A Two-stage Fuzzy Budget Allocation Model in Search Auctions

Rui Qin[1,2,3], Yong Yuan[1,2,3], Juanjuan Li[1,2,3]
[1]The State Key Laboratory of Management and Control for Complex Systems,
Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2]Qingdao Academy of Intelligent Industries, Qingdao, China
[3]Beijing Engineering Research Center of Intelligent Systems and Technology
Institute of Automation, Chinese Academy of Sciences, Beijing,China
Email: qinrui.isec@gmail.com, yong.yuan@ia.ac.cn, lijuanjuan.isec@gmail.com

***Abstract*—Budget optimization is an important issue faced by advertisers in search auctions, and has significant impact on the design of various advertising strategies. Given a limited budget on a search market during a certain period, an advertiser has to distribute her budget to a series of sequential temporal slots (i.e., days, weeks, or months), during which advertisers must avoid the budget being used up quickly, so as to keep the budget for potential clicks with better performance in the future. Considering the optimal budgets over these temporal slots as fuzzy variables, we establish a two-stage fuzzy budget allocation model, and use particle swarm optimization (PSO) algorithm to solve it in case when these optimal budgets are characterized by discrete fuzzy variables. We also conduct experiments to validate our model and algorithm. The experimental results show that our model can outperform other five budget allocation strategies in terms of reducing the revenue loss of the advertiser.**

Keywords: budget optimization; budget allocation; sponsored search auctions; two-stage fuzzy programming; PSO

I. INTRODUCTION

Search auction is an important marketing channel for many Web advertisers. In search auctions, major search engines such as Google and Yahoo! provide advertising slots for advertisers to display their advertisements alongside the organic search results on search engine result pages (SERP). Advertisers participating in search auctions should first select a set of keywords that are relevant to their products or services, and then submit a bid for each keyword. Once a search user submits a query, an auction will be triggered among the advertisers who bid for keywords matching the query. Advertisers will not pay unless their advertisements are clicked.

In search auctions, how to allocate the limited budget rationally is a significant issue faced by advertisers. Fruchter and Dou [5] used dynamic programming and derived analytical solutions for the optimal budget allocation decisions between a generic market and a specialized market. They found that in the long run, an advertiser must always spend more on the specialized market. The budget optimization problem can be formulated as an online (multiple-choice) knapsack problem [1], [2] to achieve a provably optimal competitive ratio for the advertisers. Several stochastic models were also established [3], [4], [6], [10] to help distribute a given amount of budget over a set of keywords so as to maximize the expected number of clicks.

However, most of these works are focused on the budget allocation on the level of each keyword, which is not suitable to real-world market practice. Generally speaking, there are three budget allocation scenarios at different phases in the entire life-cycle of search auctions [11], [12], including the long-term allocation across markets prior to ad campaigns, the allocation over a series of intervals (e.g. daily budget constraints) for a specific ad campaign, and the real-time adjustment of budget in a given interval. Strategies of budget allocation and adjustment at these three scenarios construct a close-loop, composite allocation strategy for search auctions through constraints and feedbacks. This paper is aimed at finding an effective solution for budget allocation in the second scenario.

This paper focuses on how to allocate the budget to a series of sequential temporal slots for a specific ad campaign on a search market. The optimal budget over each temporal slot is influenced by many factors, e.g., search demands from search users and click-through-rate of advertisements, which are not easy to know in advance. Considering the optimal budgets over these temporal slots as fuzzy variables [9], we establish a two-stage fuzzy budget allocation model based on the two-stage fuzzy programming approach as proposed in [8], where advertisers give an initial budget at the first stage, and then adjust the budget according to the distribution of these optimal budgets at the second stage. Particle swarm optimization (PSO) algorithm is used in our paper to find the optimal solution in case when the optimal budgets are characterized by discrete fuzzy variables. We also make experiments to validate our budget allocation model and algorithm. The experimental results show that our model performs better in terms of reducing the revenue loss of the advertisers, comparing with other five budget allocation strategies.

The main contributions of this paper can be summarized as follows.

- (i) We established a two-stage fuzzy budget allocation model for search auctions, with the optimal budgets characterized by fuzzy variables.
- (ii) We presented the solution procedure of the proposed model with PSO algorithm in case when the optimal budgets are characterized as discrete fuzzy variables.
- (iii) We conducted experiments to validate our proposed model and its solution algorithm.

The rest of this paper is organized as follows. In Section

II, we first state our problem, and then establish a two-stage fuzzy programming model. In Section III, we propose a solution method to the proposed model. In Section IV, we conduct experiments to evaluate the effectiveness of our budget allocation model and corresponding solution algorithm. Section V concludes this paper.

II. Problem Formulation

In this section, we first state our problems, and then present the two-stage fuzzy budget allocation model over a series of promotional slots. The notations used in this paper are listed in Table I.

TABLE I. List of notations

Notation	Definition
n	total temporal slots in the whole period
B	total budget on the search market during a certain period
c_i	clicks per unit cost of the ith temporal slot, $i = 1, 2, \ldots, n$
x_i	the allocated budget for the ith temporal slot, $i = 1, 2, \ldots, n$
$\tilde{d}_i$	the fuzzy optimal budget for the ith temporal slot, $i = 1, 2, \ldots, n$
p_i	the effective CTR of the ith temporal slot below the optimal budget, $i = 1, 2, \ldots, n$
$p_i^{'}$	the effective CTR of the ith temporal slot above the optimal budget, $i = 1, 2, \ldots, n$
I_i^+	the exceeded budget of the ith temporal slot, $i = 1, 2, \ldots, n$
I_i^-	the lacking budget of the ith temporal slot, $i = 1, 2, \ldots, n$

A. Problem Statement

Suppose the total budget of an advertiser in a search market during a certain period (i.e., a week/month/year) is given, the advertiser need to distribute the budget to a series of sequential temporal slots (i.e., several days/weeks/months), to maximize her total revenue or minimize her total loss.

Since the potential search demand in each temporal slot is not uniformly distributed, the effective clicks cannot grow with the same speed. Generally, the effective CTR may be higher in some time intervals of the temporal slot, and lower in other time intervals. Thus, there exists a critical value for the allocated budget. When the allocated budget is lower than the critical value, the clicks will have a higher effective CTR, while, when the allocated budget is higher than the critical value, the clicks will have a lower effective CTR for the part of budget exceeding the critical value.

In this paper, the critical value is regarded as the optimal budget. Besides the concept of *the optimal budget*, there are two relevant concepts called *the exceeded budget* and *the lacking budget*, represented by I^+ and I^-, respectively. We define the three terms as follows:

- **The optimal budget**: The optimal budget is the budget which can perfectly satisfy the cost for the time intervals with high effective CTR. In other words, if the advertiser has enough total budget, then her best choice is to set the budget of each temporal slot equal to the optimal budget.
- **The exceeded budget**: If the allocated budget x_i for the ith temporal slot is higher than the optimal budget d_i, then the difference between the allocated budget and the optimal budget is referred to as *the exceeded budget*, represented by $I_i^+ = x_i - d_i$. We assume this part of budget will be used up and the effective CTR is $p_i^{'} < p_i$, where p_i is the effective CTR for the part of budget below the optimal budget. This assumption is reasonable because of the fact that the allocated budget is far more than the optimal budget will never happen due to the advertiser's financial conditions.
- **The lacking budget**: If the allocated budget x_i for the ith temporal slot is lower than the optimal budget d_i, then the shortage of budget between the allocated budget and the optimal budget is referred to as *the lacking budget*, represented by $I_i^- = d_i - x_i$. The loss of effective clicks for the advertiser will rise with this part of budget increasing.

According to the above definitions, it is easy to obtain that at least one of them is equal to zero for the ith temporal slot. That is, if $x_i > d_i$, then $I_i^+ = x_i - d_i > 0$ and $I_i^- = 0$; if $x_i < d_i$, then $I_i^- = d_i - x_i > 0$ and $I_i^+ = 0$; if $x_i = d_i$, then $I_i^+ = I_i^- = 0$. Thus, we can denote I_i^+ and I_i^- as follows

$$I_i^+ = [x_i - d_i] \vee 0, I_i^- = [d_i - x_i] \vee 0,$$

where the symbol "$\vee$" is defined as

$$x \vee 0 = \begin{cases} x, & \text{if } x \geq 0 \\ 0, & \text{otherwise.} \end{cases}$$

From the above descriptions, it is obvious that the optimal budget of each temporal slot is of great significance for the advertiser in budget allocation problems. However, the optimal budget over each temporal slot is influenced by many factors, e.g., search demands from search users and CTR of advertisements, which are not easy to know in advance. However, some information of the optimal budget can be obtained according to the promotional logs and the advertiser's experience, such as the possibility distribution. Thus, in the following section, we will characterize the optimal budgets by fuzzy variables, and establish a two-stage fuzzy budget allocation model to solve this problem.

B. Formulation of Two-stage Fuzzy Budget Allocation Model

Suppose the total budget allocated to a search market during a certain period is B, and we need to allocate it to n sequential temporal slots. Let $\tilde{d}_i$ be the fuzzy optimal budget for the ith temporal slot, and c_i be clicks per unit cost of the ith temporal slot, $i = 1, 2, \ldots, n$. The objective of the advertiser is to minimize his/her loss, which includes the obtained ineffective clicks and the lost effective clicks. Thus, in the ith temporal slot, the loss can be computed in the following three cases:

Case 1: If $I_i^+ = I_i^- = 0$, then the allocated budget x_i is equal to the optimal budget, thus the loss is the ineffective clicks generated from x_i, i.e., $c_i x_i (1 - p_i)$.

Case 2: If $I_i^+ > 0$, then the allocated budget x_i is higher than the optimal budget. In this case, the allocated budget can be divided into two parts: $x_i - I_i^+$ with effective CTR p_i and I_i^+ with effective CTR $p_i^{'}$. Thus the loss includes two parts: the ineffective clicks generated from $x_i - I_i^+$, i.e., $c_i(x_i - I_i^+)(1 - p_i)$, and the ineffective clicks generated from I_i^+, i.e., $c_i I_i^+(1 - p_i^{'})$. Moreover, the exceeded budget I_i^+ can also generate some effective clicks, i.e., $c_i I_i^+ p_i^{'}$, which can cancel out part of the ineffective clicks. Therefore, the total loss can be computed as

$$\begin{aligned} &c_i(x_i - I_i^+)(1 - p_i) + c_i I_i^+(1 - p_i') - c_i I_i^+ p_i' \\ &= c_i x_i(1 - p_i) + c_i I_i^+(p_i - 2p_i'), \end{aligned}$$

of which $c_i I_i^+(p_i - 2p_i')$ is generated due to the existence of I_i^+.

Case 3: If $I_i^- < 0$, then the allocated budget x_i is lower than the optimal budget. In this case, the effective CTR for x_i is p_i, then the ineffective clicks generated from x_i is $c_i x_i(1-p_i)$. Moreover, the lacking budget I_i^- can lead to the loss of effective clicks (i.e., $c_i I_i^- p_i$) in time intervals with high effective CTR. Thus, the total loss can be computed as $c_i x_i(1 - p_i) + c_i I_i^- p_i$, of which $c_i I_i^- p_i$ is generated due to the existence of I_i^-.

Since the optimal budget for each temporal slot cannot be known in advance, we should give a budget in the first stage, and then adjust the budget based on the realization of the fuzzy optimal budget in the second stage. As discussed above, the loss will rise with I^+ or I^- increasing. The loss generated by I^+ and I^- are $c_i I_i^+(p_i - 2p_i')$ and $c_i I_i^- p_i$, respectively, which should be minimized in the second stage. With the notations in Table I, the two-stage fuzzy budget allocation model can be established as follows

$$\begin{array}{ll} \min & \sum\limits_{i=1}^{n} c_i x_i(1 - p_i) + \mathcal{Q}(\mathbf{x}) \\ \text{s.t.} & \sum\limits_{i=1}^{n} x_i \le B \\ & x_i \ge 0, i = 1, 2, \cdots, n, \end{array} \tag{1}$$

where

$$\mathcal{Q}(\mathbf{x}) = E_{\tilde{\mathbf{d}}}[Q(\mathbf{x}, \tilde{\mathbf{d}}(\gamma)] \tag{2}$$

and

$$\begin{array}{ll} Q(\mathbf{x}, \mathbf{d}(\tilde{\gamma})) & = \min \left[\sum\limits_{i=1}^{n} c_i I_i^+(p_i - 2p_i') + \sum\limits_{i=1}^{n} c_i I_i^- p_i \right] \\ \text{s.t.} & I_i^+ = [x_i - \tilde{d}_i(\gamma)] \vee 0 \\ & I_i^- = [\tilde{d}_i(\gamma) - x_i] \vee 0, i = 1, 2, \cdots, n, \end{array} \tag{3}$$

where $\mathcal{Q}(\mathbf{x})$ is called recourse function [8], and E is expected value of a fuzzy variable [9].

III. A Solution Algorithm

In this section, we first discuss the computational method of the recourse function (3) when the optimal budget of each temporal slot is characterized by a discrete fuzzy variable, and then propose a solution algorithm for our model.

A. Computing the Recourse Function (2)

Let the optimal budget $\tilde{\mathbf{d}} = (\tilde{d}_1, \tilde{d}_2, \cdots, \tilde{d}_n)$ for the n temporal slots be a discrete fuzzy vector with the following possibility distributions

$$\begin{aligned} &\hat{\mathbf{d}}^1 = (\hat{d}_1^1, \hat{d}_2^1, \ldots, \hat{d}_n^1) \text{ with possibility } \mu_1 > 0, \\ &\hat{\mathbf{d}}^2 = (\hat{d}_1^2, \hat{d}_2^2, \ldots, \hat{d}_n^2) \text{ with possibility } \mu_2 > 0, \\ &\cdots \\ &\hat{\mathbf{d}}^N = (\hat{d}_1^N, \hat{d}_2^N, \ldots, \hat{d}_n^N) \text{ with possibility } \mu_N > 0, \end{aligned}$$

and $\max_{j=1}^{N} \mu_j = 1$. Then according to Liu [8], the recourse function (2) can be computed in the following way.

Without loss of generality, we assume that for a fixed $\mathbf{x}$, the second-stage objective function satisfies the condition $Q(\mathbf{x}, \hat{\mathbf{d}}^1) \le Q(\mathbf{x}, \hat{\mathbf{d}}^2) \le \cdots \le Q(\mathbf{x}, \hat{\mathbf{d}}^N)$, then the recourse function (2) at $\mathbf{x}$ is computed by the formula

$$\mathcal{Q}(\mathbf{x}) = \sum_{j=1}^{N} \omega_j Q(\mathbf{x}, \hat{\mathbf{d}}^j), \tag{4}$$

where the corresponding weights $\omega_j, j = 1, 2, \cdots, N$ are given by the following formulas

$$\begin{aligned} &\omega_j = \tfrac{1}{2}(\max_{k=1}^{j} \mu_k - \max_{k=0}^{j-1} \mu_k) + \tfrac{1}{2}(\max_{k=j}^{N} \mu_k - \max_{k=j+1}^{N+1} \mu_k), \\ &j = 1, 2, \cdots, N \end{aligned} \tag{5}$$

($\mu_0 = 0, \mu_{N+1} = 0$) and satisfy the following constraints

$$\omega_j \ge 0, \sum_{j=1}^{N} \omega_j = \max_{j=1}^{N} \mu_j = 1.$$

B. PSO Algorithm

In this section, we will use particle swarm optimization (PSO) algorithm to solve the proposed two-stage budget allocation model (1)–(3). Inspired by social behaviors of bird flocking or fish schooling, PSO is a population based stochastic optimization technique developed by Kennedy and Eberhart in 1995 [7]. Compared with other evolutionary algorithms, PSO algorithm has much faster convergence speed and fewer parameters to adjust, which makes it particularly easy to implement. Recently, PSO algorithm has attracted much attention and been successfully applied in many fields. In PSO, the system is initialized with a population of random solutions and searches for optima by updating generations. The potential solutions, called particles, fly through the problem space by following the current optimum particles. In the search process, the velocity and position of the ith particle are updated by the following formulas

$$\begin{aligned} V_i(t+1) &= wV_i(t) + \alpha_1 r_1(P_i(t) - X_i(t)) \\ &\quad + \alpha_2 r_2(P_g(t) - X_i(t)) \end{aligned} \tag{6}$$

$$X_i(t+1) = X_i(t) + V_i(t+1), \tag{7}$$

where $i = 1, 2, \ldots, pop\_size$; w is the inertia coefficient; α_1 and α_2 are learning rates and r_1 and r_2 are two numbers randomly generated from [0,1].

The process of PSO for solving model (1)–(3) can be described as follows.

Step 1. Initialize pop_size particles with random positions and velocities, and evaluate the objective values for all particles. For each particle, the objective value is evaluated in the following way:

> Step 1.1. Compute the optimal value Q_i^* of the second-stage programming (3), $i = 1, 2, \cdots, n$.
> Step 1.2. Rearrange Q_i^*, $i = 1, 2, \cdots, n$ such that $Q_1^* \leq Q_2^* \leq \cdots \leq Q_n^*$.
> Step 1.3. Compute the corresponding weights according to the formula (5).
> Step 1.4. Compute the recourse function (2) according to the formula (4) and then the objective value.

Step 2. Set $pbest$ of each particle and its objective value equal to its current position and objective value, and set $gbest$ and its objective value equal to the position and objective value of the best initial particle;

Step 3. Update the velocity and position of each particle according to the formulas (6) and (7), respectively, and then compute the objective values for all the particles;

Step 4. For each particle, compare the current objective value with that of its $pbest$. If the current objective value is smaller than that of $pbest$, renew $pbest$ and its objective value with the current position and objective value;

Step 5. Find the best particle of the current particle swarm with the smallest objective value. If the objective value is smaller than that of gbest, then renew gbest and its objective value with the position and objective value of the current best particle.

Step 6. Repeat Step 3 to Step 5 for a given number of cycles;

Step 7. Report $gbest$ and its objective value as the optimal solution and optimal value.

IV. Experiments

In this section, we conduct experiments to validate the established two-stage fuzzy budget allocation model and its solution method.

Suppose the budget on a search engine during one week (only run the advertisements on workday) is $B = 300$, and clicks per unit cost and effective CTR of the ith day are given in Figure 1 and 1, respectively. Furthermore, suppose the optimal budget of the five days is a discrete fuzzy vector with the following possibility distributions

$$\tilde{\mathbf{d}} = \begin{cases} (47, 51, 65, 62, 75), & \text{with possibility } 0.6 \\ (40, 50, 78, 59, 73), & \text{with possibility } 0.8 \\ (38, 58, 62, 64, 78), & \text{with possibility } 1 \\ (50, 54, 68, 68, 60), & \text{with possibility } 0.7. \end{cases}$$

With the proposed method, if we set the learning rates $\alpha_1 = \alpha_2 = 2$, and the population size pop_size=100, then a run of PSO algorithm with 5000 generations returns the following optimal solution

$$\mathbf{x}^* = (39.999, 54, 68, 62.216, 60),$$

and the corresponding optimal value is 59.896.

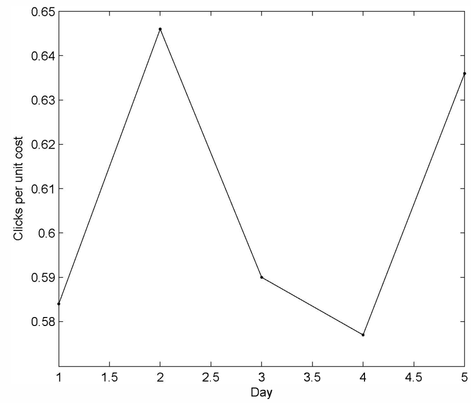

Fig. 1. Clicks per unit cost of the five days

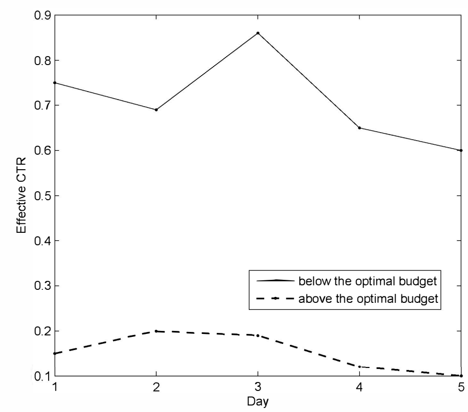

Fig. 2. Effective CTR of the five days

To validate the effectiveness of our method, we make a comparison between the results obtained by our method and the following five baseline strategies:

- BASE-Average: a strategy to allocate the budget to the five days averagely.
- BASE-0.6: a strategy to allocate the budget according to the optimal budget with possibility 0.6.
- BASE-0.8: a strategy to allocate the budget according to the optimal budget with possibility 0.8.
- BASE-1: a strategy to allocate the budget according to the optimal budget with possibility 1.
- BASE-0.7: a strategy to allocate the budget according to the optimal budget with possibility 0.7.

The BASE-Average strategy is commonly used by advertisers in practice since it is easy to implement, and the other four baseline strategies are usually used when the possibility distributions of the optimal budget is known.

The optimal solutions and the corresponding optimal values (e.g., the cumulative loss) of our strategy and the five baseline strategies are illustrated in Figure 3 and Figure 4, respectively. By cumulative loss on the jth day, we refer to the number of loss accumulated from the 1st day to the jth day, $j = 1, 2, \cdots, 5$.

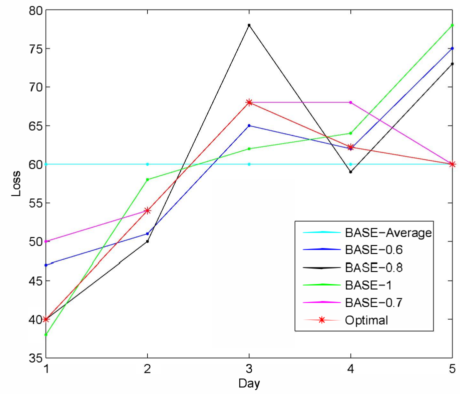

Fig. 3. Comparisons of the daily budget among our strategy and the five baseline strategies

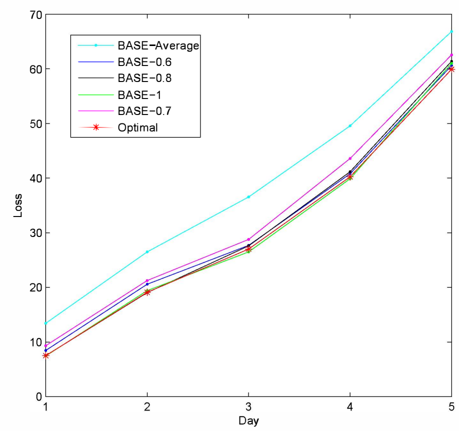

Fig. 4. Comparisons of the loss among our strategy and the five baseline strategies

From Figure 3–4, we can see that,

(1) The optimal solution of our strategy is (39.999, 54, 68, 62.216, 60), and the budget allocation strategy for the five baseline strategies are (60, 60, 60, 60, 60), (47, 51, 65, 62, 75), (40, 50, 78, 59, 73), (38, 58, 62, 64, 78) and (50, 54, 68, 68, 60), respectively.

(2) The loss of our strategy is 59.896, and the loss for the five baselines strategies are 66.830, 60.590, 61.371, 60.995 and 62.557, respectively. The loss of the five baseline strategies are all higher than our strategy, about 11.577%, 1.159%, 4.202%, 1.835% and 4.427%, respectively. It refers that our strategy outperforms the five baseline strategies.

(3) Comparing the five baseline strategies, we can see that, the loss of the BASE-Average strategy is higher than the other four baseline strategies, about 10.299%, 8.895%, 9.566% and 6.831%, respectively. It illustrates that the BASE-Average strategy is the last choice if the advertiser knows the possibility distribution of the optimal budget.

(4) Comparing the four baseline strategies considering the possibility distributions of the optimal budget, we can see that, the loss of the BASE-0.6 strategy is the lowest, followed by the BASE-1 strategy, the BASE-0.8 strategy and the BASE-0.7 strategy. It illustrates that the budget strategies obtained according to the possibility distributions of the optimal budget cannot be evaluated by its possibility.

V. Conclusions

This paper proposed a two-stage fuzzy budget allocation model for the budget allocation problem over a series of sequential temporal slots in search auctions, considering the optimal budget of each temporal slot as a fuzzy variable. We also proposed an algorithm and conducted computational experiments to illustrate the effectiveness of the proposed model. The experimental results show that our model performs better than the other five baseline budget allocation strategies, and when the possibility distribution of the optimal budget is known, consider it when making the budget allocation strategies can greatly decrease the advertiser's loss.

Acknowledgements

This work is partially supported by NSFC (#71472174, #71102117, #71232006, #61233001) and the Early Career Development Award of SKLMCCS (Y3S9021F36, Y3S9021F2K).

References

[1] Babaioff M., N. Immorlica, D. Kempe, R. Kleinberg (2007). A knapsack secretary problem with applications, Lecture Notes in Computer Science, 4627, 16–28.

[2] Chakrabarty D., Y. Zhou, R. Lukose (2007). Budget constrained bidding in keyword auctions and online knapsack problems, Proceedings of the 16th International World Wide Web Conference.

[3] 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.

[4] DasGupta B., S. Muthukrishnan (2010). Stochastic budget optimization in internet advertising, arXiv:1001.2735v1.

[5] Fruchter G. E., W. Dou (2005). Optimal budget allocation over time for keyword ads in web markets, Journal of Optimization Theory and Applications, 124(1), 157–194.

[6] Feldman J., S. Muthukrishnan, M. Pal, C. Stein (2007) Budget optimization in search-based advertising auctions, Proceedings of the 9th ACM Conference on Electronic Commerce.

[7] Kennedy J., R. C. Eberhart (1995). Particle swarm optimization, Proceedings of the IEEE International Conference on Neural Networks, Piscataway, NJ, pp.1942–1948.

[8] Liu Y.K. (2005). Fuzzy programming with recourse, International Journal of Uncertainty, Fuzziness & Knowledge-Based Systems, 13, 382–413.

[9] Liu B., Y.K. Liu (2002). Expected value of fuzzy variable and fuzzy expected value models, IEEE Transactions on Fuzzy Systems, 10(4), 445–450.

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

[11] Yang Y., J. Zhang, R. Qin, J. Li, F. Wang, Q. Wei (2012). A budget optimization framework for search advertisements across markets, IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 42(5), 1141–1151.

[12] Yang Y., R. Qin, BJ. Jansen, J. Zhang, D. Zeng (2014) Budget planning for coupled campaigns in sponsored search auctions, International Journal of Electronic Commerce, 18(3): 39–66.