Auto-Bidding in Large-Scale Auctions: Learning Decision-Making in Uncertain and Competitive Games

Jian Xu Zhilin Zhang Zongqing Lu Xiaotie Deng Michael P. Wellman
Chuan Yu Shuai Dou Yusen Huo Zhiwei Xu Zhijian Duan
Shaopan Xiong Chuang Liu Ningyuan Li Kefan Su Wei Gong
Bo Zheng `xiyu.xj@alibaba-inc.com`

Abstract

Decision-making in large-scale games is an essential research area in artificial intelligence with significant real-world impact. An agent confronts the critical task of making high-frequency strategic decisions in an uncertain and competitive environment, characterized by significant randomness and rapidly changing strategies from massive competitors. However, the shortage of large-scale, realistic game systems and datasets has hindered research progress in this area. To provide opportunities for in-depth research on this highly valuable problem, we present the Auto-Bidding in Large-Scale Auctions challenge derived from online advertising, a booming $626.8 billion industry in 2023. We have developed a standardized ad auction system for the competition, which reproduces the characteristics of real-world large-scale games and incorporates essential features that deserve research attention. We also provide a training framework with a 500-million-record dataset and several industry-proven methods as baselines to help participants quickly start and deeply optimize their strategies. Furthermore, we have prepared a comprehensive promotional strategy, raised sufficient funds, and offered varied incentives to attract more participants from diverse backgrounds. We believe that the proposed competition will provide opportunities for more researchers to gain insights and conduct research in this field, driving technical innovation for both research and real-world practical applications.

Keywords

Decision-Making, Competitive Game, Large-Scale Auction, Online Advertising, Auto-Bidding

1 Competition Description

1.1 Background and Impact

Decision-making is a key research area in artificial intelligence with wide real-world applications, especially in large-scale games such as online advertising [12], financial market [5], e-commerce [13], and energy trading [11]. In large-scale games, an agent needs to make high-frequency strategic decisions to achieve its objective under resource constraints in an uncertain and competitive environment, characterized by significant randomness and rapidly changing strategies of massive competitors. However, the shortage of **large-scale**, **realistic** game systems and datasets for research purposes, caused by restricted data access within the industry, poses a challenge for academic researchers to study this problem and verify their strategies, hindering technological advancement in the field. Therefore, we believe that establishing a platform to provide opportunities for in-depth research on decision-making within uncertain, changing, and competitive games could capture the

Preprint. Under review.

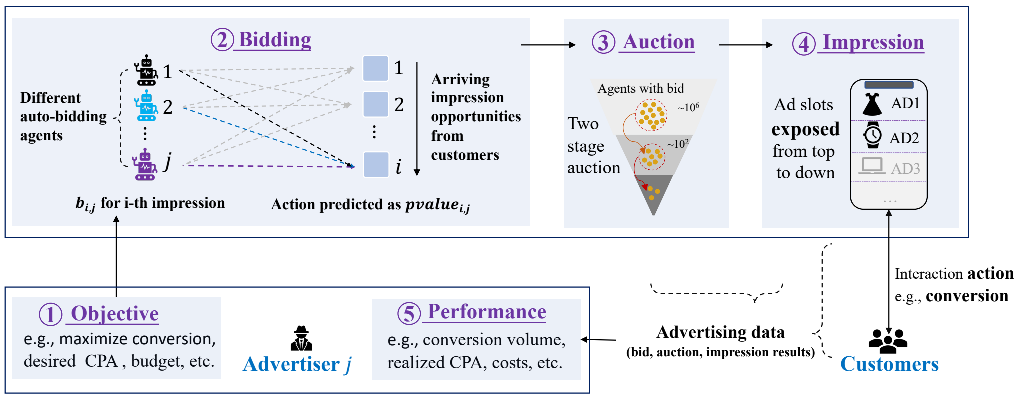

Figure 1: Overview of typical large-scale online advertising platform. Numbers 1 through 5 illustrate how an auto-bidding agent helps advertiser j optimize performance.

interest of researchers from the NeurIPS and broader artificial intelligence community, while also advancing the development of the related fields.

Therefore, we introduce the Auto-Bidding in Large-Scale Auctions challenge, a typical large-scale game-based competition derived from online advertising, which reached a market size of $626.8 billion in 2023 [2]. As shown in Figure 1, The auto-bidding agent sequentially bids for each impression across a large number of continuously arriving opportunities to maximize performance, adhering to certain constraints on behalf of the advertiser. The key challenge is how to make effective bidding decisions in large-scale auctions, which requires effective perceiving the **changes** in competitors' strategies and the **randomness** of opportunities' arrivals. In our competition, each participant is tasked with implementing an auto-bidding agent that considers the given optimization objective and constraints. They need to compete against massive agents that employ confidential strategies provided by other participants or organizers. Besides dealing with the complexity of large-scale auctions, participants also need to address essential challenges deserving of research attention in real-world scenarios.

This competition has received strong support from Alibaba, which owns the leading online advertising platform in China, and is equipped with significant resources. For the competition platform, we have developed a standardized ad auction system that allows numerous participants to evaluate their strategies' performance in large-scale games. This system effectively reproduces dynamic competition among agents, impression opportunity arrival patterns, and several essential industry characteristics. Additionally, we have established the competition website on Tianchi, which has successfully hosted over 500 competitions since 2014. The competition platform's robustness, completeness, and scalability have been effectively validated by a confidential in-house contest within our group, involving roughly 200 participants[1]. For the materials, we have prepared a large-scale advertising auction dataset consisting of approximately **500** million records and a total size of **40GB**. This dataset has undergone strict privacy protection and has passed rigorous review and approval. In addition, we have developed a training framework and provided several industry-proven methods as baselines to help participants quickly get started. For the promotion, we have provided comprehensive promotional strategies with sufficient funds and various incentives to attract participants from diverse backgrounds.

Since decision-making in large-scale games is a compelling research area in artificial intelligence with substantial influence on applications, our competition is expected to attract a wide range of participants from related fields, including researchers and practitioners in machine learning, game theory, and data science. Researchers working on large language models (LLMs) and foundation agents will also be attracted, as these technologies are playing an increasingly active role in decision-making fields [3, 9]. With our comprehensive promotional and incentive strategies, we conservatively anticipate hundreds of teams from both industry and academia to participate in our competition. We believe that our proposed competition will offer opportunities for researchers to gain insights and conduct in-depth research in large-scale decision-making, thereby driving technical innovation in both academic research and real-world applications. After the competition, we plan to **open-source** the related frameworks and datasets, which will further promote a broad impact in various fields.

[1]These participants' tasks differ from the proposal and they will be excluded from the official competition.

1.2 Novelty

The competition we propose is a brand-new event designed as a novel benchmark for large-scale decision-making, rooted in real-world practical applications. It offers a large-scale, realistic auction system and dataset, and features a distinct task with specific evaluation metrics. This initiative bridges an existing gap in the community for effectively benchmarking research questions in this field. Ultimately, it aims to facilitate in-depth research in the field and spur the development of numerous valuable new research topics.

To our best efforts, we identified two past competitions [7, 10] related to the optimization of online advertising, held more than a decade ago. These competitions were narrowly focused on distinct advertising application issues and conducted in small-scale simulated environments with fewer than 10 participants. In contrast, we distill fundamental research questions from the area of online advertising, particularly those concerning decision-making in large-scale auctions. To this end, we have developed a scalable competition platform designed to accommodate thousands of participants.

Moreover, our competition focuses on key aspects of real-world, large-scale games that are crucial for researchers, highlighting factors like environmental uncertainty and the dynamic strategies of numerous competitors. Compared to other related topics featured at NeurIPS competitions, such as reinforcement learning, our competition digs deeper into the complexities of decision-making within realistic application settings, making it a particularly meaningful and valuable avenue for research.

1.3 Data

In our competition, participant solutions will be evaluated in an auction system, as described in Section 1.5, which is designed to replicate a large-scale auction environment. This involves enriching the system with millions of impression opportunities and deploying hundreds of competing agents. To accurately reflect market dynamics, opportunity arrival patterns are derived from industrial scenarios, showcasing the inherent randomness observed in customer behaviors. These competing agents serve as various advertisers with different objectives, employing auto-bidding strategies of varying effectiveness. The data for evaluation and competing agents' strategies will be kept confidential from participants during the competition. Besides this, we will release a large-scale training dataset with 500 million records to participants at the start of the competition as described in Section 1.6.

All information concerning advertisers and customers has been anonymized and subjected to stringent processing. Data for evaluation and the training dataset have been fully prepared and have passed the rigorous approval process of Alibaba's legal department.

1.4 Tasks and Application Scenarios

Participants are tasked with designing and implementing an auto-bidding agent that is capable of making efficient bidding decisions in a large-scale auction while fulfilling advertisers' specified objectives to optimize performance.

Specifically, the task in our competition is based on a significant and practical problem encountered in real-world scenarios called Target CPA (Cost Per Action) [1]. The problem has been slightly optimized for academic inquiry. Achieving favorable results requires additional efforts to address challenges such as multiple ad slots and variance in conversion prediction, beyond those posed by complex large-scale auctions. This issue remains unsolved in both the academic and professional realms, making it a key research field.

Formally, given an advertiser j's budget B and the desired CPA C, the agent bids for N impression opportunities within an advertising period. The objective is to maximize the total conversion volume while ensuring that the realized CPA remains below C at the end of the period. All opportunities arrive sequentially, and the agent bids on each opportunity in turn. Here are the bidding and auction procedures for each opportunity i:

1. **Bidding**. **The agent bids** b_i based on $pvalue_i$[2], which represents the probability of a conversion action occurring when advertiser j's ad is exposed to the customer. Meanwhile,

[2] For simplicity, we usually omit the index j of the advertiser where it does not cause ambiguity.

other competing agents simultaneously bid using their individual strategies, denoted by b_i^-, to compete for K ad slots in opportunity i.

2. **Auction**. The ad platform runs a GSP (Generalized Second Price) auction mechanism and returns the auction results to the agents. $x_i(b_i, b_i^-)$ indicates whether to win, $k_i(b_i, b_i^-)$ denotes the ad slot won, and $c_i(b_i, b_i^-)$ represents the cost. x_i, k_i, c_i[3] depends not only on b_i but also on b_i^- .

3. **Impression**. Whether the k_i-th slot is exposed to the customer is determined by a random variable defined as $E_i \sim \text{Bernoulli}\big(exposure_{k_i}\big)$, where $exposure_{k_i}$ is the probability of exposure for the slot k_i. The actual conversion occurrence is also a random variable defined as $V_i \sim \text{Bernoulli}\big(value_i\big)$, where $value_i \sim \mathcal{N}(pvalue_i, \sigma_i^2)$, and σ_i represents the variance, e.g., the uncertainty in prediction. If an ad in the slot is not exposed, the advertiser doesn't need to pay the cost and the customer won't make a conversion on the ad.

Therefore, this task can be formalized as follows:

$$\begin{aligned}
\max_{b_1,\cdots,b_N} \quad & \sum_{i=1}^{N} x_i(b_i, b_i^-) \cdot E_i \cdot V_i \\
\text{s.t.} \quad & \sum_{i=1}^{N} x_i(b_i, b_i^-) \cdot E_i \cdot c_i(b_i, b_i^-) \leq B, \\
& \frac{\sum_{i=1}^{N} x_i(b_i, b_i^-) \cdot E_i \cdot c_i(b_i, b_i^-)}{\sum_{i=1}^{N} x_i(b_i, b_i^-) \cdot E_i \cdot V_i} \leq C.
\end{aligned}$$

The realized CPA of the advertiser j is $CPA = \frac{\sum_i x_i(b_i,b_i^-)\cdot E_i \cdot c_i(b_i,b_i^-)}{\sum_i x_i(b_i,b_i^-)\cdot E_i \cdot V_i}$.

challenge. This task is not a traditional constrained optimization problem because when bidding for each opportunity, it is impossible to know all future opportunities in advance due to the randomness of arriving patterns, making it hard to obtain a closed-form analytical solution. Advertising is a competitive game where multiple agents bid simultaneously using their confidential bidding strategies. The game is also dynamic, with competitors continually adapting their strategies. Therefore, auto-bidding agents need to perceive the dynamic gaming environment, model the connection between bid and performance, and then carefully bid for each impression opportunity by taking into account the preceding bids and performance. Additionally, participants need to consider how to adhere to the CPA constraint, given complex features such as the uncertainty of prediction and sparse data. In this competition, we'll have 2 tracks. The first is the classic track, where any optimization method is allowed. The second is the AIGB track, encouraging the use of generative models for modeling.

1.5 Standardized Large-Scale Auction System

To ensure the quality of the competition and to comprehensively demonstrate real-world large-scale games, we have developed a standardized auction system specifically for the competition, as shown in Figure 2. This system effectively reproduces dynamic competition among agents, randomness of impression opportunity arrivals, and several essential industry features deserving of research attention. To better simulate a large-scale auction, a huge amount of impression opportunities will be fed into the system and configured with hundreds of competing agents. These agents are equipped with different confidential bidding strategies developed by the organizers or submitted by participants.

To simplify the auto-bidding process for participants, we divide impression opportunities in an advertising period into T decision time steps. Given the objective, the auto-bidding agent sequentially bids at each step, using the results from step t and prior historical information to refine its strategy for step $t+1$. This approach enables agents to continuously optimize their bidding strategies in order to adapt to the changing environment. Within each step, all impression opportunities are executed independently and in parallel. At the end of the period, the system provides the final performance for the agent.

[3] For simplicity, $x_i(b_i, b_i^-)$ is generally denoted as x_i, so as k_i, c_i .

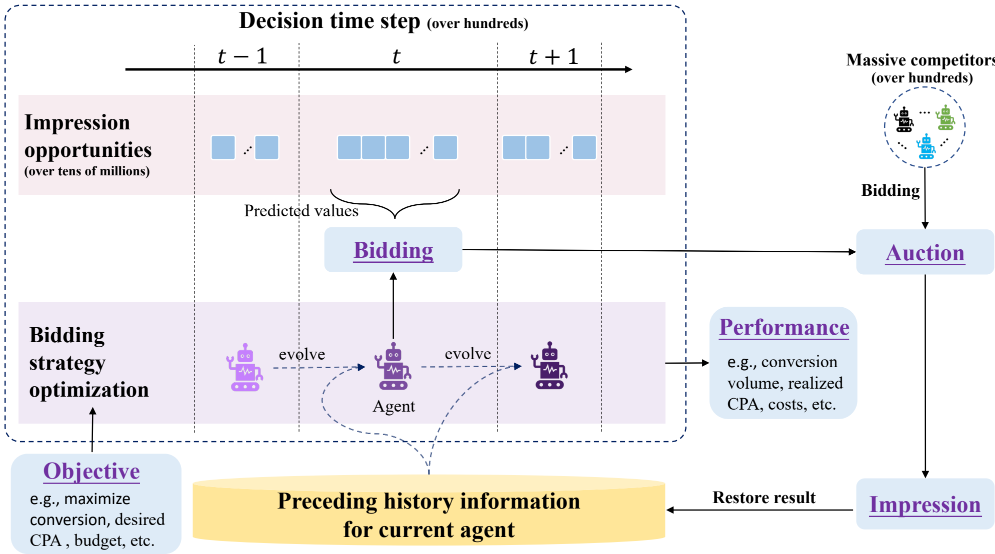

Figure 2: Overview of the standardized ad auction system. The auto-bidding agent sequentially bids for impression opportunities at each step, evolving based on the preceding information.

1.6 Training Dataset

The training dataset is derived from advertising data generated via the auction system where massive agents compete against each other. This data can be used to model the auction environment and train the auto-bidding agent. It should be noted that the training samples are non-overlapping with the evaluation data in the auction system. Participants need to avoid overfitting to the training data and strive to engineer an agent endowed with robust generalization ability.

The training dataset includes 21 advertising periods, each containing more than 500,000 impression opportunities and divided into 48 steps. Each opportunity includes 50 agents[4] with the bids. The dataset comprises over **500 million** records, totaling **40 GB** in size. Each record includes information such as the predicted conversion value, bid, auction, and impression results, among other details. The specific data format is as follows:

(c1). deliveryPeriodIndex: The index of the current delivery period.

(c2). advertiserIndex: The unique identifier of the advertiser.

(c3). advertiserCategoryIndex: The index of the advertiser's category.

(c4). budget: The advertiser's budget for a period.

(c5). CPAConstraint: The CPA constraint of the advertiser.

(c6). timeStepIndex: The index of the current decision time step.

(c7). remainingBudget: The advertiser's remaining budget before the current step.

(c8). pvIndex: The index of the impression opportunity.

(c9). pValue: The conversion action probability when the advertisement is exposed to the customer.

(c10). pValueSigma: The variance of predicted probability.

(c11). bid: The agent's bid for the impression opportunity.

(c12). xi: The winning status of the agent for the impression opportunity.

(c13). adSlot: The won ad slot.

(c14). cost: The cost needs to be paid if the ad is exposed to the customer.

(c15). isExposed: The indicator signifying whether the ad in the slot was displayed to the customer.

[4]Real-world data show that 50 agents can ensure competitive pressure for auto-bidding agent training.

(c16). conversionAction: The indicator signifying whether the conversion action has occurred.

(c17). leastWinningCost: The minimum cost to win the impression opportunity.

(c18). isEnd: The completion status of the advertising period.

1.7 Metrics

The auto-bidding agent aims to maximize conversion volume while satisfying the CPA constraint set by the advertiser. If the realized CPA exceeds the advertiser's desired CPA C, a penalty will be applied. The detailed evaluation metrics are defined as follows:

$$Score = \mathcal{P}(CPA; C) \cdot \sum_i x_i \cdot E_i \cdot V_i. \tag{1}$$

In real-world scenarios, if the advertiser's total costs exceed the budget during the advertising period, the ad platform typically stops the agent's bidding for the remaining opportunities. Therefore, the budget constraint can always be satisfied. The current evaluation metric focuses solely on the CPA constraint, where the penalty function $\mathcal{P}(CPA; C)$ for exceeding the CPA constraint is defined as:

$$\mathcal{P}(CPA; C) = \min\left\{\left(\frac{C}{CPA}\right)^{\beta}, 1\right\}$$

The penalty function is defined with a hyperparameter $\beta > 0$, typically set to 3. $\mathcal{P}(CPA; C)$ implies that the penalty is incurred only when $CPA > C$.

In each evaluation, the agent implemented by the participant is required to bid on behalf of one designated advertiser, given a specific budget, CPA, and other settings. To fully evaluate the performance of the agent across various advertisers, we will run it multiple times in the auction system using different advertiser profiles and periods, and then average the results as the final score. During evaluation, profiles and strategies of competitors' auto-bidding agents are kept confidential.

1.8 Baselines, Code, and Material Provided

Baselines. According to research [6], the optimal bid for opportunity i in the single-slot GSP mechanism is given by the formula:

$$b_i = \alpha \cdot C \cdot pvalue_i$$

Here, C denotes the CPA constraint, while α represents the parameter that the auto-bidding agent needs to optimize. In order to adapt to changes in competitors' agents and randomness of opportunity arrival, α must be dynamically adjusted at each step based on previous performance within a period. This constitutes a typical sequential decision-making problem to determine α, which can be addressed using decision-making algorithms. We provide several industry-proven methods as baselines within the training framework, including reinforcement learning [6, 8] and AIGB models [4].

Training Framework. We also provide a training framework to help participants implement and evaluate their auto-bidding agents. This framework includes three modules: data processing, agent training, and offline evaluation. The complete process for several effective baseline methods is included in the training framework. Participants can utilize this framework to develop a well-trained auto-bidding agent based on the training dataset as described in Section 1.6.

1.9 Website, Tutorial and Documentation

The competition will be hosted on the Tianchi platform, supported by Alibaba Group. Since being founded in 2014, Tianchi has hosted more than 500 competitions and established a global network of over 1200,000 developers spanning over 100 countries and territories. In addition, Tianchi has collaborated with international prestige conferences such as KDD, CVPR, IJCAI, ICDM, and CIKM to host nearly 100 competitions. The Tianchi platform offers a full range of services for the competition, from user management to discussion forums and leader-boards.

We will publish API documentation, FAQs and tutorial for the various components of the environment on the Tianchi platform. In addition to the leader-board that displays the scores of the participants, we will establish a dedicated forum on the Tianchi platform, staffed by a specialized team to address questions posed by participants. This forum will also allow participants to share experiences and insights with each other. Any modifications to rules or requirements will be promptly communicated to all participants. Ensuring that everyone is on the same page and has an equal opportunity for success is very important. We will prominently display our contact email address on the competition page to make it easy for participants to seek assistance from us. We will launch a website when the competition starts.

2 Organizational Aspects

2.1 Protocol

In this competition, participants are free to use a variety of tools and methods to optimize their auto-bidding agents. We have defined standard input and output interfaces for the participants. Consequently, participants need only to implement the '`PlayerBiddingStrategy`' class, which accepts input information and outputs bids for impression opportunities at each decision step.

Participants are required to package their local code into a Docker image and push it to the competition platform as described in Section 1.9, where it will be pulled and executed for evaluation. We will provide a base Docker image containing all necessary packages. Participants' code must meet specific time efficiency requirements, and any code exceeding the predetermined execution time threshold will be considered invalid. High-performance machines will be used for evaluation, with their specifications provided to participants beforehand. Participants may submit their work multiple times per day for performance evaluation within the competition platform. After the evaluation, participants can query their scores on the leader-board.

Recently, we organized a confidential internal competition within the group to fully test the competition platform, involving approximately 200 participants. We will continue to conduct rigorous testing before the game begins.

2.2 Rules

Some important rules are listed below. The complete rules will be published on the competition website.

1. Participants must not engage in activities that would disrupt the evaluation servers.
2. All code sharing among participants must be conducted in the public forum, ensuring transparency and accessibility to all.
3. Participants are limited to a maximum of five submissions per calendar day.
4. The winning teams are required to submit a technical report of their solution, detailing methods, data analysis, and performance analysis, as well as the complete source code.

We will closely monitor the submission platform for any signs of illegal activity, particularly if different participants submit identical or similar codes or settings. We commit to carrying out thorough investigations into all accusations. We will immediately disqualify those involved if we find clear evidence of cheating.

2.3 Competition Stages

The competition consists of two rounds:

Preliminary Round Participants need to implement their agents to compete against the pre-trained agents provided by the organizer. Each participant's evaluation is conducted separately and independently. All participants are consistently ranked according to their evaluation scores, with the top m advancing to the finals. The parameter m will be set according to participant numbers to guarantee sufficient finalists.

Final Round In the final round, advancing participants need to further optimize their agents. Final rankings will be determined using the same evaluation metrics and methods as in the preliminary round. The winning teams will be determined by the results of the final round.

2.4 Schedule and Readiness

Timeline of the competition:

- Start of the competition(registration): June 25, 2024
- Preliminary Round: July 22, 2024 - September 15, 2024
- Final Round: September 16, 2024 - October 24, 2024
- Announcement of Winning Teams: November 15, 2024

Throughout the competition, we will organize various promotional activities to draw attention and participation. Registration will remain open until the end of the preliminary round.

2.5 Competition Promotion and Incentives

To enhance the diversity of the participants, we will simplify complex domain knowledge and provide high-quality sample code and training frameworks to help those new to this research field.

Our skilled operations team will leverages sufficient resources, including Alibaba's relationship network and fund, to promote the competition. Promotional materials will be created, such as posters and videos, to highlight the benefits and features of the competition, utilizing the following channels:

- **Targeted Participant Outreach:** Identify potential participants by retrieving their contact information from papers in relevant fields, in order to reach out via email and invite them to join the competition (reach: 3000+ precise target audiences).
- **Institutional Collaboration and Invited Talks:** Collaborate with relevant research institutions, and organize invited talks to spread the word about the competition (reach: 2000+ precise target audiences).
- **Digital Platform Advertising:** Advertising on digital platforms like Google, Facebook, and Twitter to expand potential audience (reach: 1000+ precise target audiences).
- **Tianchi Community Engagement:** Invite potential participants from the Tianchi developer community by sending emails. (reach: 1,00,000+ broad target audiences).

The competition will offer a variety of awards to the winning teams. The details are as follows:

- **Cash Prize:** The total cash prize pool for this competition is $30,000 (pre-tax).
- **Alibaba & Peking University Study Tour:** The winning teams have the opportunity to visit Alibaba and Peking University, and engage in face-to-face communication with industry engineers and academic researchers.
- **Internships and Visiting Scholar:** We will prioritize offering long-term internship opportunities for students and long-term visiting scholar opportunities for researchers at Alibaba.

Through these strategies, our goal is to create an inclusive environment that encourages widespread participation, thus fostering more novel and diverse solutions.

3 Resources

3.1 Resources Provided by Organizers

We have raised $70,000 to cover the winning teams' prizes, free computing resources, and promotional expenses. We will offer GPU/CPU cloud computing resources and the complete machine learning suite to participants for free on Tianchi. Additionally, we have assembled a professional support team to provide technical assistance for all aspects of the competition. This includes maintaining the training framework, competition servers, and baselines, as well as promptly answering any questions.

3.2 Support Requested

For the winning teams, it would be beneficial if NeurIPS could provide funds to enable their in-person attendance at the conference to present their work.

4 Organizing Team[5]

Alibaba formed with research institutions, including Peking University and the University of Michigan, to create a team to prepare for this competition. The team consists of 15 members, including algorithm experts, engineering experts, research scholars, operations experts, and others. The following delineates responsibilities for each individual:

Organization, including management, coordination, securing funding and resources: **Jian Xu** (Alibaba Group), **Zhilin Zhang** (Alibaba Group), **Zongqing Lu** (Peking University), **Xiaotie Deng** (Peking University), **Michael Wellman** (University of Michigan), **Chuan Yu** (Alibaba Group), **Bo Zheng** (Alibaba Group).

Task Design and Solution Evaluation, including problem design and data preparation, testing of the competition system and evaluation of the participants' solutions: **Zhilin Zhang** (Alibaba Group), **Shuai Dou** (Alibaba Group), **Yusen Huo** (Alibaba Group), **Zhiwei Xu** (Chinese Academy of Sciences), **Kefan Su** (Peking University), **Ningyuan Li** (Peking University), **Zhijian Duan** (Peking University).

Competition Platform, including development of auction system and competition website, training framework and baselines; implementation of the competitors' strategy ; and management of competition platform: **Shaopan Xiong** (Alibaba Group), **Shuai Dou** (Alibaba Group), **Zhilin Zhang** (Alibaba Group), **Chuang Liu** (Alibaba Group), **Yusen Huo** (Alibaba Group), **Zhiwei Xu** (Chinese Academy of Sciences).

Operations, including design and implementation of promotional activities: **Wei Gong** (Alibaba Group), **Shuai Dou** (Alibaba Group), **Zhijian Duan** (Peking University),**Zhilin Zhang** (Alibaba Group).

Detailed introduction of team members

Mr. **Jian Xu** is Chief Algorithm Architect and Senior Director of Alimama, Alibaba Group. He leads the team in innovating and converting cutting-edge technologies into the Alimama ad platform, driving its continued growth. He has published over 60 papers in the areas of data mining, machine learning, and computational advertising. He also serves as a program committee member and reviewer for various top academic conferences and journals in these areas. He has (co-)organized several competitions, such as the CIKM 2019 e-commerce challenge. DBLP Link

Mr. **Zhilin Zhang** is a Staff Algorithm Engineer at Alimama, Alibaba Group. He is responsible for providing high-performance and scalable solutions for auto-bidding and ad auctions. His research interests include auction design, deep learning, and reinforcement learning. He has published 12 papers on these topics at top-tier international conferences such as KDD and WWW in recent years. Some of his representative works include online RL-based bidding, Deep GSP, and Neural Auctions. He has organized several competitions centered on reinforcement learning and online advertising techniques within the Alibaba Group. Google Scholar

Prof. **Zongqing Lu** is a tenured Associate Professor in the School of Computer Science at Peking University. He also leads the Multimodal Interaction Research Center at Beijing Academy of Artificial Intelligence (BAAI). His research interests include reinforcement learning, foundation models, and general agents. He has published more than 70 papers in prestigious journals and conferences. He serves as an inaugural associate editor for ACM JATS, area chairs for ICML and ICLR, and reviewers for journals like Nature Machine Intelligence. Google Scholar

Prof. **Xiaotie Deng** is an ACM Fellow, IEEE Fellow, and a foreign member of Academia Europaea. He is a Chair Professor at Peking University and the Director of the Multi-Agent Intelligent Research Center at Peking University. His research interests include algorithmic game theory and Internet economics. He has published over 200 papers, with over 10,000 Google Scholar citations. He has served as chairman of multiple international academic conferences and initiated the globally-circulating conference WINE. He has (co-)organized several competitions, such as the CCF Computational Economics Competition. Google Scholar

[5]This section does not count toward the 8-page limit.

Prof. **Michael P. Wellman** is a Professor and Division Chair of Computer Science & Engineering at the University of Michigan. His research has focused on computational market mechanisms and game-theoretic reasoning methods. He has served as Chair of the ACM Special Interest Group on Electronic Commerce (SIGecom), and as Executive Editor of the Journal of Artificial Intelligence Research. He is a Fellow of the Association for the Advancement of Artificial Intelligence and the Association for Computing Machinery. He has published over 200 papers, with over 22,000 Google Scholar citations. He has (co-)organized several competitions, such as the TAC Ad Auctions game. Google Scholar

Mr. **Chuan Yu** is a Senior Staff Algorithm Engineer and Director of Alimama, Alibaba Group. He is responsible for developing marketing solutions for advertisers with advanced technology and product features such as auto-bidding. His research interests include pattern recognition, reinforcement learning, and computational advertising. He has published over 20 papers at top-tier international conferences such as KDD, ICML, and NeuIPS in recent years.

Mr. **Shuai Dou** received his master's degree at Peking University. He is currently an Algorithm Engineer at Alimama, Alibaba Group. His research interests include reinforcement learning, computational advertising, and natural language processing. He has published 2 papers in international conferences including WWW and COLING.

Mr. **Yusen Huo** is an Advanced Algorithm Engineer at Alimama, Alibaba Group. His research focuses on reinforcement learning and computational advertising. He has published 3 papers in international conferences including KDD, WWW, and NeurIPS.

Mr. **Zhiwei Xu** is a Ph.D. student at Institute of Automation, Chinese Academy of Sciences. He is interning at Alibaba Group as an Algorithm Engineer and will join Alibaba Group formally in July 2024. His work mainly focuses on reinforcement learning and multi-agent systems. He has published 18 papers at conferences such as NeurIPS, AAAI, and AAMAS. Google Scholar

Mr. **Zhijian Duan** is a Ph.D. student at Peking University, where he received his Bachelor's degree in 2020. He is presently interning at Alibaba Group. His research primarily focuses on machine learning, algorithmic game theory, and computational advertising. He has published 8 papers in prominent international conferences such as ICML, NeurIPS and AAMAS. Google Scholar

Mr. **Haopan Xiong** is an Advanced Development Engineer at Alibaba Group. His research interest focuses on designing and implementing large-scale training and service frameworks for advertising. The 'XingYun' framework developed by their team has been widely used within the Alibaba Group.

Mr. **Chuang Liu** is an Advanced Development Engineer at Alibaba. His research interests focus on designing and implementing robust and scalable competition platforms. The TianChi competition platform developed by their team has successfully hosted over 500 competitions.

Mr. **Ningyuan Li** is a Ph.D. student in CFCS, Peking University and an intern at Alibaba Group. He received his bachelor's degree from the Turing Class at Peking University in 2023. His research mainly focuses on computational economics and algorithmic game theory. He has published 5 papers in international conferences including WWW, KDD, WINE and AAAI. Google Scholar

Mr. **Kefan Su** is a Ph.D. student at Peking University, where he obtained a bachelor's degree at the School of Electronic Engineering and Computer Science in 2020. He is currently an intern at Alibaba Group. His work mainly focuses on multi-agent reinforcement learning algorithms. He has published 4 papers in ICML, CVPR, TMLR, and AAMAS. Google Scholar

Mrs. **Wei Gong** is an operations expert at Alibaba Group. She leads the team in successfully promoting and operating tens of large-scale competitions on the Tianchi platform, including collaborations with prestigious international conferences such as CVRP, CIKM, KDD.

Dr. **Bo Zheng** is the Engineering VP of Alimama, Alibaba Group, responsible for the overall technical work of Alimama's advertising technology division, including algorithms, machine learning, and engineering architecture. He has published almost 100 papers in the areas of AI and computational advertising. He also serves as a program committee member and reviewer for various top academic conferences and journals in these areas. Google Scholar

References

[1] About Target CPA bidding. https://support.google.com/google-ads/answer/6268632. 2024.

[2] Online Advertising Spending Worldwide. https://www.statista.com/statistics/237974/online-advertising-spending-worldwide. 2024.

[3] Anurag Ajay, Seungwook Han, Yilun Du, Shuang Li, Abhi Gupta, Tommi Jaakkola, Josh Tenenbaum, Leslie Kaelbling, Akash Srivastava, and Pulkit Agrawal. Compositional foundation models for hierarchical planning. Advances in Neural Information Processing Systems, 36, 2024.

[4] Jiayan Guo, Yusen Huo, Zhilin Zhang, Tianyu Wang, Chuan Yu, Jian Xu, Yan Zhang, and Bo Zheng. Aigb: Generative auto-bidding via diffusion modeling. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pages 2993–3001, 2024.

[5] Ben Hambly, Renyuan Xu, and Huining Yang. Recent advances in reinforcement learning in finance. Mathematical Finance, 33(3):437–503, 2023.

[6] Yue He, Xiujun Chen, Di Wu, Junwei Pan, Qing Tan, Chuan Yu, Jian Xu, and Xiaoqiang Zhu. A unified solution to constrained bidding in online display advertising. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, pages 2993–3001, 2021.

[7] Patrick R. Jordan and Michael P. Wellman. Designing an ad auctions game for the trading agent competition. In International Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis, 2009.

[8] Haoming Li, Yusen Huo, Shuai Dou, Zhenzhe Zheng, Zhilin Zhang, Chuan Yu, Jian Xu, and Fan Wu. Trajectory-wise iterative reinforcement learning framework for auto-bidding. Proceedings of the ACM Web Conference 2024, 2024.

[9] Kolby Nottingham, Yasaman Razeghi, Kyungmin Kim, JB Lanier, Pierre Baldi, Roy Fox, and Sameer Singh. Selective perception: Learning concise state descriptions for language model actors. In NeurIPS 2023 Foundation Models for Decision Making Workshop, 2023.

[10] Mariano Schain and Yishay Mansour. Ad exchange – proposal for a new trading agent competition game. In Agent-Mediated Electronic Commerce. Designing Trading Strategies and Mechanisms for Electronic Markets, pages 133–145, 2013.

[11] Yuxin Wu, Tianyang Zhao, Haoyuan Yan, Min Liu, and Nian Liu. Hierarchical hybrid multi-agent deep reinforcement learning for peer-to-peer energy trading among multiple heterogeneous microgrids. IEEE Transactions on Smart Grid, 2023.

[12] Wei Zhang, Yanjun Han, Zhengyuan Zhou, Aaron Flores, and Tsachy Weissman. Leveraging the hints: Adaptive bidding in repeated first-price auctions. Advances in Neural Information Processing Systems, 35:21329–21341, 2022.

[13] Yiheng Zhu, Yang Zhan, Xuankun Huang, Yuwei Chen, Jiangwen Wei, Wei Feng, Yinzhi Zhou, Haoyuan Hu, Jieping Ye, et al. Ofcourse: A multi-agent reinforcement learning environment for order fulfillment. Advances in Neural Information Processing Systems, 36, 2024.