Research Article Impact Factor: 0.621 ISSN: 2319-507X
Rupali Kamble, IJPRET, 2014; Volume 2 (8): 600-607 IJPRET

INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

SURVEY PAPER ON STOCHASTIC MODELING AND OPTIMIZATION FOR ONLINE ADVERTISEMENTS

RUPALI S. KAMBLE[1], PROF. A. N. JAISWAL[2], PROF. S. P. CHHAWARE[2]

1. M. Tech Student, Dept. of CSE, Priyadarshini College of Engineering, MH, India.
2. G.H.R.I.E.T.W , Nagpur University, MH, India.

Accepted Date: 15/02/2014 ; Published Date: 01/04/2014

Abstract: In this paper, we propose a stochastic modeling and optimization for online advertisements, this stochastic modeling is used to describe how search service providers charge client companies for posting their ads based on users' queries for the keywords related to these companies' ads by using certain advertisement assignment strategies. A stochastic model is non-deterministic model. We also formulate an optimization problem for stochastic modelling to maximize the long-term average revenue for the service provider under each client's long-term average budget constraint and design an online algorithm which captures the stochastic properties of users' queries and click-through behaviors. Optimization problem arises here, so that to solving this optimization problem by making connections to scheduling problems in wireless networks, queuing theory and the stochastic networks.

Keywords: Online Advertising, Optimization, Stochastic Systems, Bidding.

PAPER-QR CODE

Corresponding Author: MS. RUPALI S. KAMBLE

Access Online On:

www.ijpret.com

How to Cite This Article:

Rupali Kamble, IJPRET, 2014; Volume 2 (8): 600-607

INTRODUCTION

Today's world is the Internet World, due to this surprising growth of the Internet has been increased day by day, so due to this surprising growth of the Internet population, the World Wide Web has become a fantastic marketing channel. Among marketing activities, advertising might be one of the most popular activities and the one that can take the most advantages of specialties of the Web. Advertising is the form of marketing communication and it is useful for encourage or manipulate audience such as listeners, readers, and viewers. The advantage of web advertising includes low cost, multimedia, interaction, easy maintenance, round-the-clock and worldwide presentation and almost unlimited advertising space, etc. Besides, Web advertising has ability of combing advertising and purchase behavior through the Internet.

An online advertising service has been the major source of revenue for search service providers such as Google, Yahoo and Microsoft. When an Internet user queries a keyword, alongside the search results, the search engine may also display advertisements from some companies which provide services or goods based on the keyword related to this client companies. These companies pay the search service providers for posting their ads with a specified amount of price for each ad on a pay-per impression or pay-per-click basis [1].

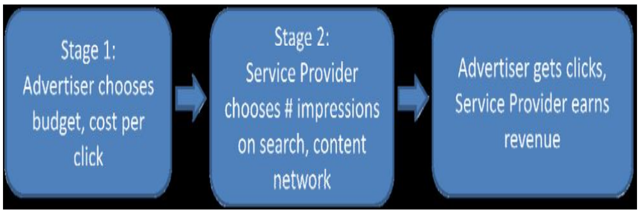

Fig. 1: Time line of the two-stage game-theoretical model.

A system for providing advertisements has been the major source of revenue for search service providers from a central server to viewers who access web sites. The central server stores both advertisements which are to be displayed and an information data base. The data base includes information about viewers included in audience, information about the characteristics of particular web sites and other information relevant to which advertisements should be displayed for particular viewers.

II. AN ARCHITECTURE FOR INTERACTIVE WEB ADVERTISING SYSTEM

We propose architecture for interactive Web advertising system as Figure 2. To satisfy the data requirements, there are three databases: 1)User profile, 2)Site database, and 3)Advertisement database. For keeping information about users, homepages in the Web site, and all advertisements separately. Furthermore, in order to let the proposed mechanisms work well, we propose six agents to collaborate with each other based on the advertising knowledge base [2].

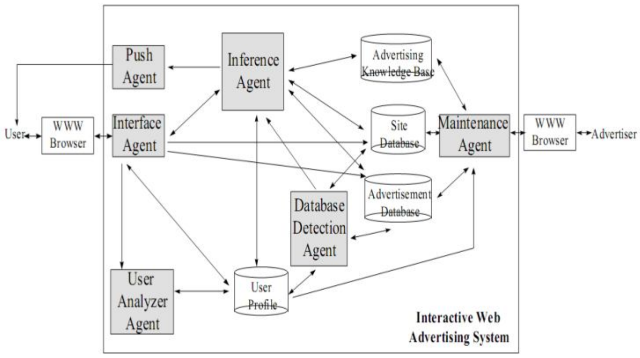

Fig.2: Architecture for Interactive Web Advertising System

Interface agent is responsible for collecting the user data and presenting the advertisements to users. Since there are diverse sources of user data, it has to decide to either update the user profile directly or ask the user analyzer agent to proceed further analysis. The responsibility of the user analyzer agent is to analyze a user's background and online behavior to explore his preference or the interest group to which he belongs. The process of maintaining the user profile is described in Figure 2. If a user has registered, the interface agent can collect the user's information from registration form, cookie, browsed advertisements, transaction records and online messages, etc. For the user who is the first time to visit the site, the interface agent may get the bookmark by asking the user to upload or capture the user's registration information from his other software installations. If the user has visited before but not registered, its information can be got from cookies. Maintenance agent allows the Web site owner to maintain the site database and advertisement database when there are changes. Another agent to maintain these databases is the database detection agent that will detect all changes in the three databases automatically. The inference process is the core operation of an interactive

Web advertising system. Finally, advertisements can be exposed to the users proactively when the advertiser thinks it is the right time to do. The push agent is designed for this purpose.

The basic requirements of an online audience analytics and targeting solution are:

- To collect user behavior across multiple web sites
- To organize and index information regarding users behavior and the content they consume
- To infer demographics and interest data
- To classify new user in real time
- To match demographics and content with ads inventory
- To deliver advertising in the appropriate format and in timely fashion.

III. KEYWORD OPTIMIZATION IN SEARCH BASED ADVERTISING

Keyword optimization for search based advertising has been the focus of some theoretical and applicative work.

Rusmevichientong and Williamson [12] have formulated a model for keyword selection in search based advertising. In their model, the advertiser has a fixed daily budget and each keyword has fixed known cost and profit. However, the keyword click-through probabilities are unknown. The numbers of queries appearing in each of the days, as well as the distribution of keywords, are generated probabilistically with known parameters. They justify their assumptions about keyword costs and distribution by identifying multiple public available data sources that may be used to estimate these parameters. The three stochastic models discussed in [13] are:

1. Fixed proportions model

In fixed proportions model only random variable is the total number of clicks in a day and the proportions of clicks of each of the keywords remains constant. For this model we first prove that an optimal solution is a fractional prefix. In a fractional prefix the advertiser bids on all sorted keywords up to a selected keyword. In this fractional prefix last keyword in the prefix, the advertiser is assigns a probability which is the probability that he will bid on queries from that keyword. Using an interchange argument they prove that there is a fractional prefix and this fractional prefix is optimal here. Further they show that finding the optimal prefix solution is non trivial since there are local maxima. Their solution to the optimization problem overcomes the infinite number of possible fractional prefix solutions by showing that it is sufficient to evaluate a polynomial set of "interesting" solutions which depends on the number of keywords and the number of different values the total number of clicks can take, so that to solving this optimization problem by making connections to scheduling problems in wireless networks, queuing theory and the stochastic networks.

2. *Independent keywords model*

In independent keywords model the number of clicks for each keyword has its own probability distribution (which can be different for different keywords). The number of clicks is independent of this model .The key distinguishing feature of this model which can be different for different keywords). For this model we prove that the prefix solution may not be an optimal solution and there exists a prefix solution which is a 2-approximation to the optimal solution.

3. Scenario model

The scenario model is attempts to capture the full generality of a joint distribution without the large number of bits needed to represent an arbitrary joint probability distribution. So for this a limited (polynomial) number of scenarios in which the exact number of clicks for each word is given. A single scenario has been taken from a given probability distribution over scenarios. In this model, the authors prove two negative results: 1)The keyword optimization problem under the scenario model is NP-hard by using a reduction from CLIQUE and 2)The gap between the optimal fractional prefix solution and the optimal (integer or fractional) solution can be arbitrarily large for the optimization problem.

IV. RELATED WORK

We will only survey the online resource allocation models, Web advertising model, auctions mechanisms.

J. Feldman and S. Muthukrishnan modeled algorithmic methods for sponsored search advertising, in this large commercial search engines have emerged as information gateways for millions of Internet users. In response to a user's query, search engines generate a ranked list of results based on sophisticated information retrieval algorithms [5].

Buchbinder *et al.* showed that matching clients to webpage slots (whether it is a single slot or multiple slots) can be solved as a maximum-weighted matching problem, In the search engine environment, advertisers link their ads to (search) keywords and provide a bid on the amount paid each time a user clicks on their ads ,when users send queries to search engines, along with the (algorithmic) search results returned for each query, the search engine displays funded ads corresponding to *ad-auctions* [6].

By modifying the algorithm in [7], Kalyanasundaram *et al.* the goal of online algorithm is to maximize the number of requests that is services and analyze this problem using standard competitive ratio.

Agrawal *et al.* designed a class of algorithms which achieve a considerably better competitive ratio with accurate estimates the number of query arrivals for each keyword, while still guarantee a reasonably good competitive ratio with inaccurate estimates [10].

Three learning-based algorithms in [5], [6], [7] achieve a near-optimal competitive ratio of based on a random-order arrival model (rather than the adversarial model in most of the earlier work), assuming small bids and knowledge of the total number of queries.

V. NOTEWORTHY CONTRIBUTION

In our implementation, the desired frame shows only the relevant ads by generating the SynSet (Synonym Set of keyword) based on the keyword entered by the user on text field, for this we use lexical database. This lexical database having two tables. First relevant URLs and Second irrelevant URLs. The Relevant Table keeps the identified relevant URLs and the Irrelevant Table keeps the identified irrelevant URLs. The first step in our algorithm puts all out-links from seed pages in a Queue. Each URL taken from the Queue is sent to a function that computes its relevance attribute values (i.e., URL words relevancy, anchor text relevancy, parent pages relevancy, and surrounding text relevancy). Then, NB classifier takes the URL with its attributes values as inputs and makes prediction of its relevancy to the search topic. In particular, we use the optimization decomposition ideas in [7], the stochastic performance bounds in [3] and the modelling of delay-sensitive flows in [8]. Borrowing from that literature, we introduce the concept of an "*overdraft*" *queue*. The overdraft queue measures the amount by which the provided service temporarily exceeds the budget specified by a client. In making the connection to wireless networks, we define something called the "*per-client revenue region,*" which is related to the concept of capacity region in queuing networks (see [9], [11]). In our context, it characterizes the revenue extractable from each client as a function of all the clients' budgets. Our online algorithm exhibits a trade-off between the revenue obtained by the service provider and the level of overdrafts. We can further modify our online algorithm so that clients can always operate strictly under their budgets. Finally, our algorithm and analysis naturally allow us to assess the impact of click-through rate estimation on the service provider's revenue.

Besides the revenue maximization model, we study another model in which the objective is to maximize the average overall click-through rate, subject to a minimum impression requirement for each client. We also show that our results can be naturally extended to handle non-stationary query arrival processes and clients which have short-term contracts with the service provider.

VI. CONCLUSION

In this paper, we propose a stochastic model to describe how search service providers charge client companies for posting their ads based on users' queries for the keywords related to these companies' ads by using certain advertisement assignment strategies. We formulate an optimization problem to maximize the long-term average revenue for the service provider under each client's long-term average budget constraint and design an online algorithm which captures the stochastic properties of users' queries and click-through behaviors. The optimization problem is solving by making connections to scheduling problems in wireless networks, queuing theory and stochastic networks. Our online algorithm is entirely oblivious to query arrivals and fully adaptive, so even non-stationary query arrival patterns and short-term clients can be handled.

REFERENCES

1. Bo Tan, and R. Srikant , "Online Advertisement, Optimization and Stochastic Networks", *IEEE transaction on Automic Control*, Vol.57, No.11, November 2012.

2. Hsiangchu Lai and Tzyy-Ching Yang, " An architecture of interactive web advertising model," *International Journal of Advertising Research* 36, 43-54.

3. I. Menache, A. Ozdaglar, R. Skrikant, and D. Acemoglu, "Dynamic Online Advertising Auctions on Stochastic Scheduling," *IEEE transaction on Information Theory*, vol. 13, no.2, pp. 411–424, Apr. 2010.

4. L. Ying, R. Srikant, A. Eryilmaz, and G. Dullerud, "A large deviations analysis of scheduling in wireless networks," *IEEE Trans. Inform. Theory, vol. 52, no. 11, pp. 5088–5098, Nov. 2006.*

5. J. Feldman and S. Muthukrishnan, "Algorithmic methods for sponsored search advertising," *IEEE Trans. Autom. Control,* vol. 53, no. 3, *pp. 91–122, 2008.*

6. N. Buchbinder, K. Jain, and J. S. Naor, "Online primal-dual algorithms for maximizing ad-auctions revenue," *IEEE Trans. Inform. Theory, vol. 4698,*pp.253–264, 2007.

7. B. Kalyanasundaram and K. Pruhs, "An optimal deterministic algorithm for online b-matching," *IEEE conference on Inform.Theory, Comp. Sci., vol. 233, no. 1–2, pp.* 319–325, 2000.

8. M. J. Neely, "Optimal energy and delay tradeoffs for multiuser wireless downlinks," *IEEE Trans. Inform. Theory, vol. 53, no. 9, pp.*3095–3113, Sep. 2007.

9. M. J. Neely, "Intelligent packet dropping for optimal energy-delay radeoffs in wireless downlinks," *IEEE Trans. Autom. Control, vol.* 54, no. 3, pp. 565–579, Mar. 2009.

10. S. Agrawal, Z. Wang, and Y. Ye, "A dynamic near-optimal algorithm for online linear programming," *IEEE transaction on online optimization techniques*, 2009.

11. M. Mahdian, H. Nazerzadeh, and A. Saberi, "Allocating online advertisement space with unreliable estimates," *IEEE Conference on Electron. Commerce (EC)*, 2007.

12. Paat Rusmevichientong and David P. Williamson. 2006. An adaptive algorithm for selecting profitable keywords for search-based advertising services. In Proceedings of the 7th ACM conference on Electronic commerce (EC'06). ACM, New York, NY, USA, 260-269. DOI=10.1145/1134707.1134736 http://doi.acm.org/10.1145/1134707.1134736

13. S. Muthukrishnan, Martin Pal, and Zoya Svitkina. 2007. Stochastic models for budget optimization in search-based advertising. In Proceedings of the 3rd international conference on Internet and network economics (WINE'07), Xiaotie Deng and Fan Chung Graham (Eds.). Springer-Verlag, Berlin, Heidelberg, 131-142.