Evolutionary Prediction of Online Keywords Bidding

Liwen Hou, Liping Wang, and Jinggang Yang

535, Fahuazhen Rd., Department of Management Information System
Shanghai Jiaotong University, Shanghai, China

Abstract. Online keywords bidding as a new business model for search engine market facilitates the prosperity of internet economy as well as attracts myriad small business to directly target their customers. In order to effectively manage each advertising campaign the manager needs to figure out smart strategy to bid. This paper successively developed two models (static and dynamic) with four crucial variants at keyword bidding (bid, rank, click through and impression) based on the bayesian network to assist the prediction of bidding. Herein dynamic model, evolved from the static model, takes the influence of click through on the last period into account and extends the strategy space. Empirical study by aid of data from the largest online travel agency in China is carried out to test both models and the results indicate that they are effective while the dynamic model is more attractive in terms of prediction accuracy. Finally further research directions along this paper is shown.

Keywords: Dynamic Bayesian Network, Keyword Bidding, Click Through.

1 Introduction

As the giant in search engine market Google earns much money from keywords auction, well appeared as couple of sponsored links listed in the right side of the searched pages. Many others quickly get involved afterwards and ignite this industry. Since being able to provide a feasible approach to advertising through internet while against the expensive traditional media and assumed to directly target customers in favor of those smaller and unknown companies sponsored link Ad does give an edge to business and has boomed since then, just as seen in Fig.1.

Myriad business resorts to sponsored link to get known, even larger enterprises who perceived this synthetic power and advantages to combine online marketing with physical campaign also consider to capitalizing on part of budget to achieve it. Advertisers, however, firstly exited about this new channel and now got confused in the sense that they have no chance to know in detail the bidding process as well as the consumption of their Ad budget, which may stymie the evaluation to their advertising alternatives. In addition there typically exists a many-to-many relationship between the keyword set and advertiser set. That is, each of the majority of keywords will be bided by several advertisers who in turn manage a keyword portfolio. As a result any intention to occupy the highest position in the sponsored links is doomed to exhaust out the Ad budget much quickly and armed with strong incentive to explore lower position. Such extreme asymmetric game definitely aggravates the existence and undulation of the instant equilibria, which pushes advertisers trapped in continuous

G. Psaila and R. Wagner (Eds.): EC-Web 2008, LNCS 5183, pp. 124 – 133, 2008.

and irrational bidding. For example a generic keyword "flower" was bided from ￥0.21 to ￥13.05 for the first position on Google during Dec.2006 in China because 96 companies got involved and triggered off intense competition at that time. Many advertisers afterwards complained that they harvested less than they paid out for this campaign due to lacking the optimal budget management. Consequently the most urgent need for advertisers is to figure out some strategy to bid smartly and avoid suffering as much as possible otherwise the "stupid money" will still stick to them in case of mismanaged campaigns.

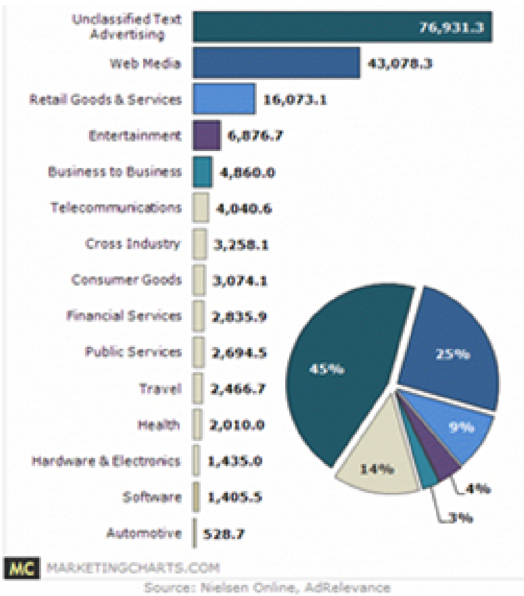

Fig. 1. Sponsored link by industries (11.2007)

At present scholars have exploited this field by different methodologies but still in its infancy so related literature is mostly very recent. Research by Animesh et al (2005) considered different profiles of bidding strategy, and examined the underlying reasons for the customers' quality signaling under a search-experience-credence framework. They also examined the relationship between advertisers' quality and their bidding strategies in online setting. Benjamin et al. (2005) aimed at analyzing the strategic behavior of keyword auctions. They found strategic behavior impaired the benefit of both search engine and market. Feng et al. (2005) inspected four kinds of mechanisms with respect to the correlation of advertiser's willingness to pay and term relevance in the steady state setting and two alternative approaches under dynamic situation. Brendan et al (2004) simulated PPC (pay per click) keyword auction based on intelligent agent technology. A series of rules are provided to the agent imitating bidders and the real bidding process is illustrated in the Overture platform (now part of Google) given such exogenous variables as profit function. Agent technology was also used by Sunju et al. (2004), who elaborately induced agents to intelligently bid using a Markov chain model. They also identified the optimal bidding strategies in an experiment by combining those parameters addressing the auction situation in a given testbed. Young-Hoon (2005) proposed an integrated framework to simulate bidding behavior for keyword auctions on a website. They concluded that out of the four key components willingness to bid is the only significant factor.

This paper will address the prediction of online keywords bidding using bayesian network (BN) technology in an evolutionary sequence from the static to dynamic model according to influencing relationship of bidding profile. The contributions of the paper can be unfolded as the following three aspects: method, evolution view and empirical test. BN is rarely applied to online keyword auction in previous research but does function as an effective vehicle to reflect the causal linkage among such four core variants of keyword bidding strategy as *bid, rank, click_through* (abbr. *CT*) and *impression*. On the other hand Google put on new policy in 2005 to facilitate the promotion of online advertising for small business, which led to the extension of bidding prediction from only one period to successive two periods conditional on new parameters to be learned simultaneously supplemented. This evolution of prediction approximates the reality more by taking dynamics into account so that advertisers can lay down the budget in a relative long term. In order to verify the prediction models this paper collected data from the largest online travel agency in China to test them. The results show that our models are valid and attractive and may be adopted by advertisers with little modification.

The rest of the paper is organized as follows: Section 2 develops a static prediction model based on BN and then addresses types of the selected nodes which represent those parameters (or variants) mentioned above. Considering the request to cover sequential campaigns Section 3 develops a dynamic BN prediction model with the aid of new learning functions. Section 4 verifies our models using empirical data collected from the largest online travel agency in China and the results are proved to be acceptable. Finally, Section 5 summarizes the paper and shows the further research direction.

2 Static Prediction

2.1 Methodology

BN as one of the advanced modeling methods representing the joint probability distributions of variants (i.e. nodes in the network) is consisted of structure model and parameter model. The former, based on a diagram, illustrates the interactive relationship (most likely being causal link) among those variants while the latter addresses the conditional probabilities of variants given their parent nodes. The technology is rather powerful due to couple of potentials, such as multifactor input and prior knowledge or experiences involved. The details can be found in reference [8]. Towards this end people often employ bayes technology in artificial intelligence to assist uncertain reasoning in different domains like medical diagnosis and heuristic search. For online bidding it is also applicable because not only the causal links between these variants need to be simultaneously described but also the bidding history is an important prior for the subsequent bidding strategy. In this paper BN is associated with system modeling for decision, a generic application pro tango.

In general the construction of a BN involves three steps: ①identifying the dominant variants; ②ascertaining interdependent and independent relationships among those variants; ③learning and determining probability parameters for the existing relationship. Herein we will only concentrate the last step because it is the toughest analysis and needs more consideration.

2.2 Model

In online keyword auction *bid*, *rank*, *CT* and *impression* are four the most essential variants with respect to bidding strategy. Herein *impression* refers to appearing times resulted from searching. *Bid* and *rank* have obvious causal link while the *rank* is the cause of *CT* as well if the assumption that higher *rank* brings about more clickstream held. In addition *CT* will change with *impression* in the sense that the number of search of one keyword accrued in some period given *CT* = *impression* * *CT_Rate*. Assuming the higher rank the keyword the more possible to click it we can construct a directed acyclic graph with these four variants shown in Fig.2, which implies: *rank* = *f* (*rank* * *CT*) and *CT*=*g* (*rank*, *impression*). Both functions can further scratch up an implicit probability listed in the following:

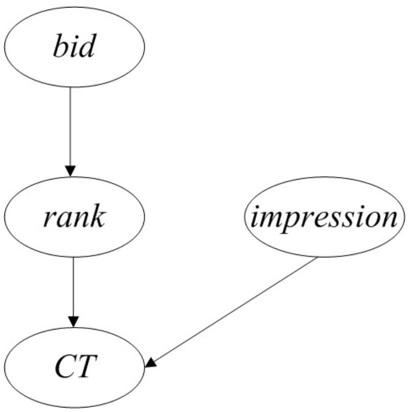

Fig. 2 . Static BN model

$$P(bid, rank, impression, CT) =$$
$$P(bid)*P(rank \mid bid)*P(impression)*P(CT \mid rank, impression)$$

On the other hand total probability formula tells that:

$$P(bid, rank, impression, CT) =$$
$$P(bid)*P(rank \mid bid)*P(impression)*P(CT \mid bid, rank, impression)$$

Both formulae above show the following relationship

$$P(CT \mid bid, rank, impression) = P(CT \mid rank, impression)$$

which means *CT* is independent to *bid* given *rank*, congruent with the reality.

2.3 Parameter Learning

Just as asserted previously that parameter learning is rather challengeable for bayesian prediction and the prerequisite to do that is firstly make clear the type (i.e. discrete, continuous and mixed) of nodes inasmuch as different types are associated with different learning algorithms. We choose the mixed type for all nodes because *bid* and *rank* can be viewed as the discrete variants (also called Tabular nodes) and *impression* and *CT* as the continuous (also called Gaussian nodes). This judgment is a little arbitrary though, it is acceptable if observing change frequency of each variant in fixed period.

The Essential of parameter learning based on history data is to adjust the conditional probabilities associated with the four variants. Therefor the complete maximum likelihood estimate (MLE) needs to be employed to address the learning in the presence of existed data set. Herein a powerful Toolbox BayesNet coded by Kevin

Murphy is used to carry on the prediction process, where five functions shown in the following are involved:

mk_bnet: establishment of structure of BN
bnet.CPD{i}: assign the type to the node i in BN
learn_params: parameter learning function
enter_evidence: variant constraints of reasoning
jtree_inf_engine: inference engine used during parameter learning

3 Dynamic Prediction

3.1 Quality Score

This paper technically aims at Google auction whose ranking algorithm differs from other search engines like Yahoo due to incorporating *CT* rate (abbr. CTR) instead of completely bidding oriented. This restriction, however, ruled out those websites with low CTR appearing on the top position of right side. In 2005 Google changed his ranking algorithm so that part of the websites previously excluded from the anterior searched pages is reignited to strive for favorable positions. This is the effect of quality score defined by Google as the base to evaluate keyword quality and decide lowest bidding. Besides CTR quality score is related with Ad relevance, history records of keyword effect and other minor factors, among which history effect plays the fundamental role in our model because this impact results in a dynamic situation. Towards this point the static model evolves into the dynamic one, inheriting previous performance and allows trial and error.

3.2 Model

The introduction of quality score oscillates the assumption about static model that *rank* is only affected by *bid*. Instead history *CT* of a keyword also heralds such influence. This modification not only brings out a dynamic model of keyword bidding but also allows advertisers to allocate Ad budget more reasonable in long period. This dynamic bayesian model, shown as Fig.3, is developed by incorporating dynamic mechanism into the static model. Dynamic BN as a schematic representation to the complex stochastic process arguments BN by modeling the process of time sequence and the change with time of the stochastic variants set. In this sense dynamic bayesian model paves a path to explore the prediction of keyword bidding with respect to evolutionary view, a more realistic approach for advertisers.

3.3 Parameter Learning

Similar to that of static model dynamic model needs to learn those conditional probabilities as well. Nevertheless a difference between both is the type of nodes. Herein node *CT* and *impression* in the previous phase is considered as the parent node of *rank* and should be labeled as discrete because mixed learning algorithm is still under mature and hence inapplicable to our model. Therefore all nodes in this dynamic model are attributed to discrete (i.e. tabular) type.

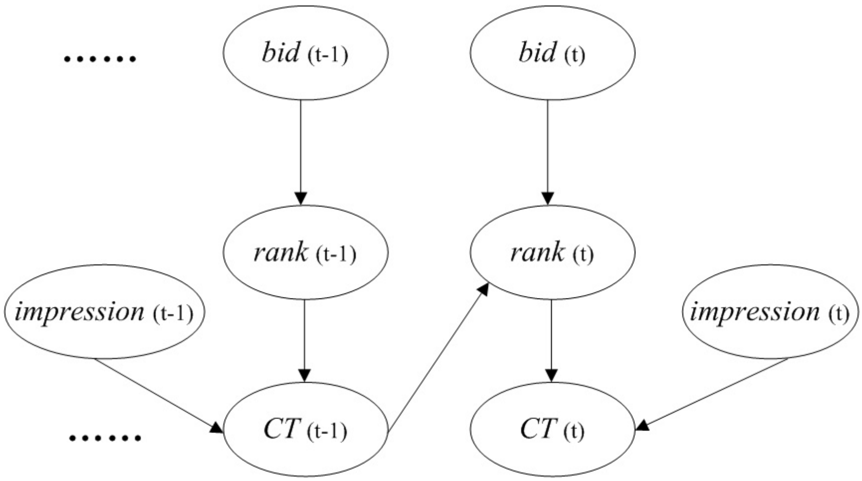

Fig. 3. Dynamic BN model

An accompanying issue to this node treatment is to discretize the continuous variants *CT* and *impression*. Therefor one should firstly select the number of values of continuous variants (i.e., those values can mimic the shape of the variant) to be discretized and then segment the continuous interval corresponding to some period to guarantee each subinterval encompassing one value of one of the continuous variants. Thus the variant can be represented by those discretized values if the number of subintervals enough. Several algorithms such as equal width interval and weighted information-loss discretization can serve for this purpose. This paper chooses the former algorithm, the simplest one because of no category information about type needed. In addition we assume all conditional probabilities of each variant obey uniform distribution so that they can be adjusted adequately by training.

In order to realize the dynamic bayesian model three more functions in the BayesNet Toolbox are involved and shown as the following

mk_dbn: establishment of dynamic structure of BN
dbn.CPD{i}: assign the type to the node i in dynamic BN
jtree_2TBN_inf_engin: dynamic inference engine of two-period interaction

4 Case Analysis

4.1 Data Description

In order to test the validity and effect of the models developed above we luckily acquired the experimental data from the largest online travel agency in China. The dataset, spanning from Dec.21 to Dec.28 2007 on Google platform, includes 299 time slices, each of which tracks such information as bidding time, average bidding, click through, rank and so on. The statistics of the dataset is summarized in the Tab.1.

Table 1. Statistics of bidding dataset

Variants	Average	S.t.	Median	Min	Max
bid	1.01	0.55	1.01	0.01	2.00
rank	8.91	3.47	8.84	1.56	15.77
impression	3005.47	328.00	3038	2074	3956
click through	222.14	90.61	196	46	478
conversion rate	22.26	11.12	19	5	63

Analyzing the relationship of those variants in the dataset we found that the number of search of single keyword within a time slice will change with slice sequence. That is, the slice close to right end incurs more searching times. Fig.4 shows the test result of *impression* using normplot function in Matlab to analogize the normal distribution.

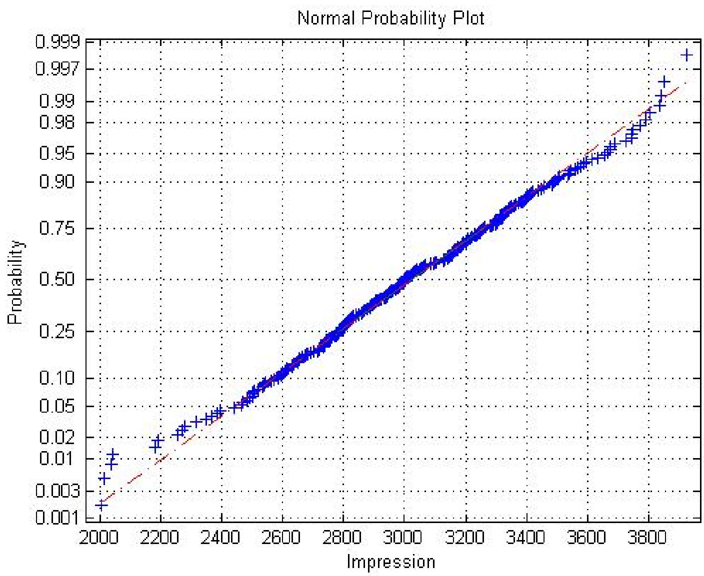

Fig. 4. Normal distribution test of *impression*

Suppose the more the data points juxtapose a line the more likely the distribution of the variant is paralleled to normal distribution. So the searching times of any single keyword, whatever *impression* or other variants, can be assumed to obey normal distribution. Actually we experimented *rank* and obtain the same result.

4.2 Prediction Results

In the presence of structure model of BN shown in Fig.2 and Fig.3 we split the sample dataset into two parts. The first part including 200 time slices serves as the training set which feeds data to the BN models and all parameters (i.e. a series of conditional probabilities) will be computed using the functions in the BayesNet mentioned above. The process is iterative so that the final results can converge to some value. The second part consisted of the rest of 99 slices acts as the tester to make sure the predicted results acceptable.

Fig.5 shows the true value (from the 99 time slices given P(*rank* | *bid*)), predicted value (from our model) and their gap with respect to *rank* from top to down when the static model works.

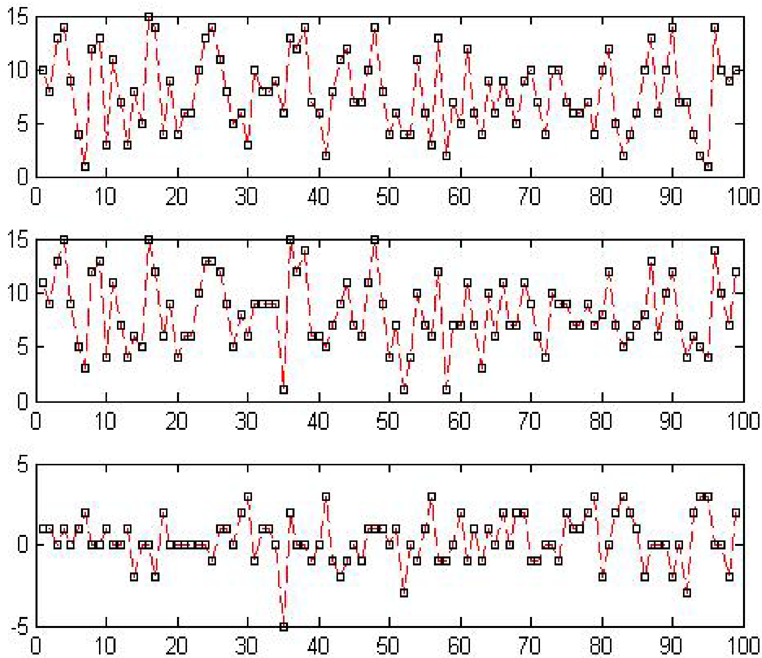

Fig. 5. True value, predicted value and their gap for *rank* based on static model

Similarly we can obtain the schemas of other three variants using this static bayesian model. While the dynamic model roughly remains the same process dynamic inference engine is embedded and works for the learning. Fig.6 shows the predicted result with respect to *rank* as well.

Intuitively dynamic bayesian model seems to be more accurate than its counterpart because more zero points appears in the gap diagram if glimpsing the graph. Actually it is true when the general statistics associated with both predictions above are compared. See Tab.2.

Obviously dynamic model indicates smaller deviation with respect to predicted mean and standard error in terms of variant *rank*. This means dynamic model not only evolves to be more desirable and effective for keyword bidding management but also proves that *CT* in the previous round does influence the *rank* in this round, as shown

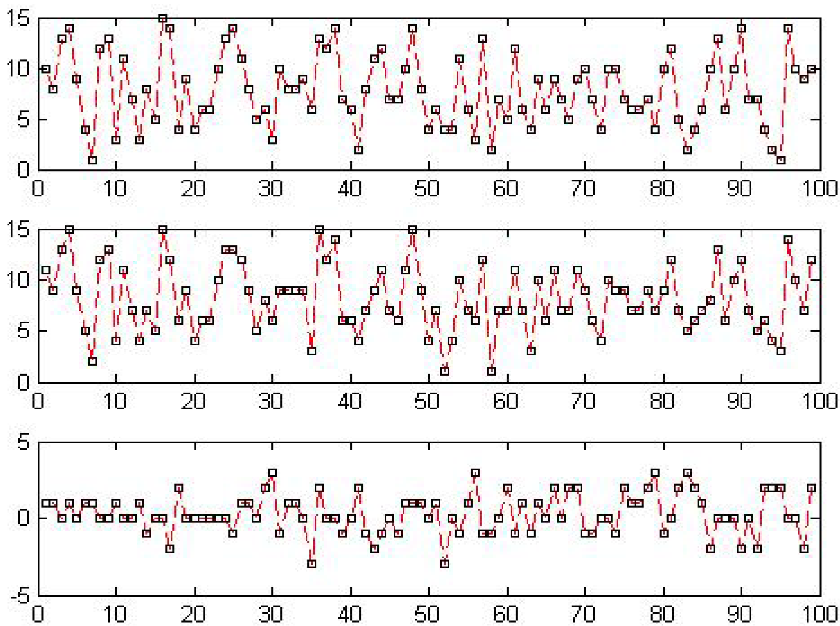

Fig. 6. True value, predicted value and their gap for *rank* based on dynamic model

Table 2. Comparison of rank prediction of two models

Deviation statistics	Static	Dynamic
Predicted mean	0.30303	0.213131
Predicted standard error	1.48086	0.91434

in Fig.3. Notice that during the prediction the keyword remains thoroughly coincident while the searching times in the meantime obey normal distribution. So the prediction error can be further dwindled once the information about searching times of a specific keyword acquired.

5 Conclusion

The importance and attractiveness of online marketing have been realized and fulfilled for couple of years. As the frontier of the practice sponsored link provides effective and efficient channel for business to approach customers. The fact that billions of dollars swarm into this market facilitates the prevalence of keywords auction as well as stimulates the prosperity of Internet economy. However myriad advertisers relapsed into an adverse state because of ignorance to bidding strategies and competitors behavior. This paper developed two models based on BN to help advertisers predict such crucial variants as *bid, rank, CT* and *impression* in online keyword auction. The results indicate that *CT* in the previous round does influence *rank* in the current round and hence dynamic bayesian model exceeds the static model in terms of prediction precision. The contributions of this paper lie in the introduction of evolutionary prediction from static to dynamic model by incorporating both knowledge on history bidding and impact of previous click and meanwhile the empirical study by aid of data from the largest online travel agency in China.

Definitely this research can be enhanced from two aspects. One is the inclusion of more variants to dynamic model. Since the history values of *bid, rank* and *impression* instead of only *CT* all may influence the *rank* prediction of current round the model can evolve to be more complete by assigning more parent nodes to rank. Another improvement results from the attribution of keywords. Travel and ticket are the examples of perishable goods which have obvious expiration so their inventory level will influence their bidding while it may escape such issue in the case of imperishable industrial goods. So managers may tailor different bidding strategies for different goods. Summarily we expect further research to underpin those aspects above.

References

1. Animesh, A., Ramachandran, V., Viswanathan, S.: Online Advertisers Bidding Strategies for Search, Experience, and Credence Goods: An Empirical Investigation. Working paper (2005)
2. Animesh, A., Ramachandran, V., Viswanathan, S.: Quality Uncertainty and Adverse Selection In Sponsored Search Markets, working paper (2005)
3. Edelman, B., Ostrovsky, M.: Strategic Bidder Behavior in Sponsored Search Auctions. Int. J. Internet Marketing and Advertising 2, 92–108 (2005)
4. Kitts, B., Leblanc, B.: Optimal Bidding on Keyword Auctions. Electronic Markets, special issue: innovative auction markets 14, 186–201 (2004)
5. Feng, J., Bhargava, H.K., Pennock, D.M.: Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms. working paper (2005)
6. Park, S., Durfee, E.H., Birmingham, W.P.: Use of Markov Chains to Design an Agent Bidding Strategy for Continuous Double Auctions. Journal of Artificial Intelligence Research 22, 175–214 (2004)
7. Park, Y.-H., Bradlow, E.T.: An Integrated Model for Bidding Behavior in Internet Auctions: Whether, Who, When, and How Much. Journal of Marketing Research XLII, 470–482 (2005)
8. Jensen, F.V.: Bayesian Networks and Decision Graphs. Springer, Heidelberg (2001)
9. Murphy, K.: Bayes Net Toolbox for Matlab (Accessed November 15, 2007) (2007), `http://www.cs.ubc.ca/~murphyk/Software/BNT/bnt.html`
10. Kitts, B., Leblanc, B.: Optimal Bidding on Keyword Auctions. Electronic Markets 14, 186–201 (2004)