TRANSIENT SIMULATION OF BALANCED BIDDING IN KEYWORD AUCTIONS

Maurizio Naldi[a], Giuseppe D'Acquisto[b]

[a]Università di Roma "Tor Vergata", Rome, Italy
[b]Università di Roma "Tor Vergata", Rome, Italy

[a]naldi@disp.uniroma2.it, [b]giuseppe.dacquisto@inwind.it

EXTENDED ABSTRACT
The synchronous version of the Balanced Bidding strategy for keyword auctions is examined through simulation. In the case of a uniform distribution for advertisers' valuation of clicks, advertisers prefer the highest positioned slots. The slot assignment process matches advertisers' expectations with a probability decaying with the number of slots according to a power law and depending negligibly on the number of advertisers.

Keywords: Search engines, Auctions, Keywords, Balanced Bidding

1. INTRODUCTION

Search engines, such as Google or Yahoo!, present users with a set of hyperlinks in response to their queries. In addition to the links deemed relevant to the query by the search engine (often named organic links), a number of sponsored links are presented as well, see Battelle (2005), associated to the query through the keywords specified in it. Advertisers are willing to pay for their ads to appear on the seach engine's response. Such sponsored links are then generally assigned through auctions, and the resulting revenues represent a significant source of income for search engines, see Edelman et alii (2007). On the other hand advertisers are interested inmaking their advertising strategy as efficient as possible. Such auctions have been studied for some time now in the context of game theory, where the players in the game are the auctioneer (the search engine) and the advertisers (see e.g. Varian (2007)). The related studies have been devoted mainly to examining if, and under which conditions, the game exhibits a Nash equilibrium.

Such conditions are typically linked to the assignment and pricing rule on one side and to the advertisers' bidding strategies on the other side. As to the former issue the Generalized Second Price (GSP) rule has reached a wide consensus, which however leaves the field open as to the bidding strategy for the advertiser.

Recently the Balanced Bidding (BB) strategy has been proposed by Cary et alii (2008), where advertisers update their bid at each auction round by exploiting the intelligence gathered in the previous rounds. In this process each advertiser identifies at each round his optimal slot as that maximing his utility. In Cary et alii (2008) the convergence to a Nash equilibrium has been studied, but the optimal slot determination process has not been explored in detail, though its relationship with the subsequent slot assignment is central to the advertiser's satisfaction. In this paper we analyse the characteristics of preferences for slots observed for advertisers as resulting from the repetition of such keyword auctions, when advertisers follow the BB strategy. In particular we examine the way advertisers distribute their preferences among the slots on auction, and how the auction's results match their expectations as to the assigned slot.

2. KEYWORDS AUCTIONS

Search engines act in response to users' queries for websites containing the information of interest. In such queries the information of interest is synthetically expressed as a string of keywords, possibly connected through Boolean operators. For the example, in response to the query "*sea* AND *winds* NOT *ice*" the search engine will return pointers to all the documents containing the first two terms but not the third one. The hyperlinks returned by the search engine are typically named organic links. The search engine can add to this list (and show e.g. on the right-hand side of the screen) a number of sponsored links. The available positions for sponsored links are named slots. Such links are provided by advertisers, who are willing to pay to have their ad appear on the screen in relation to a query containing a specific keyword. Hence for any query there are a number of potential fillers of the screen space devoted to sponsored links. It is assumed that the advertisers choose to run for keywords that are actually related to their product. The payment rules may be freely defined in the contract relationship between the search engine manager and the advertiser, but the most established agreement follows the pay-per-click model, where the advertiser pays a pre-determined amount of money each time the user actually clicks on the ad. Since the number of slots is generally smaller than the number of interested advertisers (i.e. advertisers who have opted to run for a keyword appearing in the query), slots represent a scarce resource and a natural way to assign them to the advertisers is through auctions, namely keywords auctions. Hence advertisers declare how much they are willing to pay for a click, and an auction is run for the slots among the advertisers whose keywords match the query. We have therefore a number of slots $S \in Z^{+}$ and a larger number of advertisers $A \in Z^{+}$, with $A < S$. Actually, a new auction is run every time a query is submitted, among the advertisers submitting bids for keywords matching the query. For any given keyword we have then a sequence of repeated auctions. As will be seen in Section 4, the repetition of auctions allows advertisers to update their bids by taking into account their past observations of other bidders' behaviour and of the output of previous auction runs.

In order to make the assignment process as effective as possible the auctioneer has to carefully design the auctioning rules, which boils down to choosing: a) the assignment rule (i.e. the way advertisers are assigned the slots); b) the price setting rule (i.e. the price an advertiser has to pay when the user clicks on its ad).

As to the first issue this is unanimously solved by using a straightforward ordering of slots and advertisers. Slots are indexed progressively by their vertical position on the screen (the slot appearing on top of the screen is assigned index 1 by convention; the slot appearing on the bottom of the ad-devoted space has index *S*) and evaluated by their click-through rate. The click-through rate θ_i of slot *i* is the probability that the user clicks on that slot. Its estimate can be obtained by dividing the number of users who clicked on an ad on a web page by the number of times the ad was delivered (impressions), see Sherman and Deighton (2001). It is generally accepted that the click-though rate is a declining function of the slot's position, i.e. $\theta_i > \theta_{i+1}$, when $i = 1,\ldots,S-1$; a statistical study reported in Brooks (2004) supports this assumption. Hence, top-positioned slots are more valuable than bottom-positioned slots. As to the precise shape of the click-through rate decaying function, we consider a Zipf distribution for the probability that the user clicks on a given slot. Namely the probability that the user clicks on the slot *j* is

$$\theta_j \propto \frac{1}{j^{\alpha}}, \qquad (1)$$

where $\alpha \in R^{+}$ is the Zipf parameter.

For convenience (with no consequence on the following results) we adopt the normalizing condition $\sum_{j=1}^{S} \theta_j = 1$, so that we are actually considering the probability of clicking on a specific slot conditioned to the user clicking on a slot (or, alternatively, the user clicks on a slot with probability 1). Though in this paper we implicitly consider the click-through rate being a function of the slot's position only, other authors have considered the more general case of click-through rates being function of advertisers as well , see Feldman and Muthukrishnan (2008) and Aggarwal et alii (2006). Advertisers are likewise ordered in a decreasing function by the value of their bid.

If we now denote by b_i the bid submitted by the *i*-th advertiser, and then by $b_{(j)}$ the *j*-th highest bid, the assignment rule states that the *k*-th slot is assigned to the advertiser submitting the bid $b_{(k)}$. For convenience we introduce the function $\Pi(k)$ returning the index of the advertiser who's assigned the *k*-th slot.

Setting the price is a less straightforward matter. A well-known mechanism is the truthful Vickrey-Clarke-Groves rule, which would lead each participating advertiser to bid its true valuation, see Clarke (1971) and Groves (1973). However, search engines do not adopt the VCG mechanism in practice, but rather the Generalized Second Price (GSP) rule, which is described in detail in Section 3.

3. GENERALIZED SECOND PRICE MECHANISM

In this Section we review the basic characteristics of GSP as a price setting mechanism. In GSP the natural assignment rule is maintained whereby the advertiser submitting the *k*-th highest bid $b_{(k)}$ is assigned the *k*-th slot. However the price he pays is equal to the next lower bid, i.e. $p_{(k)} = b_{(k+1)}$. Advertisers who are not assigned a slot pay nothing. The most important decision advertisers have to take is then to choose their bids. As a reference they have their own private valuation of clicks: in the simplest scenario the *i*-th advertiser values a click worth v_i (i.e. the click value doesn't depend on the slot position itself and doesn't vary as the auction is repeated). In general any bid of the generic *i*-th advertiser will satisfy the inequality $b_i \le v_i$. The expected utility of the advertiser receiving the the *k*-th slot is $\theta_k \left(v_{\Pi(k)} - b_{(k)} \right)$. The most important property of the VCG mechanism is that it induces the advertiser to declare its private valuation, so that $b_i = v_i$ (truthfulness property). On the contrary, GSP is not truthful, hence advertisers' bids are limited by the above inequality only. If we consider the static game associated to GSP-driven auction, a Nash equilibrium has been shown to exist, see Varian (2007). However, in the dynamic version resulting from the repetition of the auction, bidders can update their bid at each new issue of the auction by taking advantage of the knowledge they have gained from the past auction occurrences (the bids submitted by all the other bidders).

4. BALANCED BIDDING STRATEGY

If the advertiser dosn't submit a truthful bid, he has full freedom to choose for his bid any value satisfying the inequality recalled in Section 3. Cary et alii have proposed the Balanced Bidding (BB) strategy, see Cary et alii (2008), under which the advertiser chooses his next bid b so as to be indifferent between successfully winning the targeted slot k at the price p_k at which it was awarded at the previous auction run, or winning the slightly more desirable slot *k-1* at price b. In this context a set of A advertisers, who compete for S slots and have their private valuations $\{v_1, v_2, \ldots, v_S\}$ for a click, are assigned the S slots according to the GSP rule. The resulting BB strategy leads the winning i-th advertiser to:

1. targeting the slot k_i^* that maximizes its utility (optimal slot), i.e. $k_i^* = \arg\max_k \{\theta_k (v_i - p_k)\}$;
2. setting its next bid b according to the expression $b_i = v_i - \dfrac{\theta_{k_i}}{\theta_{k_{i-1}}}\left[v_i - p_{k_i^*}\right]$.

Losing advertisers' bids instead equal their valuations.

For the asynchronous version of BB (advertisers update their bids one at a time) Cary et alii have proved that the dynamic system where all bidders play this strategy converge to a unique fixed point, which is also the Nash equilibrium of the static game, see Cary et alii (2008). However, the convergence time depends on the number of bidders and may take some hundreds of auction runs.

In this paper we focus on the transient behaviour of the auction in the synchronous case, i.e. under the hypothesis that each bidder plays the Balanced Bidding strategy at each round.

5. MEASURES OF AUCTION'S SUCCESS

In order to evaluate the characteristics of the auction a number of metrics can be considered. In Naldi and D'Acquisto (2008) some have been proposed to reflect the interest of the bidders as well as that of the auctioneer. Since in the Balanced Bidding strategy bidders submit their bids after a process identifying the optimal slot, it is natural to consider their bids as referred to a specific slot (though the slot is not mentioned explicitly in the submission process). If we consider auctions as a matching process between the objects on auctions (the slots) and the bids, a measure of success is naturally given by the probability that a slot is assigned to the bidder for which that slot is optimal. We name such probability the slot matching probability. If we indicate by Y_i the optimal slot of the bidder who is assigned the slot i, the formal representation of such probability is then $\omega(i) = P[Y_i = i]$ for a generic i, or, if we lose the details of the specific slot, the average

$$\Omega = \frac{1}{S}\sum_{i=1}^{S}\omega(i). \tag{2}$$

If all the users were to designate the same slot as their optimal one, the slot matching probability would be equal to the inverse of the number of slots. Though the advertisers' preferences are not so unanimous, they are far from being uniformly distributed, so that the same slot may be regarded as optimal by multiple advertisers while some other is not even for a single advertiser. Another issue of interest is then the distribution of such preferences, i.e. the values of $\psi(i,j) = P[Y_i = j]$ when $j = 1,2,\ldots,S$ and $i = 1,2,\ldots,A$. Here, similarly, we can consider just the average over the set Λ of winning advertisers

$$\Psi(j) = \frac{1}{S}\sum_{i\in\Lambda}\psi(i,j). \tag{3}$$

In this paper we focus on these two measures: the slot matching probability and the distribution of preferences. We resort to MonteCarlo simulation, by running N_{sim} times a simulation cycle consisting of T repetitions of the auction for S slots. In Cary et alii (2008), where the Balanced Bidding strategy has been proposed, the asynchronous version of that strategy is considered, where a single advertiser (randomly chosen) is given the chance to update his bid at a time. This leads to a quite slow convergence towards the VCG results: roughly 200 rounds are needed for the players' payoffs to converge. Here we instead take the much more realistic assumption that all advertisers actually update their bid at each round. Such asumption is expected to reduce considerably the convergence time. Our study will therefore concentrate on the transient behaviour of the auction, i.e. that pertaining to the first batch of repetitions.

For evaluation purposes we have to set some working hypotheses, in particular concerning the advertisers' valuations and the click-through rate (for which we assume the Zipf distribution). A relevant role in the auction's outcome depends on the distribution of bidders' valuations (which remains unchanged during the subsequent repetitions of the auction). In Naldi and D'Acquisto (2008) a number of distributions were considered: uniform, triangular, Gaussian, exponential, and Pareto. They can be roughly divided into two sets, respectively comprising those showing a large dispersion of bids and those where bids cluster around a common value. In the present early study we consider just the uniform distribution as a representative of the first category. The valuations are then represented by the i.i.d. random variables $V_1, \ldots, V_A$ following a standard uniform distribution, so that the coefficient of variation (standard deviation-to-expected value ratio) is $1/\sqrt{3}$.

6. DISTRIBUTION OF PREFERENCES

As stated in the previous section we first consider the distribution of preferences, i.e. the probability that a given slot position is deemed as optimal by the advertisers. Since the definition of optimality is based on the evaluation of the utility, which is positive only if the advertiser's private valuation is larger than the current price for that slot, not all advertisers declare an optimal slot. Hence the distribution is evaluated for the restricted group of addvertisers who do exhibit an optimal slot. The evaluation is conducted by simulation. We consider N_{sim} MonteCarlo simulation runs, each consisting of T repetitions of the auction for S slots. Within each run any auction is conducted according to the GSP mechanism, where each advertiser adopts the Balanced Bidding strategy described in Section 4. At the end of the full set of MonteCarlo simulation runs we can estimate the distribution of preferences $\Psi(j)$ as a function of the slot's position $j = 1,\ldots,S$. The estimator's expression is

$$\widehat{\Psi}(j) = \frac{1}{N_{sim} \cdot T \cdot S} \sum_{i=1}^{N_{sim}} \sum_{t=1}^{T} \sum_{k=1}^{S} I_{\left[x_k^{(t,i)} = j\right]}, \quad (4)$$

where $I_{[\bullet]}$ is the indicator function (equal to 1 if its logical argument is satisfied and zero otherwise) and $x_k^{(t,i)}$ is the optimal slot (as resulting after the auction repetition t) of the bidder who is assigned the slot k at the repetition t in the i-th simulation run. Here we report some results obtained under the following conditions:

- Number of simulations runs $N_{sim} = 10000$
- Number of slots $S = 5,10$
- Number of advertisers $A = S + 1$
- Number of repetitions $T = 100$
- Uniform distribution of advertisers' private valuations
- Zipf distribution for the probability of users clicking on a given slot, with the Zipf parameter $\alpha = 0.5,1,2$

We briefly review these assumptions. The number of simulation runs is large enough to allow for an excellent accuracy for the values at hand. The relative standard error of this crude MonteCarlo estimator $\widehat{\Psi}(j)$ is in fact

$$\varepsilon = \sqrt{\frac{1 - \Psi(j)}{N_{sim} \cdot T \cdot S \cdot \Psi(j)}}. \quad (5)$$

As to the number of slots, the size of the screen space available for sponsored links coupled with the visibility requirements can hardly allow for more than 10 such links. When we come to the assumption on the number of advertisers, actually we could imagine a number much larger than the number of slots. However, by choosing A as the minimum integer larger than the number of slots, we reduce to a minimum the computational load, while obtaining a final result quite accurate also for larger values of A (as briefly shown later). In Figure 1 and Figure 2 we draw the distribution of preferences when the number of slots is 5 and 10 respectively. As expected the figures for the 10 slots case are generally lower than the 5 slots case, since preferences distribute among a larger number of potentially optimal slots. In both cases we can however note that the highest slot is the most preferred one, but at the same time the preference doesn't decay monotonically with the slot position. In fact, in some cases we see a slight upsurge of the preference probability for the lowest slot. Counter-intuitive is also the impact of the Zipf parameter: Highly skewed click-through rate distributions (i.e. having larger values of the Zipf parameter) produce more balanced distributions of preferences.

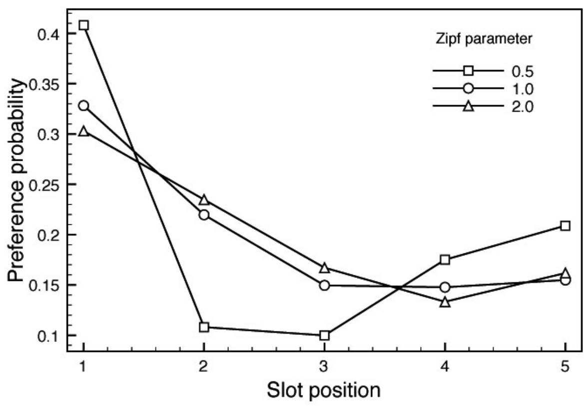

Figure 1: Probability of Advertisers' Preferences with 5 Slots

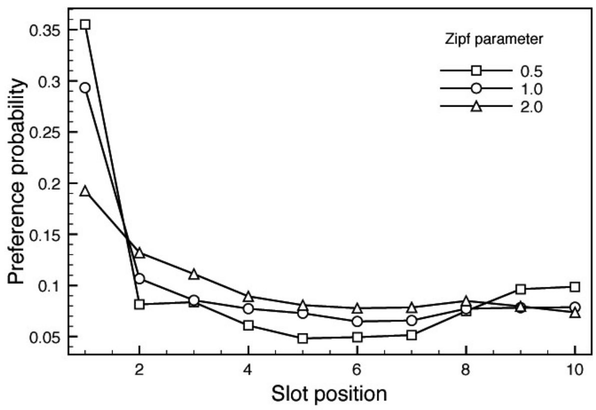

Figure 2: Probability of Advertisers' Preferences with 10 Slots

Finally, we come to the assumptions on the number of advertisers. The results so far shown, obtained for $A = S + 1$, keep valid as long as the preference distribution depends negligibly on the number of advertisers. In Figure 3 we show the same distribution, obtained for $A = 6$ and $A = 50$, where multiplying the number of advertisers tenfold produces a very limited variation on the estimated preference probability, comparable to the accuracy of the simulation method. Hence the results obtained above can be deemed accurate enough for larger values of the number of advertisers as well.

7. SLOT MATCHING PROBABILITY

We now turn to the slot matching probability, which measures how well the slot assignment satisfies the advertisers' expectations, and can therefore be considered as a measure of success of the auction.

After running N_{sim} MonteCarlo simulation runs, each consisting of T repetitions of the auction for k slots, the slot matching probability is estimated as

$$\hat{\Omega} = \frac{1}{N_{sim} \cdot T \cdot S} \sum_{i=1}^{N_{sim}} \sum_{t=1}^{T} \sum_{k=1}^{S} I_{\left[x_k^{(t,i)} = k\right]}, \qquad (6)$$

where $I_{[\bullet]}$ is the indicator function (equal to 1 if its logical argument is satisfied and zero otherwise) and $x_k^{(t,i)}$ is the optimal slot (as resulting after the auction repetition t) of the bidder who is assigned the slot k at the repetition t in the i-th simulation run.

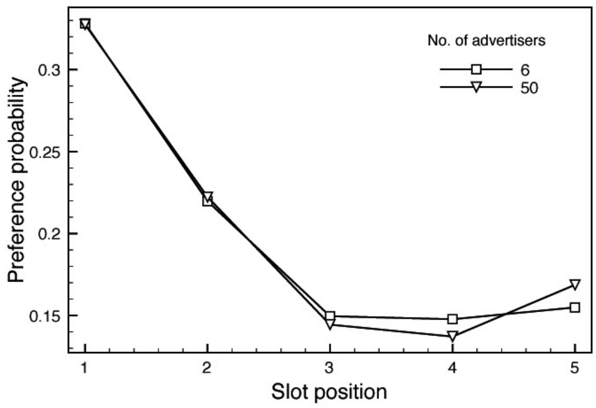

Figure 3: Impact of the Number of Advertisers on Preference Distribution

We report here the results obtained under the following conditions:

- Number of simulations runs $N_{sim} = 10000$
- Number of slots $S = 2,10$
- Number of advertisers $A = S + 1$
- Number of repetitions $T = 100$
- Uniform distribution of advertisers' private valuations
- Zipf distribution for the probability of users clicking on a given slot, with the Zipf parameter $\alpha = 0.5,1,2$

In Figure 4 the slot matching probability appears as a fast decaying function of the number of slots. Matching appears to be rarer for lower values of the Zipf parameter (i.e. as the click rate becomes more uniform over the set of slots).

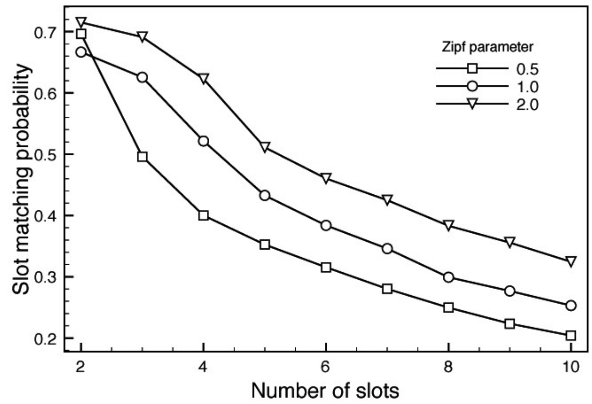

Figure 4: Slot Matching Probability

A tentative fitting can be considered by using the power law model

$$\Omega = \frac{\gamma}{S^{\beta}}, \qquad (7)$$

where γ and β are two constants. For the three values of the Zipf parameter considered in this paper we obtain by regression the values reported in Table 1 along with the resulting R^2 goodness of fit index. We note that, though the fit is generally good, the power law exponent is not a monotonic function of the Zipf parameter.

Table 1: Fitted Power Law Parameters

α	γ	β	R^2
0.5	1.1147	0.7237	0.9935
1	1.4683	0.7574	0.9969
2	1.4417	0.6378	0.9915

In order to analyse the transient dynamics we have also evaluated the slot matching probability through a sliding window of width equal to 20 auction rounds over a total length of $T = 500$ repetitions. The quantity of interest is now a function of the ending time T_{end} of the sliding window

$$\hat{\Omega}(T_{end}) = \frac{1}{N_{sim} \cdot 20 \cdot k} \sum_{i=1}^{N_{sim}} \sum_{t=T_{end}-20}^{T_{end}} \sum_{j=1}^{k} I_{\left[y_j^{(t,i)} = j\right]}, \qquad (8)$$

where $T_{end} = 21,\ldots,500$.

A rough analysis of the length of the transient can be obtained by visual inspection of the resulting slot matching probability. In Figure 5 (obtained for $S = 5$ slots and for the Zipf parameter $\alpha = 1$) the curve stabilizes well before the first 40 rounds, much earlier than the convergence time reported in Cary et alii (2004) for the asynchronous version of the auction.

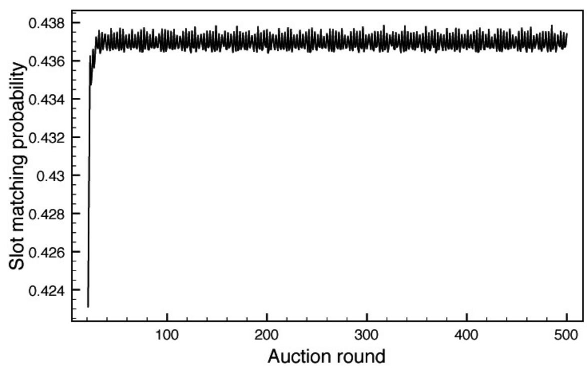

Figure 5: Transient in Slot Matching Probability

8. CONCLUSIONS

In this paper we have examined some characteristics of the Balanced Bidding strategy applied to repeated auctions for keywords. We have considered the synchronous version of that strategy (all the bidders update their bids at each auction round). As to the performance indices we have focussed on the probability that a slot is assigned to the advertiser for which it is optimal, which can be considered as a measure of success of the advertiser's bidding strategy. The slot matching probability depends negligibly on the number of bidders, while decays fast with the number of slots on sale. For that relationship we provide a tentative power law fitting, where however the power law exponent doesn't vary monotonically with the Zipf parameter. We have also examined the way bidders' preferences distribute among the available slots. The highest slots are the most preferred ones, but the lowest slots may exhibit a preference upsurge. Such phenomenon, which makes the preference distribution non monotonic, is more prominent the lower the Zipf parameter.

REFERENCES

Aggarwal, G. , Goel, A., Motwani, R., 2006. Truthful auctions for pricing search keywords. *EC '06: Proceedings of the 7th ACM conference on Electronic commerce*, 1–7, 11-15 June 2006, New York, NY, USA.

Battelle, J., 2005. *The Search*. Portfolio Hardcover.

Brooks, N., 2004. The Atlas Rank Report: HowSearch Engine Rank Impacts Traffic. *Technical report, Atlas Institute*.

Cary, M. , Das, A., Edelman, B., Giotis, I., Heimerl, K., Karlin, A.R., Mathieu, C., Schwarz, M., 2008. On Best-Response Bidding in GSP Auctions. *NBER Working Paper No. W13788*.

Clarke, E.H., 1971. Multipart pricing of public goods. *Public Choice*, 11(1):17–33.

Edelman, B. , Ostrovsky, M. , Schwarz, M.,2007. Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords. *The American Economic Review*, 97(1):242–259.

Feldman, J. , Muthukrishnan, S., 2008. Algorithmic methods for sponsored search advertising. *arXiv Preprint 0805.1759*.

Groves, T., 1973. Incentives in Teams. *Econometrica*, 41(4):617–31.

Naldi, M., D'Acquisto, G., 2008. Performance of the Vickrey auction for digital goods under various bid distributions. *Performance Evaluation*, 65:10–31.

Sherman, L. , Deighton, J., 2001. Banner advertising: Measuring effectiveness and optimizing placement. *Journal of Interactive Marketing*, 15(2):60–64.

Varian, H., 2007. Position auctions. *International Journal of Industrial Organization*, 25:1163–1178.

AUTHORS BIOGRAPHY

Maurizio Naldi graduated cum laude in 1988 in Electronic Engineering at the University of Palermo and then received his Ph.D. in Telecommunications Engineering from the University of Rome “Tor Vergata”. After graduation he pursued an industrial career, first at Selenia as a radar designer (1989–1991), and then in the Network Planning Departments of Italcable (1991–1994), Telecom Italia (1995–1998), and WIND (1998–2000) where he was appointed Head, Traffic Forecasting & Network Cost Evaluation Group. In the 1992–2000 period he was active in the standardization bodies (ETSI and ITU), in particular as Associate Rapporteur for Broadband Traffic Measurements and Models at ITU Study Group 2. Since 2000 he is with the University of Rome at Tor Vergata, where he is now Aggregate Professor.

Giuseppe D'Acquisto graduated cum laude in 1995 in Electronic Engineering and received the PhD in Telecommunications in 1999 from the University of Palermo-Italy, with a thesis on rare event simulation. After PhD he started a consulting career, working for Telco Operators and ICT companies, with a focus on market forecasting, traffic engineering and cost accounting. He collaborates with the Universities of Rome Tor Vergata and Palermo in researches in the area of simulations and stochastic optimization. He is the author of more than 20 publicatons on these topics.