT&T.

1.1.1 The birth of sponsored search and contextual advertising

Online advertising has been around for over a decade. The *sponsored search* paradigm was created in 1998 by Bill Gross of Idealab with the founding of `Goto.com`, which became Overture in October 2001, was acquired by Yahoo! in 2003 and is now Yahoo! Search Marketing [Jansen, 2007]. Meanwhile, Google started its own service *AdWords* using *Generalized Second Price Auction* (GSP) in February 2002, adding quality-based bidding in May 2002 [Karp, 2008]. In 2007, Yahoo! Search Marketing followed, added quality-based bidding as well [Dreller, 2010]. It is worth mentioning that Google paid 2.7 million shares to Yahoo! to solve the patent dispute, as reported by The Washington Post [2004], for the technology that matches ads with search results in sponsored search. Web search has now become an integral part of daily life, vastly reducing the difficulty and time once associated with satisfying an information necessity. Sponsored search allows advertisers to buy certain keywords to promote their business when users use such a search engine and greatly contributes to its continuing a free service.

On the other hand, in 1998, *display* advertising began as a concept *contextual advertising* [Anagnostopoulos et al., 2007, Broder et al., 2007]. Oingo, started by Gilad Elbaz and Adam Weissman, developed a proprietary search algorithm based on word meanings and built upon an underlying lexicon called WordNet. Google acquired Oingo in April 2003 and renamed the system *AdSense* [Karp, 2008]. Later, Yahoo! Publish Network, Microsoft adCenter and Advertising.com Sponsored Listings amongst others were created to offer similar services [Kenny and Marshall, 2001]. The contextual advertising platforms evolved to adapt to a richer media environment, including video, audio and mobile networks with geographical information. These platforms allowed publishers to sell blocks of space on their webpages, video clips and applications to make money. Usually such services are called an *advertising network* or a *display network* and are not necessarily run by search engines, as they can consist of huge numbers of individual publishers and advertisers. Sponsored search ads can also be considered a form of contextual ad that matches with simple context: query keywords; but it has been emphasised due to its early development, large market volume and research attention.

1.1.2 The arrival of ad exchange and real-time bidding

Around 2005, new platforms focusing on real-time bidding (RTB) based buying and selling impressions were created. Examples include ADSDAQ, AdECN, DoubleClick Advertising Exchange, adBrite, and Right Media Exchange, which are now known as *ad exchanges*. Unlike traditional ad networks, these ad exchanges aggregate multiple ad networks together to balance the demand and supply in marketplaces and use auctions to sell an ad impression in real time when it is generated by a user visit [Yuan et al., 2013]. Individual publishers and advertising networks can both benefit from participating in such businesses. Publishers sell impressions to advertisers who are interested in associated user profiles and context while advertisers, on the other hand, can contact more publishers for better matching and buy impressions in real-time together with their user data. At the same time, other similar platforms with different functions emerged [Graham, 2010] including (i) *demand side platform* (DSP), which serves advertisers managing their campaigns and submits real-time bidding responses for each bid request to the ad exchange via algorithms, and (ii) *supply side platform* (SSP), created to serve publishers managing website ad inventory. However, real-time bidding (RTB) and multiple ad networks aggregation do not change the nature of such marketplaces (where buying and selling impressions happen), but only make the transactions in real-time via an auction mechanism. For simplicity, we may use the term "ad exchange" in this monograph to better represent the wider platforms where trading happens.

1.2 The major technical challenges and issues

Real-time advertising generates large amounts of data over time. Globally, DSP Fikisu claims to process 32 billion ad impressions daily [Zhang et al., 2017] and DSP Turn reports to handle 2.5 million per second at peak time [Shen et al., 2015]. The New York Stock Exchange, to better envision the scale, trades around 12 billion shares daily.[1] It is fair to say the volume of transactions from display advertising has already surpassed that of the financial market. Perhaps even more importantly, the display advertising industry provides computer scientists and economists a unique opportunity to study and understand the Internet traffic, user behaviour and incentives, and online transactions. Only this industry aggregates nearly all the Web traffic, in the form of ads transactions, across websites and users globally.

With real-time per-impression buying established together with the cookie-based user tracking and syncing (the technical details will be explained in Chapter 2), the RTB ecosystem provides the opportunity and infrastructure to fully unleash the power of user behavioural targeting and personalisation [Zhang et al., 2016a, Wang et al., 2006, Zhao et al., 2013] for that objective. It allows machine driven algorithms to automate and optimise the relevance match between ads and users [Raeder et al., 2012, Zhang et al., 2014a, Zhang and Wang, 2015, Kan et al.,

[1] According to Daily NYSE group volume, `http://goo.gl/2EflkC`, accessed: 2016-02.

2016].

RTB advertising has become a significant battlefield for Data Science research, acting as a test bed and application for many research topics, including user response (e.g. click-through rate, CTR) estimation [Chapelle et al., 2014, Chapelle, 2015, He et al., 2014, Kan et al., 2016], behavioural targeting [Ahmed et al., 2011, Perlich et al., 2012, Zhang et al., 2016a], knowledge extraction [Ahmed et al., 2011, Yan et al., 2009], relevance feedback [Chapelle, 2014], fraud detection [Stone-Gross et al., 2011, Alrwais et al., 2012, Crussell et al., 2014, Stitelman et al., 2013], incentives and economics [Balseiro et al., 2015, Balseiro and Candogan, 2015], and recommender systems and personalisation [Juan et al., 2016, Zhang et al., 2016a].

1.2.1 Towards information general retrieval (IGR)

A fundamental technical goal in online advertising is to automatically deliver the right ads to the right users at the right time with the right price agreed by the advertisers and publishers. As such, RTB based online advertising is strongly correlated with the field of Information Retrieval (IR), which traditionally focuses on building relevance correspondence between information needs and documents [Baeza-Yates et al., 1999]. The IR research typically deals with textual data but has been extended to multimedia data including images, video and audio signals [Smeulders et al., 2000]. It also covers categorical and rating data, including Collaborative Filtering and Recommender Systems [Wang et al., 2008]. In all these cases, the key research question of IR is to study and model the relevance between the queries and documents in the following two distinctive tasks: retrieval and filtering. The retrieval tasks are those in which information needs (queries) are *ad hoc*, while the document collection stays relatively static. By contrast, information filtering tasks are defined when information needs stay static, whereas documents keep entering the system. A rich literature can be found from the probability ranking principle [Robertson, 1977], the RSJ and BM25 model [Jones et al., 2000], language models of IR [Ponte and Croft, 1998], to the latest development of learning to rank [Joachims, 2002, Liu, 2009], results diversity [Wang and Zhu, 2009, Agrawal et al., 2009, Zhu et al., 2009a] and novelty [Clarke et al., 2008], and deep learning of information retrieval [Li and Lu, 2016, Deng et al., 2013].

We, however, argue that IR can broaden its research scope by going beyond the applications of Web search and enterprise search, turning towards *general retrieval* problems derived from many other applications. Essentially, as long as there is concern with building a correspondence between two information objects, under various objectives and criteria [Gorla et al., 2013], we would consider it a general retrieval problem. Online advertising is one of the application domains, and we hope this monograph will shed some light on new information general retrieval (IGR) problems. For instance, the techniques presented on real time advertising are built upon the rich literature of IR, data mining, machine learning and other relevant fields, to answer various questions related to the relevance matching between ads and users. But the difference and difficulty, compared to a typical IR problem, lies in its consideration of various economic constraints. Some of the constraints are related to incentives inherited from the auction mechanism, while others relate to disparate objectives from the participants (advertisers and publishers). In addition, RTB also provides a useful case for relevance matching that is bi-directional and unified between two matched information objects [Robertson et al., 1982, Gorla, 2016, Gorla et al., 2013]. In RTB, there is an inner connection between ads, users and publishers [Yuan et al., 2012]: advertisers would want the matching between the underlying users and their ads to eventually lead to conversions, whereas publishers hope the matching between the ads and their webpage would result in a high ad payoff. Both objectives, among others, require fulfilment when the relevancy is calculated.

1.3 The organisation of this monograph

The targeted audience of this monograph is academic researchers and industry practitioners in the field. The intention is to help the audience acquire domain knowledge and to promote research activities in RTB and computational advertising in general.

The content of the monograph is organised in four folds. Firstly, chapters 2 and 3 provide a general overview of the RTB advertising and its mechanism. Specifically, in Chapter 2, we explain how RTB works, as well as the mainstream user tracking and synchronising techniques that have been popular in the industry; In Chapter 3 we introduce the RTB auction mechanism and the resulting forecasting techniques in the auction market. Next, we cover the problems from the view of advertisers: in Chapter 4, we explain various user response models proposed in the past to target users and make ads more fit to the underlying user's patterns, while in Chapter 5, we present bid optimisation from advertisers' viewpoints with various market settings. After that, in Chapter 6, we focus on the publisher's side and explain dynamic pricing of reserve price, programmatic direct, and new type of advertising contracts. The monograph then concludes with attribution models in Chapter 7 and ad fraud detection in Chapter 8, two additional important subjects in RTB.

There are several read paths depending on reader's technical backgrounds and interests. For academic researchers, chapters 2 and 3 shall help them understand the real-time online advertising systems currently deployed in the industry. The later chapters shall help industry practitioners grasp the research challenges, the state of the art algorithms and potential future systems in this field. As these later chapters are on specialised topics, they can be read in independently at a deeper level.

References

Vibhanshu Abhishek, Peter Fader, and Kartik Hosanagar. Media exposure through the funnel: A model of multi-stage attribution. *Available at SSRN 2158421*, 2012.

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.

Gunes Acar, Christian Eubank, Steven Englehardt, Marc Juarez, Arvind Narayanan, and Claudia Diaz. The web never forgets: Persistent tracking mechanisms in the wild. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 674–689. ACM, 2014.

Deepak Agarwal, Souvik Ghosh, Kai Wei, and Siyu You. Budget pacing for targeted online advertisements at linkedin. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1613–1619. ACM, 2014.

Rakesh Agrawal, Sreenivas Gollapudi, Alan Halverson, and Samuel Ieong. Diversifying search results. In *Proceedings of the Second ACM International Conference on Web Search and Data Mining*, pages 5–14. ACM, 2009.

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.

Amr Ahmed, Abhimanyu Das, and Alexander J Smola. Scalable hierarchical multitask learning algorithms for conversion optimization in display advertising. In *Proceedings of the Seventh ACM International Conference on Web Search and Data Mining*, pages 153–162. ACM, 2014.

Sumayah A. Alrwais, Alexandre Gerber, Christopher W. Dunn, Oliver Spatscheck, Minaxi Gupta, and Eric Osterweil. Dissecting ghost clicks: Ad fraud via misdirected human clicks. In *Proceedings of the 28th Annual Computer Security Applications Conference*, ACSAC '12, pages 21–30, New York, NY, USA, 2012. ACM. ISBN 978-1-4503-1312-4. . URL `http://doi.acm.org/10.1145/2420950.2420954`.

Kareem Amin, Michael Kearns, Peter B. 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.

Aris Anagnostopoulos, Andrei Z. Broder, Evgeniy Gabrilovich, Vanja Josifovski, and Lance Riedel. Just-in-time contextual advertising. In *Proceedings of the Sixteenth ACM Conference on Conference on Information and Knowledge Management*, pages 331–340. ACM, 2007.

Eva Anderl, Ingo Becker, Florian V Wangenheim, and Jan Hendrik Schumann. Mapping the customer journey: A graph-based framework for online attribution modeling. *Available at SSRN 2343077*, 2014.

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*, 2005.

Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.

Ricardo Baeza-Yates, Berthier Ribeiro-Neto, et al. *Modern information retrieval*, volume 463. ACM press New York, 1999.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*, 2014.

Santiago R. Balseiro and Ozan Candogan. Optimal contracts for intermediaries in online advertising. *Available at SSRN 2546609*, 2015.

Santiago R. Balseiro, Omar Besbes, and Gabriel Y. Weintraub. Repeated auctions with budgets in ad exchanges: Approximations and design. *Management Science*, 61(4):864–884, 2015.

Joel Barajas, Ram Akella, Aaron Flores, and Marius Holtan. Estimating ad impact on clicker conversions for causal attribution: A potential outcomes approach. In *Proceedings of the 2015 SIAM International Conference on Data Mining*, pages 640–648. SIAM, 2015.

Joel Barajas Zamora. *Online display advertising causal attribution and evaluation*. PhD thesis, UC Santa Cruz: Electrical Engineering, 2015.

Adam Barth. The web origin concept. Technical report, Internet Engineering Task Force, 2011.

BBC. Google to charge advertisers viewed for seen ads. `http://www.bbc.com/news/business-25356956`, 2013. Accessed: 2016-07.

Yoshua Bengio, Pascal Lamblin, Dan Popovici, Hugo Larochelle, et al. Greedy layer-wise training of deep networks. *Advances in Neural Information Processing Systems*, 19:153, 2007.

James O Berger. *Statistical decision theory and Bayesian analysis*. Springer Science & Business Media, 1985.

Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973. ISSN 00223808.

Károly Boda, Ádám Máté Földes, Gábor György Gulyás, and Sándor Imre. User tracking on the web via cross-browser fingerprinting. In *Nordic Conference on Secure IT Systems*, pages 31–46. Springer, 2011.

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.

Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.

Leo Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

Leo Breiman, Jerome H. Friedman, Charles J Stone, and Richard A Olshen. *Classification and regression trees*. CRC press, 1984.

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.

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.

Han Cai, Kan Ren, Weinan Zhang, Kleanthis Malialis, and Jun Wang. Real-time bidding by reinforcement learning in display advertising. In *The Tenth ACM International Conference on Web Search and Data Mining*. ACM, 2017.

Nicolo Cesa-Bianchi, Claudio Gentile, and Yishay Mansour. Regret minimization for reserve prices in second-price auctions. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1190–1204. Society for Industrial and Applied Mathematics, 2013.

Tanmoy Chakraborty, Eyal Even-Dar, Sudipto Guha, Yishay Mansour, and S. Muthukrishnan. Selective call out and real time bidding. In *International Workshop on Internet and Network Economics*, pages 145–157. Springer, 2010.

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.

Olivier Chapelle. Offline evaluation of response prediction in online advertising auctions. In *Proceedings of the 24th International Conference on World Wide Web*, pages 919–922. ACM, 2015.

Olivier Chapelle, Eren Manavoglu, and Romer Rosales. Simple and scalable response prediction for display advertising. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 5(4):61, 2014.

Bowei Chen and Jun Wang. A lattice framework for pricing display advertisement options with the stochastic volatility underlying model. *Electronic Commerce Research and Applications*, 14(6):465–479, 2015.

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 1:1–1:9. ACM, 2014.

Bowei Chen, Jun Wang, Ingemar J. Cox, and Mohan S. Kankanhalli. Multi-keyword multi-click advertisement option contracts for sponsored search. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 7(1):5, 2015.

Tianqi Chen, Zhao Zheng, Qiuxia Lu, Weinan Zhang, and Yong Yu. Feature-based matrix factorization. *arXiv preprint arXiv:1109.2271*, 2011a.

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, 2011b.

Charles L. A. Clarke, Maheedhar Kolla, Gordon V. Cormack, Olga Vechtomova, Azin Ashkan, Stefan Büttcher, and Ian MacKinnon. Novelty and diversity in information retrieval evaluation. In *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 659–666. ACM, 2008.

Ronan Collobert and Jason Weston. A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th International Conference on Machine Learning*, pages 160–167. ACM, 2008.

Jonathan Crussell, Ryan Stevens, and Hao Chen. Madfraud: Investigating ad fraud in android applications. In *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, pages 123–134. ACM, 2014.

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.

Dorota M. Dabrowska. Non-parametric regression with censored survival time data. *Scandinavian Journal of Statistics*, pages 181–197, 1987.

Wenyuan Dai, Gui-Rong Xue, Qiang Yang, and Yong Yu. Transferring naive bayes classifiers for text classification. In *AAAI*, volume 7, pages 540–545, 2007.

Brian Dalessandro, Claudia Perlich, Ori Stitelman, and Foster Provost. Causally motivated attribution for online advertising. In *Proceedings of the Sixth International Workshop on Data Mining for Online Advertising and Internet Economy*, page 7. ACM, 2012.

Brian Dalessandro, Daizhuo Chen, Troy Raeder, Claudia Perlich, Melinda Han Williams, and Foster Provost. Scalable hands-free transfer learning for online advertising. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1573–1582. ACM, 2014.

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.

Neil Daswani, Chris Mysen, Vinay Rao, Stephen Weis, Kourosh Gharachorloo, and Shuman Ghosemajumder. Online advertising fraud. *Crimeware: Understanding New Attacks and Defenses*, 40(2):1–28, 2008.

Li Deng, Xiaodong He, and Jianfeng Gao. Deep stacking networks for information retrieval. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3153–3157. IEEE, 2013.

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.

Josh Dreller. A brief history of paid search advertising. `https://goo.gl/Rls5GQ` (Accessed: 13-Jun-2017), 2010.

Michael Durbin. *All about high-frequency trading.* McGraw Hill Professional, 2010.

Peter Eckersley. How unique is your web browser? In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 1–18. Springer, 2010.

Benjamin Edelman and Michael Ostrovsky. Strategic bidder behavior in sponsored search auctions. *Decision support systems*, 43(1):192–198, 2007.

Benjamin Edelman and Michael Schwarz. Optimal auction design in a multi-unit environment: The case of sponsored search auctions. *Unpublished manuscript, Harvard Business School*, 2006. URL `http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Optimal+Auction+Design+in+a+Multi-unit+Environment+:+The+Case+of+Sponsored+Search+Auctions#0`.

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.

Eyal Even Dar, Jon Feldman, Yishay Mansour, and S. Muthukrishnan. Position auctions with bidder-specific minimum prices. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5385 LNCS:577–584, 2008. ISSN 03029743. .

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.

Maryam Feily, Alireza Shahrestani, and Sureswaran Ramadass. A survey of botnet and botnet detection. In *2009 Third International Conference on Emerging Security Information, Systems and Technologies*, pages 268–273. IEEE, 2009.

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.

Jerome H. Friedman. Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, pages 1189–1232, 2001.

Jerome H. Friedman. Stochastic gradient boosting. *Computational Statistics & Data Analysis*, 38(4):367–378, 2002.

Jerome H. Friedman, Trevor Hastie, and Robert Tibshirani. *The elements of statistical learning*, volume 1. Springer series in statistics Springer, Berlin, 2001.

Gian M. Fulgoni. Fraud in digital advertising: A multibillion-dollar black hole. *Journal of Advertising Research*, pages JAR–2016, 2016.

Sahin Cem Geyik, Abhishek Saxena, and Ali Dasdan. Multi-touch attribution based budget allocation in online advertising. In *Proceedings of 20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pages 1–9. ACM, 2014.

Arpita Ghosh, Benjamin I.P. 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.

Richard Gomer, Eduarda Mendes Rodrigues, Natasa Milic-Frayling, and M. C. Schraefel. Network analysis of third party tracking: User exposure to tracking cookies through search. In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM International Joint Conferences on*, volume 1, pages 549–556. IEEE, 2013.

Google. Ad traffic quality resource center. `http://www.google.com/ads/adtrafficquality/index.html`, 2016. Accessed: 13-Jun-2017.

Jagadeesh Gorla. *A bi-directional unified Model for information retrieval.* PhD thesis, University College London, 2016.

Jagadeesh Gorla, Neal Lathia, Stephen E. Robertson, and Jun Wang. Probabilistic group recommendation via information matching. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 495–504. ACM, 2013.

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*, pages 13–20, 2010.

Rob Graham. A brief history of digital ad buying and selling. `https://goo.gl/BZbuWY` (Accessed: 13-Jun-2017), 2010.

William H. Greene. Censored data and truncated distributions. *Available at SSRN 825845*, 2005.

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.

Ramakrishna Gummadi, Peter B. Key, and Alexandre Proutiere. Repeated auctions under budget constraints: Optimal bidding strategies and equilibria. In *Eighth Workshop on Ad Auctions*, 2012.

Hamed Haddadi. Fighting online click-fraud using bluff ads. *ACM SIGCOMM Computer Communication Review*, 40(2):21–25, 2010.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1026–1034, 2015.

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.

Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N. Sainath, and Brian Kingsbury. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

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.

Interactive Advertising Bureau. What is an untrustworthy supply chain costing the us digital advertising industry?, 2015.

Bernard J. Jansen. Sponsored Search: Is Money a Motivator for Providing Relevant Results. *IEEE Computer*, 2007.

Thorsten Joachims. Optimizing search engines using clickthrough data. In *Proceedings of the eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 133–142. ACM, 2002.

Norman L. Johnson. *Survival models and data analysis*, volume 74. John Wiley & Sons, 1999.

K. Sparck Jones, Steve Walker, and Stephen E. Robertson. A probabilistic model of information retrieval: development and comparative experiments: Part 2. *Information Processing & Management*, 36(6):809–840, 2000.

Yuchin Juan, Yong Zhuang, Wei-Sheng Chin, and Chih-Jen Lin. Field-aware factorization machines for CTR prediction. In *Proceedings of the 9th ACM Conference on Recommender Systems*, 2016.

Ren Kan, Weinan Zhang, Yifei Rong, Haifeng Zhang, Yong Yu, and Jun Wang. User response learning for directly optimizing campaign performance in display advertising. In *Proceedings of the ACM International Conference on Information & Knowledge Management*. ACM, 2016.

Alex Kantrowitz. Inside google's secret war against ad fraud. *Ad Age*, May 2015.

Edward L. Kaplan and Paul Meier. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282): 457–481, 1958.

Niklas Karlsson and Jianlong Zhang. Applications of feedback control in online advertising. In *American Control Conference (ACC), 2013*, pages 6008–6013. IEEE, 2013.

Scott Karp. Google adwords: A brief history of online advertising innovation. `https://goo.gl/bh1HVW` (Accessed: 13-Jun-2017), 2008.

Bill Kee. Attribution playbook - google analytics. `http://services.google.com/fh/files/misc/attribution_playbook.pdf`, 2012. Accessed: 13-Jun-2017.

David Kenny and John F. Marshall. Contextual marketing: The real business of the internet. `http://hbswk.hbs.edu/archive/2124.html` (Accessed: 13-Jun-2017), 2001.

Brendan Kitts and Benjamin Leblanc. Optimal bidding on keyword auctions. *Electronic Markets*, 14(3):186–201, 2004.

Robert Kleinberg and Tom Leighton. The value of knowing a demand curve: Bounds on regret for online posted-price auctions. In *Proceedings of 44th Annual IEEE Symposium on Foundations of Computer Science.*, pages 594–605. IEEE, 2003.

Yehuda Koren, Robert Bell, Chris Volinsky, et al. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.

Kevin J. Lang, Benjamin Moseley, and Sergei Vassilvitskii. Handling forecast errors while bidding for display advertising. In *Proceedings of the 21st International Conference on World Wide Web*, pages 371–380. ACM, 2012.

Yann LeCun and Yoshua Bengio. Convolutional networks for images, speech, and time series. *The Handbook of Brain Theory and Neural Networks*, 3361 (10):1995, 1995.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

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.

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.

Bin Li, Qiang Yang, and Xiangyang Xue. Transfer learning for collaborative filtering via a rating-matrix generative model. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 617–624. ACM, 2009.

Hang Li and Zhengdong Lu. Deep learning for information retrieval. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1203–1206. ACM, 2016.

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.

Xuejun Liao, Ya Xue, and Lawrence Carin. Logistic regression with an auxiliary data source. In *Proceedings of the 22nd International Conference on Machine Learning*, pages 505–512. ACM, 2005.

Tie-Yan Liu. Learning to rank for information retrieval. *Foundations and Trends in Information Retrieval*, 3(3):225–331, 2009.

Babak Loni, Yue Shi, Martha Larson, and Alan Hanjalic. Cross-domain collaborative filtering with factorization machines. In *European Conference on Information Retrieval*, pages 656–661. Springer, 2014.

Ashish Mangalampalli, Adwait Ratnaparkhi, Andrew O. Hatch, Abraham Bagherjeiran, Rajesh Parekh, and Vikram Pudi. A feature-pair-based associative classification approach to look-alike modeling for conversion-oriented user-targeting in tail campaigns. In *Proceedings of the 20th international conference companion on World wide web*, pages 85–86. ACM, 2011.

Steven A. Matthews. *A technical primer on auction theory I: Independent private values*, volume 1096. Center for Mathematical Studies in Economics and Management Science, Northwestern University, 1995.

R. Preston McAfee. The design of advertising exchanges. *Review of Industrial Organization*, 39(3):169–185, 2011.

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.

Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5):22, 2007a.

Aranyak Mehta, Amin Saberi, Umesh Vazirani, and Vijay Vazirani. Adwords and generalized online matching. *Journal of the ACM (JACM)*, 54(5):22, 2007b.

Flavio M. Menezes and Paulo Klinger Monteiro. *An introduction to auction theory*. OUP Oxford, 2005.

Aditya Krishna Menon, Krishna-Prasad Chitrapura, Sachin Garg, Deepak Agarwal, and Nagaraj Kota. Response prediction using collaborative filtering with hierarchies and side-information. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 141–149. ACM, 2011.

Paul Robert Milgrom. *Putting auction theory to work*. Cambridge University Press, 2004.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

S. Muthukrishnan. Ad exchanges: Research issues. In *International Workshop on Internet and Network Economics*, pages 1–12. Springer, 2009.

S. Muthukrishnan, Martin Pál, and Zoya Svitkina. Stochastic models for budget optimization in search-based advertising. *Algorithmica*, 58(4):1022–1044, 2010.

R. B. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981. ISSN 0364-765X. .

Moni Naor and Benny Pinkas. Secure accounting and auditing on the web. *Computer Networks and ISDN Systems*, 30(1):541–550, 1998.

Richard J. Oentaryo, Ee-Peng Lim, Michael Finegold, David Lo, Feida Zhu, Clifton Phua, Eng-Yeow Cheu, Ghim-Eng Yap, Kelvin Sim, Minh Nhut Nguyen, et al. Detecting click fraud in online advertising: a data mining approach. *Journal of Machine Learning Research*, 15(1):99–140, 2014a.

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, 2014b.

Martin J. Osborne and Ariel Rubinstein. *A course in game theory*. MIT press, 1994.

Michael Ostrovsky and Michael Schwarz. Reserve prices in internet advertising auctions: A field experiment. *Search*, pages 1–18, 2009. ISSN 15565068. .

Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.

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.

Jay M. Ponte and W Bruce Croft. A language modeling approach to information retrieval. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 275–281. ACM, 1998.

Yanru Qu, Han Cai, Kan Ren, Weinan Zhang, Yong Yu, Ying Wen, and Jun Wang. Product-based neural networks for user response prediction. In *IEEE 16th International Conference on Data Mining*, 2016.

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.

Troy Raeder, Ori Stitelman, Brian Dalessandro, Claudia Perlich, and Foster Provost. Design principles of massive, robust prediction systems. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1357–1365. ACM, 2012.

Kan Ren, Weinan Zhang, Yifei Rong, Haifeng Zhang, Yong Yu, and Jun Wang. User response learning for directly optimizing campaign performance in display advertising. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*, pages 679–688. ACM, 2016.

Steffen Rendle. Factorization machines. In *IEEE 10th International Conference on Data Mining*, pages 995–1000. IEEE, 2010.

Steffen Rendle. Factorization machines with libfm. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(3):57, 2012.

Steffen Rendle and Lars Schmidt-Thieme. Pairwise interaction tensor factorization for personalized tag recommendation. In *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, pages 81–90. ACM, 2010.

Steffen Rendle, Zeno Gantner, Christoph Freudenthaler, and Lars Schmidt-Thieme. Fast context-aware recommendations with factorization machines. In *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 635–644. ACM, 2011.

Stephen E. Robertson. The probability ranking principle in ir. *Journal of Documentation*, 33(4):294–304, 1977.

Stephen E. Robertson, ME Maron, and William S Cooper. The unified probabilistic model for ir. In *International Conference on Research and Development in Information Retrieval*, pages 108–117. Springer, 1982.

Paul Samuelson. Rational theory of warrant pricing. *Industrial Management Review*, 6:13–31, 1965.

Xuhui Shao and Lexin Li. Data-driven multi-touch attribution models. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 258–264. ACM, 2011.

Lloyd S. Shapley. A value for n-person games. Technical report, DTIC Document, 1952.

Jianqiang Shen, Burkay Orten, Sahin Cem Geyik, Daniel Liu, Shahriar Shariat, Fang Bian, and Ali Dasdan. From 0.5 million to 2.5 million: Efficiently scaling up real-time bidding. In *IEEE International Conference on Data Mining*, pages 973–978. IEEE, 2015.

Enno Shioji and Masayuki Arai. Neural feature embedding for user response prediction in real-time bidding (rtb). *arXiv preprint arXiv:1702.00855*, 2017.

Ritwik Sinha, Shiv Saini, and N. Anadhavelu. Estimating the incremental effects of interactions for marketing attribution. In *International Conference on Behavior, Economic and Social Computing*, pages 1–6. IEEE, 2014.

Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.

Ashkan Soltani, Shannon Canty, Quentin Mayo, Lauren Thomas, and Chris Jay Hoofnagle. Flash cookies and privacy. In *AAAI Spring Symposium: Intelligent Information Privacy Management*, volume 2010, pages 158–163, 2010.

Kevin Springborn and Paul Barford. Impression fraud in on-line advertising via pay-per-view networks. In *USENIX Security*, pages 211–226, 2013.

Ori Stitelman, Claudia Perlich, Brian Dalessandro, Rod Hook, Troy Raeder, and Foster Provost. Using co-visitation networks for detecting large scale online display advertising exchange fraud. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1240–1248. ACM, 2013.

Brett Stone-Gross, Ryan Stevens, Apostolis Zarras, Richard Kemmerer, Chris Kruegel, and Giovanni Vigna. Understanding fraudulent activities in online ad exchanges. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, pages 279–294. ACM, 2011.

Anh-Phuong Ta. Factorization machines with follow-the-regularized-leader for ctr prediction in display advertising. In *2015 IEEE International Conference on Big Data*, pages 2889–2891. IEEE, 2015.

Kalyan T. Talluri and Garrett J. van Ryzin. *The Theory and Practice of Revenue Management*. Springer, 2005.

Matthew E. Taylor and Peter Stone. Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10: 1633–1685, December 2009. ISSN 1532-4435.

The Washington Post. Google ends its dispute with yahoo. http://www.washingtonpost.com/wp-dyn/articles/A52880-2004Aug9.html (Accessed: 13-Jun-2017), 2004.

Kurt Thomas, Elie Bursztein, Chris Grier, Grant Ho, Nav Jagpal, Alexandros Kapravelos, Damon McCoy, Antonio Nappa, Vern Paxson, Paul Pearce, Niels Provos, and Moheeb Abu Rajab. Ad injection at scale: Assessing deceptive advertisement modifications. In *2015 IEEE Symposium on Security and Privacy*, pages 151–167. IEEE, 2015.

Stephen Tyree, Kilian Q. Weinberger, Kunal Agrawal, and Jennifer Paykin. Parallel boosted regression trees for web search ranking. In *Proceedings of the 20th International Conference on World Wide Web*, pages 387–396. ACM, 2011.

Nevena Vratonjic, Mohammad Hossein Manshaei, Maxim Raya, and Jean-Pierre Hubaux. Isps and ad networks against botnet ad fraud. In *International Conference on Decision and Game Theory for Security*, pages 149–167. Springer, 2010.

Jun Wang and Bowei Chen. Selling futures online advertising slots via option contracts. In *Proceedings of the 21st International Conference on World Wide Web*, pages 627–628. ACM, 2012.

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.

Jun Wang, Arjen P. De Vries, and Marcel J. T. Reinders. Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 501–508. ACM, 2006.

Jun Wang, Arjen P. De Vries, and Marcel J. T. Reinders. Unified relevance models for rating prediction in collaborative filtering. *ACM Transactions on Information Systems (TOIS)*, 26(3):16, 2008.

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.

David A. Wooff and Jillian M. Anderson. Time-weighted multi-touch attribution and channel relevance in the customer journey to online purchase. *Journal of Statistical Theory and Practice*, 9(2):227–249, 2015.

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.

Baichun Xiao, W. Yang, and J. Li. Optimal reserve price for the generalized second-price auction in sponsored search advertising. *Journal of Electronic Commerce Research*, 10(3):114–129, 2009.

Jian Xu, Kuang-chih Lee, Wentong Li, Hang Qi, and Quan Lu. Smart pacing for effective online ad campaign optimization. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2217–2226. ACM, 2015.

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.

Lizhen Xu, Jason A Duan, and Andrew Whinston. Path to purchase: A mutually exciting point process model for online advertising and conversion. *Management Science*, 60(6):1392–1412, 2014.

Jun Yan, Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen. How much can behavioral targeting help online advertising? In *Proceedings of the 18th International Conference on World Wide Web*, pages 261–270. ACM, 2009.

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.

Shuai Yuan, Ahmad Zainal Abidin, Marc Sloan, and Jun Wang. Internet advertising: An interplay among advertisers, online publishers, ad exchanges and web users. *arXiv preprint arXiv:1206.1754*, 2012.

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.

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.

Haifeng Zhang, Weinan Zhang, Yifei Rong, Kan Ren, Wenxin Li, and Jun Wang. Managing risk of bidding in display advertising. In *The Tenth ACM International Conference on Web Search and Data Mining*. ACM, 2017.

Weinan Zhang. *Optimal Real-Time Bidding for Display Advertising*. PhD thesis, University College London, 2016.

Weinan Zhang and Jun Wang. Statistical arbitrage mining for display advertising. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1465–1474. ACM, 2015.

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.

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, 2014a.

Weinan Zhang, Shuai Yuan, Jun Wang, and Xuehua Shen. Real-time bidding benchmarking with ipinyou dataset. *arXiv preprint arXiv:1407.7073*, 2014b.

Weinan Zhang, Ye Pan, Tianxiong Zhou, and Jun Wang. An empirical study on display ad impression viewability measurements. *arXiv preprint arXiv:1505.05788*, 2015.

Weinan Zhang, Lingxi Chen, and Jun Wang. Implicit look-alike modelling in display ads: Transfer collaborative filtering to ctr estimation. In *European Conference on Information Retrieval*, pages 589–601. Springer, 2016a.

Weinan Zhang, Tianming Du, and Jun Wang. Deep learning over multi-field categorical data. In *European Conference on Information Retrieval*, pages 45–57. Springer, 2016b.

Weinan Zhang, Kan Ren, and Jun Wang. Optimal real-time bidding frameworks discussion. *arXiv preprint arXiv:1602.01007*, 2016c.

Weinan Zhang, Yifei Rong, Jun Wang, Tianchi Zhu, and Xiaofan Wang. Feedback control of real-time display advertising. In *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*, pages 407–416. ACM, 2016d.

Weinan Zhang, Tianxiong Zhou, Jun Wang, and Jian Xu. Bid-aware gradient descent for unbiased learning with censored data in display advertising. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM, 2016e.

Ya Zhang, Yi Wei, and Jianbiao Ren. Multi-touch attribution in online advertising with survival theory. In *IEEE International Conference on Data Mining*, pages 687–696. IEEE, 2014c.

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, 2014d.

Yuyu Zhang, Hanjun Dai, Chang Xu, Jun Feng, Taifeng Wang, Jiang Bian, Bin Wang, and Tie-Yan Liu. Sequential click prediction for sponsored search with recurrent neural networks. In *Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014e.

Xiaoxue Zhao, Weinan Zhang, and Jun Wang. Interactive collaborative filtering. In *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management*, pages 1411–1420. ACM, 2013.

Yunhong Zhou, Deeparnab Chakrabarty, and Rajan Lukose. Budget constrained bidding in keyword auctions and online knapsack problems. In *International Workshop on Internet and Network Economics*, pages 566–576. Springer, 2008.

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, 2009a.

Mu Zhu. Recall, precision and average precision. *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo*, 2, 2004.

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, 2009b.