Dynamic Budget Adjustment in Search Auctions

Jie Zhang[1], Yanwu Yang[1], Rui Qin[1], Daniel Zeng[1,2] and Xin Li[3]

[1]The State Key Lab of Intelligent Control and Management of Complex Systems,
Chinese Academy of Sciences, Beijing 100190, China

[2]Department of Management Information Systems, The University of Arizona, USA

[3]Department of Information Systems, City University of Hong Kong, Kowloon Tong, Hong Kong

{zhangjie.isec, yangyanwu.isec, qinrui.isec}@gmail.com, zeng@email.arizona.edu, xin.li@cityu.edu.hk

Abstract: With serious advertising budget constraints, advertisers have to adjust their daily budget according to the performance of advertisements in real time. Thus we can leave precious budgets to better opportunities in the future, and avoid the surge of ineffective clicks for unnecessary costs. However, advertisers usually have no sufficient knowledge and time for real-time advertising operations in search auctions. We formulate the budget adjustment problem as a state-action decision process in the reinforcement learning (RL) framework. Considering dynamics of marketing environments and some distinctive features of search auctions, we extend continuous reinforcement learning to fit the budget decision scenarios. The market utility is defined as discounted total clicks to get during the remaining period of an advertising schedule. We conduct experiments to validate and evaluate our strategy of budget adjustment with real world data from search advertising campaigns. Experimental results showed that our strategy outperforms the two other baseline strategies.

Keywords: search auctions; budget adjustment; reinforcement learning; dynamical adjustment

1 Introduction

Recently there have been a rapid growth of search auctions when ``economics meet search" [6]. In search auctions, a bidding contract is triggered once as an information request is submitted. The high volume of search demands makes the bidding a continuous and infinite process. Once any advertiser amends her search advertisements at any time, ranking results and cost-per-clicks would be changed accordingly. Thus, advertisers have to continuously monitor the market and adjust their advertising strategies to respond to marketing dynamics. Budget is an endogenous factor in auctions [2] that heavily influences advertising strategies in search auctions. Advertisers need to make good use of their limited budgets to maximize revenues. In search auctions, how to rationally allocate the limited budget is a critical issue. An effective advertising budget allocation strategy should be able to dynamically allocate and adjust advertising budget according to status of the marketing environment.

Advertisers usually do not have sufficient knowledge and time for real-time advertising operations in search auctions. Thus, their budget is usually fixed, being taken as various constraints for other advertising strategies in search advertising practices. On the other hand, there have been some research efforts along the line of budget allocation over keywords. However, these works are not operationally suitable to practical paradigms of search auctions provided by major search engines (e.g., Google). Most of these studies [5, 7] fall into the category of bidding strategies, rather than budget strategies. By considering the entire lifecycle of search advertising, budget decisions in search auctions occur at three levels [8]: allocation across search markets, temporal distribution over a series of slots (e.g. day) and adjustment of the remaining budget (e.g., the daily budget). The objective of this work is to explore an effective solution for dynamic adjustment of the daily budget, in order to achieve high-return clicks and avoid less effective clicks.

In this paper, we formulate the budget adjustment problem as a state-action decision process in the reinforcement learning (RL) framework, as the strategic decision of advertising budget can be viewed as a special multi-stage dynamic decision problem. Since major state variables in this problem such as the daily budget are continuous values, we model advertising decisions of the budget adjustment based on a continuous-time, continuous-state reinforcement learning approach [3]. We extend this continuous reinforcement learning (CRL) approach to fit budget decision scenarios in search auctions, from three perspectives. First, as some search auction providers (e.g. Google) permit only a limited number of times for daily budget adjustments per day, we amend CRL to deal with continuous-time, continuous-state, discrete-realtime-action cases with step functions, where actions can be taken at any time during a given period (e.g. day). Consider that an advertiser wants to make optimal policies of budget adjustment for N intervals and the advertising budget assigned to these N intervals is bounded. Secondly, dynamical systems' behaviors of agents are usually given as a prior. However, in search auctions we don't have access to such a system of differential equations in advance. Thus we employ Back Propagation (BP) neural networks with some rules based on search auction scenarios for system fitting. Thirdly, we provide a novel numeric solution for our CRL-based budget adjustment model. The market utility is defined as discounted total clicks to get during the remaining period of an advertising schedule. Furthermore, we design some experiments to validate and evaluate our strategy of budget adjustment with real world data from search advertising campaigns. Experimental results illustrate the superiority of our strategy over the two other baseline strategies. The rest of this paper is organized as follows. In Section 2, we state the problems and challenges of dynamical budget adjustment in search auctions, and present our budget adjustment strategy based on continuous reinforcement learning. Section 3 provides a numerical solution for our proposed model. Section 4 reports some experimental results to validate our model, through comparing with two baseline strategies. Section 5 concludes this work.

2 Dynamic Budget Adjustment in Search Auctions

2.1 The Problem Statement

In search auctions, an advertiser usually sets a daily budget for each campaign in which a set of keywords are selected and bids over these keywords are assigned. If the daily budget is used up before the end of the advertising schedule, her advertisements will not be shown up in the rest of the day, which may result in the loss of potential clicks. On the other hand, if the daily budget is set too high, there is a risk of wasting money on ineffective clicks without valuable actions expected by advertisers. As the marketing environment of search auctions changes over time, an effective budget adjustment strategy should entitle advertisers to dynamically adjust the daily budget.

The strategic decision of advertising budget can be viewed as a special multi-stage dynamic decision problem with Markov properties [4, 1]. Decisions at time t depends on both current marketing state and decisions at time $t-1$. This makes RL an appropriate technique to model budget adjustment in search auctions. Nevertheless, there are still several challenges need to be addressed. First, search auctions are a continuous process in term of high volume of search demands. It demands CRL with flexible components suitable to encode various states and actions in search auctions. Secondly, due to the dynamical complexity of search auctions, it is impossible to get an explicit dynamic system in CRL, e.g., a system of differential equations, as a prior. Thirdly, budget constraints make the decision space discontinuous, which hinders the derivation of rewards in the direction of policy to get necessary conditions for optimal actions.

2.2 The Budget Adjustment Model

Suppose that an advertiser has the daily budget d and bids over keywords of interest q. Let c denote the cost per click (CPC), p the effective CTR, and b the remaining daily budget during the advertising schedule. Then, $\boldsymbol{X} = \{x_1, x_2, ...\}$ represents a set of environment states $\boldsymbol{x} = (c, p, b)$ in search auctions, $\boldsymbol{U} = \{u_1, u_2, ...\}$ a set of actions $\boldsymbol{u} = (d, q)$ that the advertiser can take, and a policy $\boldsymbol{\mu}$ can be identified with a mapping from the set of states $\boldsymbol{x}$ to the set of actions $\boldsymbol{u}$, that is, to select advertising actions based on the current state.

The environment state in search auctions is time-varying, thus we represent it with differential equations, $\dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t), \boldsymbol{u}(t))$, which denotes the changing rate of the environment state. We don't have access to such a system of differential equations (e.g., f) in advance, thus we employ BP neural networks based on some rules of search auction scenarios for system fitting.

The number of times for budget adjustment is limited by some search markets such as Google. If an advertiser adjusts the remaining daily budget and bids simultaneously, then $\boldsymbol{u}$ can be given as a step function (e.g. N pieces)

$$\boldsymbol{u}(t) = \begin{cases} \boldsymbol{u}_1, & if\ 0 \le t < t_1 \\ \ldots \\ \boldsymbol{u}_N, & if\ t_{N-1} \le t < T. \end{cases}$$

The control variable $\boldsymbol{u}$ consists of $\boldsymbol{u}_1, t_1, \boldsymbol{u}_2, t_2, \ldots, \boldsymbol{u}_N$. In other words, it includes $2N-1$ parameters in total. The optimal action $\boldsymbol{u}^* = (\boldsymbol{u}_1^*, t_1^*, \boldsymbol{u}_2^*, t_2^*, \ldots, \boldsymbol{u}_N^*)$ characterizes a policy of budget adjustment with maximized profits in terms of effective clicks. At time t_m, the values of $\boldsymbol{u}_1, t_1, \boldsymbol{u}_2, t_2, \ldots, \boldsymbol{u}_{t_m}, t_m$ are already chosen by the advertiser, leaving only $2N-1-2m$ parameters to determine.

At time t, if action $\boldsymbol{u}(t)$ is taken based on state $\boldsymbol{x}(t)$, then the system state transits to $\boldsymbol{x}(t+\Delta t)$. Then, the cost from t to $t+\Delta t$ can be represented by the advertising expenditure $b(t) - b(t+\Delta t)$, and the advertiser gets $p(t)(b(t) - b(t+\Delta t))/c(t)$ effective clicks. Let $r(t)$ represent the instant reward at time t, when $\Delta t \to 0$, $r(t) = r(\boldsymbol{x}(t), \boldsymbol{u}(t)) = -\frac{b'(t)}{c(t)}p(t)$, which is a continuous function on time t. The current estimate of the value function $V(t) = \int_t^T e^{-\frac{s-t}{\tau}} r(\boldsymbol{x}(s), \boldsymbol{u}(s))\mathrm{d}s$, where $\tau \in (0, 1]$ is a discount factor. It represents the total discounted effective clicks from time t to T.

Then, the objective is to find a policy that can maximize the future total discounted reward $V(t)$. We formulate this problem in the reinforcement learning (RL) framework, as follows,

$$\begin{aligned} \boldsymbol{u}(t) = \quad & \underset{\boldsymbol{u} \in U(t)}{\operatorname{argmax}}\, V(t) \\ \text{s.t.} \quad & \dot{\boldsymbol{x}}(t) = f(\boldsymbol{x}(t), \boldsymbol{u}(t)) \\ & V(t) = \int_t^T e^{-\frac{s-t}{\tau}} r(\boldsymbol{x}(s), \boldsymbol{u}(s))\mathrm{d}s \end{aligned} \tag{1}$$

where $U(t)$ is the feasible set of actions, and varies over time.

The peculiar characteristics of model (1) from other RL models lie in three aspects specific to distinctive features in search auctions. First, rather than taking f as a determined function, it can not be given explicitly in advance due to the dynamics of system environments, which is trained from the field data of advertising performance collected. Secondly, $\boldsymbol{u}$ in model (1) is a step function with limited steps, instead of a continuous variable. However, in our case, actions of budget adjustment can be taken at any time during a given period (e.g. day). Thirdly, the advertising schedule makes our model time-bounded.

3 The Solution

In this section, we provide a numerical solution for our CRL-based budget adjustment model. First, we deal with the function f. We can extract some priori knowledge about f from the field historical data collected from practical advertising campaigns, and get a better estimation of the function f with the optimal action $\boldsymbol{u}^*$ on the current state. Thus, we train a BP neural network to approximate the function f, and use reward-based mechanism to make the approximation more and more perfect. In details, we initialize f arbitrarily, then at time t it can be improved based on the reward at $t-1$; after many iterations, the approximation function of f will approach to the real

f. The training and learning processes of f might be time-consuming, however we can accelerate them with some domain knowledge in search auctions.

Secondly, we discuss the estimation of $\partial V^*(\boldsymbol{x})/\partial \boldsymbol{x}$. A universal approximation $V(x;w)$ is used to approximate the total discounted reward function V. Utilizing the Hamilton-Jacobi-Bellman equation, the optimal total reward V^* should satisfy conditions [3] $\dot{V}(t) = \frac{1}{\tau}V(t) - r(t)$ and

$$\frac{1}{\tau}V^*(\boldsymbol{x}(t)) = \max_{\boldsymbol{u}[t,T]}\left[r(\boldsymbol{x}(t),\boldsymbol{u}(t)) + \frac{\partial V^*(\boldsymbol{x})}{\partial \boldsymbol{x}} f(\boldsymbol{x}(t),\boldsymbol{u}(t))\right].$$

If we define $\delta(t) \equiv r(t) - \frac{1}{\tau}V(t) + \dot{V}(t)$, which is called the TD error. Denote $E(t) \equiv \frac{1}{2}|\delta(t)|^2$. Minimizing E with gradient descent algorithm, we can obtain the optimal reward $V^*(\boldsymbol{x};w)$, thus $\partial V^*(\boldsymbol{x})/\partial \boldsymbol{x}$.

Thirdly, we propose a way to find the optimal action $\boldsymbol{u}^*$. In CRL, the control variable $\boldsymbol{u}$ can be chosen in the whole action space U, so it can be obtained by many kinds of optimization algorithms such as gradient descent algorithm. However, in search auction, $\boldsymbol{u}$ can only be a step function with limited number of times, which makes it extremely difficult to find the solution $\boldsymbol{u}^*$. This problem is beyond the solving ability of traditional optimization algorithms. With consideration of some characteristics of $\boldsymbol{u}$, we solve this problem with the following method. Suppose we are at time t, and the last time for adjustment is t_m. That is, actions $\boldsymbol{u}_1^*, \cdots, \boldsymbol{u}_{t_m}^*$ and the real adjustment time $t_1^*, \cdots, t_m^*$ before time t is known certainly, leaving $\boldsymbol{u}_{t_{m+1}}^*, \cdots, \boldsymbol{u}_{t_{N-1}}^*$ and $t_{m+1}^*, \cdots, t_{N-1}^*$ unknown. Thus, at t, we have $2N-1-2m$ parameters to determine, which can be obtained by maximizing

$$r(\boldsymbol{x}(t),\boldsymbol{u}) + \frac{\partial V^*(\boldsymbol{x})}{\partial \boldsymbol{x}} f(\boldsymbol{x}(t),\boldsymbol{u}).$$

We utilize Q-Learning to find the optimal action $\boldsymbol{u}^*$ that minimize the temporal difference E during the current day. In detail, we initialized $Q(\boldsymbol{x},\boldsymbol{u})$ arbitrarily, where $Q(\boldsymbol{x},\boldsymbol{u})$ is the estimated utility function, which tells how good the action $\boldsymbol{u}$ is in a given state $\boldsymbol{x}$; then at time $t+1$, we will choose $Q(\boldsymbol{x}(t+1),\boldsymbol{u}(t+1))$ based on the state and action set at time t, i.e.,

$$Q(\boldsymbol{x}(t),\boldsymbol{u}(t)) = r(\boldsymbol{x}(t),\boldsymbol{u}(t)) + \gamma \max_{\boldsymbol{u}(t+1)} Q(\boldsymbol{x}(t+1),\boldsymbol{u}(t+1))$$

where $\gamma \in (0,1)$ is the relative value of delayed vs. immediate rewards. Our aim is to find the action that maximize Q, which is obviously a recursive process. After a limited number of iterations, we can obtain the optimal action $\boldsymbol{u}^*$ and then the optimal policy μ^*.

The above three steps for solving model (1) are proceeding at any time t, and influenced by each other. Repeat the three steps from time 0 to T, until the optimal action $\boldsymbol{u}^*$ is obtained ultimately.

4 Experimental Validations

4.1 Experiment Setup

We collected data on field reports and logs for practical search advertising campaigns of an e-commerce website. In search auctions, a click is an action initiating a visit to a website via a sponsored link, and if a click is an intentional click that has a realistic probability of generating values once the visitor arrives at the website, then it is a *valid click*, otherwise it is *invalid*. In this work, we consider that the generated value will be obtained by some kinds of user behaviors such as purchase, registration, staying on the landing page for more than 5s, surf more than 2 links, bookmarking or downloading relevant pages. We give the concept of effective click-through rate (CTR) as follows.

Definition 1 (Effective CTR). *Effective CTR is the ratio of valid clicks and total clicks, i.e.,*

$$\textit{Effective CTR} = \frac{\textit{valid clicks}}{\textit{total clicks}}.$$

Note that the effective CTR is equivalent to the conversion rate if these kinds of user behaviors are defined as conversion actions by advertisers. Figure 1 illustrates some data characteristics including CPC, effective CTR and the remaining budget throughout the day for instance with a fixed daily budget \$100 and bids \$3.00. These data can be used to initialize and train function f.

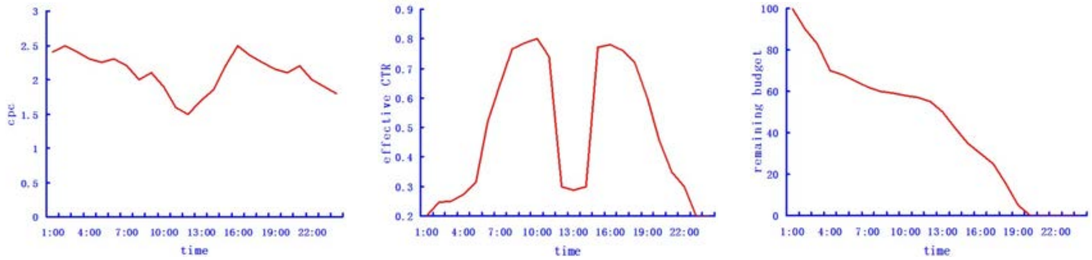
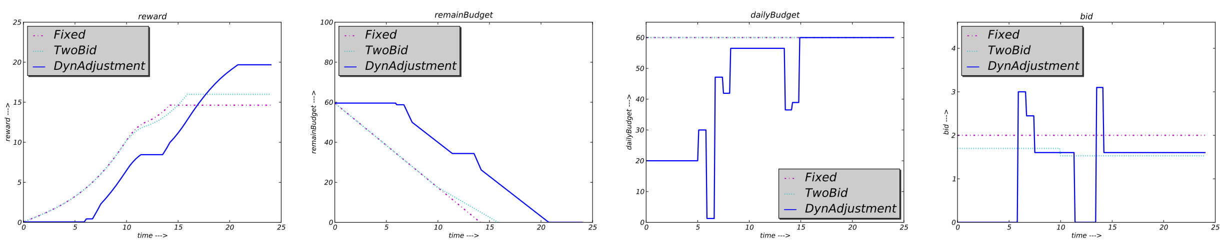


Figure 1: CPC, effective CTR and the remaining budget during a given day

4.2 Experimental Results

In experiments, we intend to compare our strategy with two baseline strategies, namely Fixed-strategy and TwoBid-strategy. Fixed-strategy is a strategy considers both the daily budget and bids fixed throughout the day. This strategy can be usually seen when the advertisers lack of time and knowledge for advertising decisions. TwoBid-strategy [5], randomizes bidding between a value q_1 on all keywords, and the other value q_2 on all keywords until the budget is exhausted. Note that TwoBid-strategy doesn't consider daily budget adjustment. Our proposed strategy (denoted as DynAdjustment), is capable of both daily budget and bid adjustment.

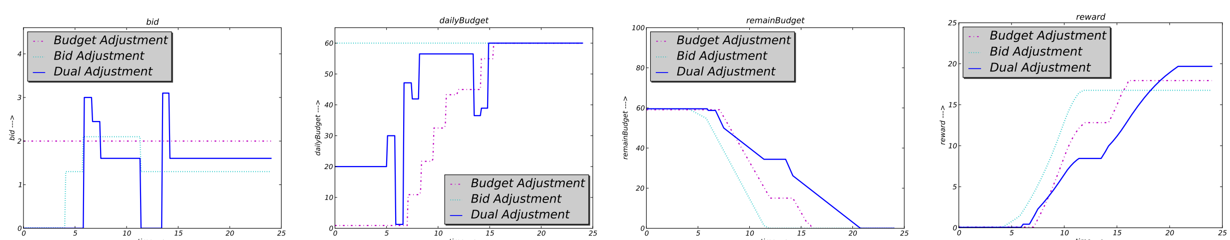
The comparisons of rewards and policies for these three strategies are shown in Figure 2. From Figure 2 (a), we can see our strategy provides the advertiser the highest revenue with 19.681 effective clicks, followed by TwoBid-strategy with 16.000 effective clicks. Fixed-strategy has the smaller number of 14.634 effective clicks. Our strategy outperforms these two strategies, with 23.020% and 34.491%, respectively. This can also be explained by the fact that our strategy can avoid wasting the budget on ineffective clicks, as shown in Figure 2 (b), through daily budget adjustment (Figure 2 (c)) and bid adjustment (Figure 2 (d)).



(a) (b) (c) (d)

Figure 2: Comparisons of budget adjustment strategies

Both daily budget and bid adjustment can increase advertising revenues for advertisers. We made experiments to validate the effect of these two operations individually, by controlling parameters in our model. Figure 3 depicts the comparison results. The revenue in terms of effective clicks from daily budget and bid adjustment are 17.951 and 16.759, respectively. It is surprising to know, the effect of daily budget adjustment is larger than the effect of bid adjustment in our case. This phenomenon might occur occasionally, which is interesting to explore empirically in future.



(a) (b) (c) (d)

Figure 3: Comparisons of effects of the daily budget and bid adjustment

4.3 Managerial Insights

Our paper provides critical managerial insights for advertisers in search auctions. On the one hand, advertisers usually take the budget as simple constraints, and put a lot of efforts to find effective operations as defined by various kinds of markets (e.g. search auctions). This work indicates that a simple strategy for budget adjustment can, to some degree, improve advertising effects in terms of effective clicks. On the other hand, most advertisers pay more attention to bidding strategies in order to either minimize the loss or maximize advertising performance. In this research, we provide opportunities for advertisers to adjust both the daily budget and bids over keywords of interest. It is proved that, dual-adjustment of these two factors together could significantly facilitate the journey to advertising goals.

5 Conclusions

In this paper, we propose a CRL-based budget strategy for the daily budget and bid adjustment. By considering dynamics of search marketing environments, our budget strategy could deal with continuous-time, continuous-state, discrete-realtime-action cases with step functions, where actions can be taken at any time during a given period (e.g. day). We also provide a numerical solution to our model, and conduct some experiments to validate it with real world data from practical advertising campaigns. Experimental results illustrate the superiority of our strategy over the two other baseline strategies. This work also report some preliminary results of dual adjustment of the daily budget and bids. In ongoing works, we are intended to explore (a) the theoretical basis and possible empirical evidences for co-optimization of advertising budget and bids; (b) inter-operation of dynamical budget adjustment across several search markets; and (c) efficient computational algorithms to facilitate online implementations.

Acknowledgements

We are thankful to anonymous reviewers who provided valuable suggestions that led to a considerable improvement in the organization and presentation of this manuscript. This work is partially supported by NSFC (71071152, 70890084, 71025001, 71102117, 60875049, 60921061) and CAS (2F09N05, 2F09N06).

References

[1] Archak N., V. S. Mirrokni, S. Muthukrishnan (2010). Budget optimization for online advertising campaigns with carryover effects, 6th Workshop on Ad Auctions, Harvard.

[2] Benoit, J.P., V. Krishna (2001). Multiple-object auctions with budget constrained bidders, Review of Economic Studies, 68(1), 155--179.

[3] Doya, K. (2000). Reinforcement learning in continuous time and space, Neural Computation, 12, 215--245.

[4] Du R., Q. Hu , S. Ai (2007). Stochastic optimal budget decision for advertising considering uncertain sales responses, European Journal of Operational Research, 183, 1042--1054.

[5] Feldman J., S. Muthukrishnan, et al. (2007). Budget optimization in search-based advertising auctions, Proceedings of the 8th ACM Conference on Electronic Commerce, California USA.

[6] Jansen, B. J., A. Spink (2007). Sponsored search: Is money a motivator for providing relevant results?, IEEE Computer, 40(8), 50--55.

[7] Muthukrishnan S., M. Pál, Z. Svitkina (2010). Stochastic models for budget optimization in search-based advertising, Algorithmica, 58, 1022--1044.

[8] Yang, Y., J. Zhang, R. Qin, J. Li, F. Wang, W. Qi (2011). A budget optimization framework for search advertisements across markets, IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans, Accepted.