Implementing Sponsored Search in Web Search Engines: Computational Evaluation of Alternative Mechanisms

Juan Feng
University of Florida, Warrington College of Business, Gainesville, Florida 32611, jane.feng@cba.ufl.edu

Hemant K. Bhargava
University of California Davis, Graduate School of Management, Davis, California 95616, bhargava@computer.org

David M. Pennock
Yahoo! Research Labs, Pasadena, California, 91103, pennockd@yahoo-inc.com

The practice of *sponsored search* advertising—where advertisers pay a fee to appear alongside particular Web search results—is now one of the largest and fastest growing source of revenue for Web search engines. We model and compare several mechanisms for allocating sponsored slots, including stylized versions of those used by `Overture` and `Google`, the two biggest brokers of sponsored search. The performance of these mechanisms depends on the degree of correlation between providers' willingness to pay and their relevance to the search term. Ranking providers based on the product of relevance and bid price performs well and is robust across varying degrees of correlation. Ranking purely based on bid price fares nearly as well when bids and relevance are positively correlated (the expected regime), and is further enhanced by adding an editorial filter. Regardless of the allocation mechanism, sponsored search revenues are lower when users' attention decays quickly at lower ranks, emphasizing the need to develop better user interfaces and control features. The search engine can address initial inscience of relevance scores by modifying rank allocations over time as it observes clickthroughs at each rank. We propose a rank-revision strategy that weights clicks on lower ranked items more than clicks on higher ranked items. This method is shown to converge to the optimal (maximum revenue) ordering faster and more consistently than other methods.

Key words: internet search engines; preferential placement; sponsored search; slot allocation; auctions

History: Accepted by Amit Basu, Area Editor for Knowledge and Data Management; received December 2002; revised January 2004, October 2004; accepted January 2005.

---

1. Introduction

Internet search engines index billions of Web pages and employ information retrieval algorithms to display, in response to a user's query, links to Web pages deemed relevant to the query. These pages might represent commercial firms selling goods or services, information sites, government entities, etc. Lawrence and Giles (1999) estimated the publicly indexable Web at 800 million pages, containing 6 terabytes of text data on 2.8 million servers; today, Internet statistics indicate that the publicly indexable Web contains billions of pages, served by over 40 million Web servers. Due to this vast amount of information, search engines act as an information gateway to many search and decision-making tasks. More than 50% of Web users visit a search engine every few days, the leading search engine (`Google`) gets over 250 million search requests each day, over 13% of traffic to commercial sites was generated by search engines, and over 40% of product searches on the Web were initiated via search engines. For us, the term *search engine* encompasses various applications of these indexing-retrieval technologies, including pure Web search engines (e.g., `Google`), information portals with search functionality (e.g., `Yahoo!`), metasearch engines (e.g., `Metacrawler`), niche search engines (e.g., `CiteSeer`), and comparison shopping engines (e.g., `mySimon`, `Shopping.com`).

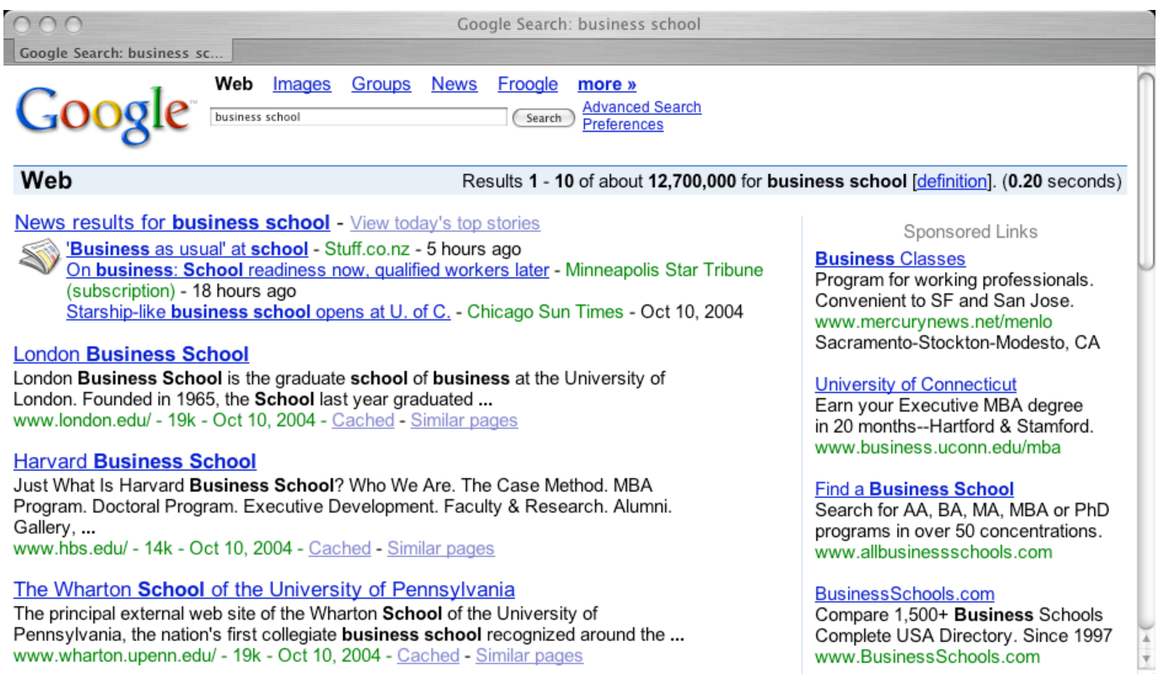


Figure 1: Sponsored and algorithmic search results in Google.

Due to the critical influence of search engines on Web users' actions, many commercial firms have realized the importance of gaining a high position on the search results for specific queries. Entire niche industries exist touting services to boost a Web page's ranking on the popular search engines, in part by reverse engineering the search engines' information retrieval algorithms. The expectation of increased traffic from good placement on a search page has led to the creation of a market for *sponsored search* (or *paid placement*—we use the terms interchangeably, as dictated by context) where search engines can charge a fee for prominent positioning within a "sponsored" section in the results page. For example, a digital camera retailer may pay a search engine to appear among the sponsored results when users search for "digital cameras." Usually, sponsored results are shown on top of, or to the side of, the standard unpaid search results (also called *algorithmic* results). In general, sponsored listings are explicitly marked as sponsored results or advertising, and the FTC has advised search engines to disclose paid links appropriately, see Tantono et al. (2002). Figure 1 displays a screen shot with sponsored and algorithmic search results in an Internet search engine.

In recent years, sponsored search has become an important and fast-growing revenue source for Internet search engines. Total industry revenue increased from approximately US$0.9 billion in 2002 to about US$4.6 billion in the first half of 2004. Leading firms include `Google` and `Yahoo`'s `Overture` division. `Overture` (formerly `GoTo.com`, recently acquired by `Yahoo!`), is credited with pioneering paid placement advertising on the Internet, and acts as a broker between advertisers and information gatekeepers like `Yahoo!` and `MSN`. `Overture`'s success has prompted several other companies to adopt similar business models involving paid placement, most prominently `Google`, the leading Internet search engine. Other sponsored search providers include `LookSmart`, `FindWhat`, and `eSpotting`.

Section 2 presents an overview of management challenges related to sponsored search. Section 3 presents four alternative mechanisms by which a search engine might award sponsored slots and ranks. In Section 4, we report results of computational simulations of the equilibrium performance of these four mechanisms. Ranking by the product of bid price times clickthrough weakly dominates other mechanisms in all tested regimes, though the mechanism is statistically equal to ranking by bid price in the expected region of positive correlation between willingness to pay and relevance. Editorial filtering helps the basic rank-by-bid mechanism significantly. Section 5 studies two alternative approaches for the dynamic ranking problem, where the search engine periodically revises the ranking of listings within the paid slots based on its observations about user clickthroughs. Section 6 provides concluding remarks.

2. Problem Overview and Related Research

The study of sponsored search is crucial to understanding the future design, quality levels, and market structure of Web search engines. We elaborate on the research focus of this paper and summarize other management problems and research related to paid placement. We begin by discussing essential characteristics of sponsored search.

2.1. Characteristics of Sponsored Search

Advertising in traditional media (e.g., magazines and television) is typically sold on a per-impression basis, or according to the number of people exposed to the ad. Banner ads on the Web are also typically sold per impression. On the other hand, sponsored search is typically sold on a *per-click* basis. Advertisers pay only when a user clicks on their ad; in a sense they are paying for leads rather than exposure. Traditional advertising (and to a large extent banner advertising on the Web) is typically priced via an informal process of estimation and negotiation. The Internet, however, supports more efficient and mechanized pricing via real-time auctions that can capture the advertisers' true willingness to pay, and track it over time. The sponsor-search industry typically runs separate auctions for different search terms: for example, the search queries "plasma television" and "investment advice" are associated with two distinct auctions. What's being sold in each auction is the right to appear alongside the results for that search query (actually what's being sold is the right to the users' proceeding *clicks* if any). In practice, hundreds of thousands of advertisers compete for positions alongside several million search queries every day. So, for example, travel vendors like `Expedia` and `Orbitz` may compete in an auction for the right to appear alongside the result of a user's search on "Las Vegas travel." Generally auctions are dynamic, meaning that advertisers can change their bids at any time, and a new auction clears every time a user enters the search query. In this way, advertisers can adapt to changing environments, for example boosting their bids for "Harry Potter" around the release date of the latest book in the series.

2.2. Management Problems and Challenges

How should search engines award and charge for placement in the sponsored section of the search results page? The industry standard of charging per click rather than per impression makes the mechanism-design problem unusual. If the search engine instead auctioned off impressions, with each specific part of the results page sold in a separate auction, then lessons from traditional auction theory might apply. However, because the "commodity" being auctioned is a click rather than an impression, the price signal coming from advertisers leaves unanswered where on the page to display each impression.

We assume a specific portion of the results page is reserved for k sponsored results, and that the mechanism must choose which k listings to show and in what order. Let $R_{pp}(k)$ represent the revenue from sponsored search when the search engine allocates k paid slots. One approach is simplified to display the top k listings in descending order according to the advertiser's bid. Assuming an open-outcry format, this approach is transparent to advertisers, who can easily see at any time how much they need to bid to appear in any given position. Another approach is to factor in directly the likelihood of each ad being clicked (by assessing its relevance, by measuring past clickthroughs, or both). Since $R_{pp}(N)$ is the product of click price and click volume, factoring in click volume may improve revenue, though reduces transparency for bidders. Search engines also have incentives to refuse irrelevant or offensive ads: such ads may cause users to defect to competing search engines, ultimately reducing revenues. An important policy decision is how much to rely on automatic filtering, and how many resources to allocate to human editorial review of sponsored listings.

Myopically maximizing current-period revenue is not a good strategy for a search engine. All advertising revenues, including revenues from sponsored search, depend on the search engine's number of users (volume of traffic); moreover direct revenues from, for example, subscriptions to premium services, depend on the number of users. Let R_{users} denote revenue obtained directly from user subscriptions, plus advertising sold on a per-impression basis (e.g., typical banner ads). Then the search engine's total revenue is $R = R_{pp} + R_{users}$. The search engine faces a tradeoff: increasing k may increase R_{pp} in the short term, but may turn off users, thereby lowering total traffic and, in turn, lowering R_{users} and lowering R_{pp} in the longer term. Thus, besides the optimal ranking mechanism, the search engine must also choose the number of paid slots by finding the optimal tradeoff between sponsorship and user retention. Bhargava and Feng (2002) provide a theoretical model to explain and analyze this tradeoff. In this paper, although we verify the existence of an optimal k by explicitly modeling user retention, we focus more on the related problem of determining, given some fixed k, the relative revenue performance of alternative mechanisms for awarding slots.

Management of sponsored search is ultimately a game-theoretic balance among users, advertisers, and the search engine. And when sponsored search is managed by a broker, e.g., `Overture`'s relationship with `MSN`, `Google`'s relationship with `AOL`, and `LookSmart`'s relationship with `Lycos`—the broker is a fourth party to consider in the balance. The search engine's (or broker's) most direct goal is to maximize revenue. However, revenue is entirely dependent on keeping both advertisers and users from defecting to other search engines. So, the sponsored search mechanism design problem must simultaneously consider a number of factors, including direct revenue, utility for users, utility for advertisers, and, in the case of broker-affiliate relationships, utility for the affiliate.

Other management challenges include detecting and ignoring robot clicks and fraudulent clicks by people with malicious intent—for example a competing advertiser who wants to force costs onto their competitor, or an affiliate who actually benefits monetarily from additional clicks. Two sub-industries of sponsored search experiencing rapid growth are: (1) *local advertising*, where advertisers can target users in specific geographic regions, and (2) *contextual advertising*, where listings are placed beside news stories and other Web content, with the intent of marrying content with the most relevant ads.

2.3. Academic Research

The growth of paid placement has attracted recent research on this topic. Hoffman and Novak (2002) discuss the trend in Internet advertising towards per-click pricing rather than the traditional per-impression model. Asdemir et al. (2002) study the pricing choice between impressions and clicks, while Hu (2003) uses contract theory to show that performance-based pricing models can give the publisher proper incentives to improve the effectiveness of advertising campaigns. Weber and Zheng (2003) study the implementation of paid placement strategies, and find that the revenue-maximizing search engine design bases rankings on a weighted average of relative quality performance and bid amount. Rolland and Patterson (2003) propose a methodology using expert systems to improve the matching between advertisers and web users. Kumar et al. (2003) study the optimal advertising schedule to maximize the Web site's (search engine's) revenue under a hybrid pricing model (based on both the number of impressions of the ad and the number of clicks on the ad).

3. Ranking Mechanisms for Sponsored Search

The positive correlation between top placement and increased traffic creates significant demand among businesses for top placement on search engines, especially for popular and commercially-relevant search terms. However, since Web users face negative utility if the search engine becomes impartial, most search engines limit the number of paid placement requests they accept. Thus, the sponsored slots are a scarce resource that need to be allocated carefully. In this section, we describe the mechanics of sponsored search, develop a model of the search engine's placement revenues, and describe four alternative mechanisms for choosing the allocation of paid slots to advertisers. The "v ranking" first-price auction mechanism is inspired by `Overture`'s approach. In practice, Overture advertisers employ bid proxy agents that automatically bid 1 cent above their nearest competitor up to their maximum willingness to pay, effectively simulating a second price auction, but these two formats should produce the same equilibrium outcomes under the setup of this paper. The "$v \times \alpha$" mechanism is inspired by `Google`'s second-price auction that combines bid price with the listing's click history. The "posted price ranking" is a stylized version of how portals and other content destinations sell banner ads (and more like how traditional media sell advertising space). The "α ranking" focuses on relevance of a link, and so should be the preferred format from the perspective of search-engine users. The v and $v \times \alpha$ mechanisms are stylized versions of those run by `Overture` and `Google`, respectively, that we believe capture crucial aspects of the problem. However, several details are not fully addressed in our model, including query-matching algorithms, human editorial intervention, non-search-based advertising, new pricing models (e.g., pay per *conversion*), budgeting across time and across auctions, geographic targeting, marketing efforts, brand awareness, user interfaces, legal controls, strategic alliances, etc.

3.1. Revenue Model

Consider the search engine's placement revenue for a single search term. Suppose that s listing companies, interested in advertising for this term, compete for k paid slots on the search engine's results page. Let v_j be advertiser j's willingness to pay (WTP) per click for preferential placement on this term. Let α_j represent the "true" relevance score of listing j, encoding how useful the link is to users. We assume $0 < v_j, \alpha_j < 1$. Let $f(\alpha_j, v_j)$ denote the joint density function.

The expected revenue generated by an allocation is a function of the price of each slot and the clickthroughs generated at each slot. The clickthroughs that a listing generates depends on both its relevance and its rank within the sponsor-search section, because users are inherently more likely to click on higher-ranked items. To compute the expected number of clickthroughs for an item j at position i, our simulation employs an exponentially decaying attention model with factor $\delta > 1$. Formally, we compute the average clickthrough as α_j/δ^{i-1}. Exponential decay of attention is a fairly standard assumption, see Breese et al. (1998), that is borne out in practice: actual clickthrough data obtained from `Overture` during 2003 for the top five positions across all affiliates—including `Yahoo!`, `MSN`, and `AltaVista`—are fitted extremely well ($R^2 = 0.997$) by an exponential decay model with $\delta = 1.428$.

Define $r : I \rightarrow J$ to be the search engine's ranking function that allocates position i to listing company j. The set J also contains a fictitious null provider, since the search engine may not fill all slots. Let P_i represent the payment for position i. For the clickthrough-based mechanisms, let p_i represent the payment per clickthrough at position i, so that in equilibrium $P_i = p_i \frac{\alpha_j}{\delta^{i-1}}$ where $j = r(i)$.

In our model, the total traffic that the search engine attracts is a function of its overall quality, which is influenced by how the search engine allocates its sponsored slots. To be precise, the overall quality of the sponsored portion is the average relevance score of all the paid listings ($\sum_i \alpha_{r(i)}/k$). Search-engine users may have different sensitivities toward sponsored results under different search scenarios (for example, users generally treat commercial information search differently from non-commercial information search). To model the sensitivity, we introduce a market sensitivity factor λ (≥ 1, higher λ leads to a greater reduction in demand). For a given λ, we write the aggregate user traffic attracted by the search engine as $(\sum_i \alpha_{r(i)}/k)^\lambda$.

Hence the search engine's placement revenue is:

$$\left(\frac{\sum_i \alpha_{r(i)}}{k}\right)^\lambda \cdot E[(\sum_i^k P_i)] = \left(\frac{\sum_i \alpha_{r(i)}}{k}\right)^\lambda \cdot \sum_i^k \int_0^1 \int_0^1 P_i f(\alpha_{r(i)}, v_{r(i)}) d\alpha_{r(i)} dv_{r(i)} \tag{1}$$

where k is the number of paid listings that the search engine decides to display out of the s total bidders.

3.2. Allocation Mechanisms

There are several different approaches for allocating and pricing sponsored slots. `Yahoo`'s `Overture` division screens listings according to both automated and manual editorial policies to control for relevancy, then ranks all qualified advertisers according to how much they are willing to pay per click. For any given query, paid slots are allocated to the top k bidders in order of their bids. Listings may be shut down if they do not achieve sufficient clickthroughs. `Google`'s `AdWords Select` sponsored search program relies largely on automated editorial filtering. `Google` ranks qualified listings according to the product of the advertiser's bid, times the actual clickthrough of the listing. Since advertiser payments are per click, boosting higher click-rate listings may help improve overall revenue. Some search engines post a (reserve) price for paid slots and award the slots to the highest bidders.

3.2.1. v Ranking: Highest Payers at the Top

This mechanism allocates slots based on the listing company's willingness to pay. For a certain keyword, each listing company j makes a bid B_j for payment per click. The listings associated with the highest k bids are displayed, ranked according to their bids. We assume that companies pay what they bid (first-price auction). Due to the *per-click* payment format of paid placement, each listing firm's bid is independent of which slot it might be allocated (which influences its expected clickthroughs), hence we can use standard auction theory to derive the equilibrium bidding strategy. Therefore B_j is the expected highest value among the remaining $n-1$ bidders given that this value is below v_j. Hence, the firm with i^{th} highest value makes the i^{th} highest bid. The highest k bidders win and are assigned positions according to their bids, so under this mechanism $r(i)$ is the firm with i^{th} highest valuation. The search engine realizes a price per click $v_{r_{i+1}}$ for slot i.

This mechanism is meant as a stylized version of `Overture`'s approach. Note however that our formulation ignores many factors present in the commercially deployed mechanism, including editorial control, bid proxy agents, inexact query matching, budget constraints, and less-than-perfect correspondence between relevance and clickthrough rate, among other factors. To address this gap, Section 4.6 extends the analysis to account for the editorial control observed in practice.

3.2.2. $v \times \alpha$ Ranking: Relevance and Bid Price Jointly Determine Rank

Again, every listing company j bids B_j per click. We compute the product $B_j\alpha_j$ (where α_j approximates the expected number of clickthroughs for the listing). The listings associated with the highest k products are displayed, ranked according to this product. The actual price paid by each winner is a variant of the standard second-price auction, and computed as follows. Let $r(i)$ be the firm that wins slot i. Then, if $B_{r(i)} > B_{r(i+1)}$ then firm $r(i)$ pays $B_{r(i+1)}$; otherwise, it pays the least amount $\hat{B}_i$ such that $\hat{B}_i \alpha_{r(i)} \geq B_{r(i+1)} \alpha_{r(i+1)}$. In other words, the winning firm pays either the bid just below it, or an amount such that its total predicted payment just exceeds that of firm $r(i+1)$. Even though this is no longer a standard second-price auction, we make the simplification that the listing firms bid their true value, which is reasonable since the firm does not pay its own winning bid. Therefore, the price paid by the firm in slot i is either $p_i = v_{r(i+1)}$ or $p_i = \frac{v_{r(i+1)} \cdot \alpha_{r(i+1)}}{\alpha_{r(i)}}$.

3.2.3. α Ranking

This mechanism selects the highest k bids and ranks the bidders by their expected number of clickthroughs. The sponsored slots are assigned according to this ranking, and all winners pay the highest rejected bid. Consequently, every listing company bids its true willingness to pay (per click) v_j. Thus this is a generalized version of second price auction. Hence $p_i = v_{r(k+1)}$ is the $(k+1)^{th}$ highest valuation.

3.2.4. Posted-Price Mechanism

In this mechanism, the search engine sets a reserve price for each of the positions it has for a certain period. Each listing company submits its bid to be listed for that period. The highest k bids are the potential winners, except that the search engine may fill fewer than k slots if certain of the winners do not meet the reserve-price constraints, as explained below. For each i if the ith highest bid is higher than the reserve price for the ith position, the listing company who submits this bid is admitted for the ith position and pays the reserve price; otherwise, it is compared to the $(i+1)^{th}$ reserve price. If it is higher than the $(i+1)^{th}$ reserve price, it is admitted and pays the $(i+1)^{th}$ reserve price, but this way the total number of sponsored positions will also be reduced by 1, and so on.

In setting up the auction, the search engine determines the k reserve prices by computing the expected revenue potential for each of the k positions. In our simulation, for every sample correlation (from -1 to 1) we generate s pairs of α and v, calculate the average order statistics for the product αv over 100 runs, and use this as the reserve price. Note that this procedure imposes a quality control for the winner-determination rule: those listings with higher relevance score are more likely to have a greater revenue potential (because they are most likely to generate traffic), so they are the ones most able to pay the reserve price.

Table 1: An Example of Ranking and Payment Schedule for the Four Mechanisms

	First slot		Second slot	
	winner	payment	winner	payment
v ranking	A	0.7	B	0.4
$v \times \alpha$ ranking	B	0.4	C	0.3
α ranking	B	0.4	A	0.4

3.3. Example

Two sponsored slots are available for a certain keyword, and there are four advertisers (A, B, C, D) interested in these slots. The amounts they're willing to pay per click are $v_i = (0.8, 0.7, 0.4, 0.2)$, while their true relevance scores are $\alpha_i = (0.3, 0.7, 0.8, 0.2)$. Table 1 states the winner and prices for each slot under each mechanism. The posted-price mechanism is omitted here due to its simplicity.

4. Revenue Comparisons

Although the four mechanisms share the same functional form for the expected revenue function, there are significant differences in detail. The mechanisms have different allocation rules $r : I \to J$, so $r(i)$ is different for each mechanism even for the same i—consequently, the variables $\alpha_{r(i)}$ and $v_{r(i)}$ are different. The prices p_i are endogenous, so p_i differs across the four mechanisms. Hence the revenue function is in fact quite different for each mechanism. Moreover, the revenue function is stochastic (based on the joint density function f), so the differences between the mechanisms make it hard to uncover meaningful results even for expected revenues. Finally, our interest goes beyond a comparison of the static revenue function. In Section 5, we extend our analysis to cover dynamic ranking, which considers the case where the search engine revises its allocations by observing the actual number of clickthroughs received by the firms occupying the paid slots. For these reasons, we adopt a systematic computational simulation as the basis for our analysis, leading to more intuitive and straightforward characterization of results.

This section describes our simulation experiments to compare the steady-state performance of alternative mechanisms. Section 4.1 motivates the design of the experiments, and the rest of the section provides the results. Section A in the Online Supplement to this paper on the journal's website summarizes the design of the simulation code, including flowcharts of the simulation process (Figures 1 and 2 in the Supplement).

4.1. Experiment Design

Intuitively, because of the pay-per-click format of sponsor-search pricing, the comparative revenue performance of the different mechanisms is influenced by the extent of correlation between the listing firms' willingness to pay for slots, and their ability to attract traffic (which presumably is reflective of their relevance to the search term). The extent of correlation is unobservable to the search engine. However, its impact on revenue performance varies across the mechanisms, so the choice of mechanism requires a good understanding of the relationship between correlation and relative performance. Our computational analysis accounts for this requirement by treating v and α as joint random variables, and by systematically varying the degree of correlation ρ between these variables over multiple simulation runs. We normalize the variables to lie in $[0, 1]$. We model $f(\alpha_j, v_j)$ as a truncated bivariate normal distribution between 0 and 1. The means are 0.5 and the covariance can vary between -0.167 and 0.167, making the correlation between these two factors vary from -1 to 1. We expect the region of positive correlation to be the most realistic, because advertisers generally have an incentive to request relevant placement in order to attract traffic from genuinely interested users. Moreover, we expect that, in the long run, advertisers themselves will withdraw irrelevant or ineffective advertisements, retaining only those that genuinely attract interested consumers and are cost effective.

Another factor that affects the performance of each mechanism is the intensity of competition for sponsored slots. Is this effect identical across all mechanisms? Intuitively, we believe the effect should be different, because of the differences in how each mechanism considers the anticipated clickthroughs for the bidding firms. We also note that the intensity of competition is affected by two variables, the number of interested advertisers (not controllable by the search engine) and the number of slots made available for paid placement (which is set by the search engine). Hence our analysis covers variations in both the number of available slots k and the number of firms bidding for paid placement s. Finally, we expect the performance of different mechanisms to be affected differentially by the extent of attention decay over the sponsored slots, so we vary δ across our simulations.

4.2. Impact of Correlation between WTP and Relevance

First we compute the relative sensitivity of each mechanism to the correlation between an advertiser's WTP and relevance. The following result is derived with $k = 5$ available slots, $s = 15$ advertisers who desire paid placement, and an attention decay factor $\delta = 2$. For each sample value of correlation in $[-1, 1]$, we compute the average revenue over 200 runs. Each run consists of a random draw from $f(\alpha_j, v_j)$ for each of the s content provider's WTP and relevance.

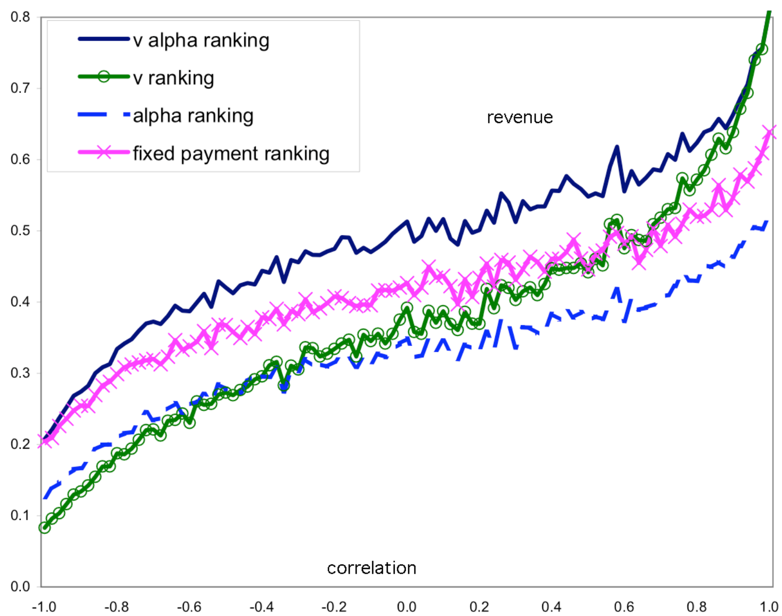


Figure 2: Average Revenue vs. Correlation for the Four Tested Ranking Strategies

Finding 1 *For every mechanism, the revenue earned is increasing in the correlation between the content provider's relevance score and its willingness to pay. The effect is most pronounced for the v ranking mechanism.*

Finding 1 is intuitive, since a greater correlation results in picking more relevant links, as indicated in Figure 2. The v ranking mechanim, because it ignores relevance, benefits the most from an increase in the correlation. This underscores the need for stronger editorial control for relevance. Interestingly, this is consistent with industry practice: `Overture`, which ranks by v, does generally expend more resources on human editorial control than `Google`, which employs a form of $v \times \alpha$ ranking.

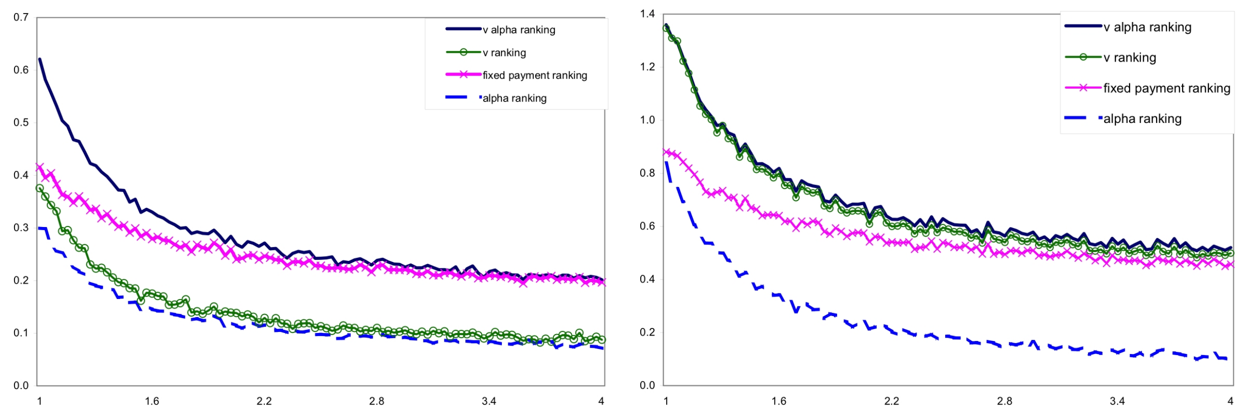
Finding 2 *$v \times \alpha$ ranking weakly dominates the other three mechanisms, and strongly dominates in the region of negative correlation. The posted-price mechanism performs better than v ranking under negative correlation, but worse under positive correlation. α ranking is always dominated by the other mechanisms.*

Figure 2 illustrates the robustness and dominance of the $v \times \alpha$ ranking. The v ranking mechanism performs well when v and α are positively correlated, but quite poorly when ρ is negative (in this case the posted-price mechanism does nearly as well as $v \times \alpha$ ranking). The poor performance of v ranking in the region of negative correlation occurs because the mechanism systemically picks the providers who will achieve lower clickthroughs, hence earns lower revenues. The posted price mechanism performs well under negative correlation because the search engine moves the risk of enrolling some less relevant listings to the listing companies. These results are summarized in Finding 2. Table 1 in the Online Supplement shows the two-sample t test assuming unequal variances for the revenues generated by the v ranking and $v \times \alpha$ ranking mechanisms when the correlation is negative, showing a significant difference. Table 2 (Supplement) shows the same t-test for the case where the correlation is positive, showing no significant difference.

Even though $v \times \alpha$ ranking performs well across all values of ρ, the search engine's choice of mechanism is not straightforward. This is because use of the $v \times \alpha$ mechanism requires knowledge of relevance scores. In the absence of good estimates of relevance scores, the $v \times \alpha$ curve in Figure 2 merely represents an upper bound on revenue, corresponding to perfect estimation of α. In this case, knowledge of ρ can guide the choice of mechanism. When management is confident that v and α are highly correlated, then it may be best to award slots by bid (v ranking, which performs close to the upper bound on revenue). Conversely, when v and α are negatively correlated or independent, then the posted-price mechanism comes close to the revenue upper bound. Finally, estimates of relevance scores can be improved over time, indicating that search engines should employ a dynamic mechanism that embeds learning of relevance scores into the allocation of slots. We explore this issue in Section 5.

4.3. Impact of Attention-Decay Factor

Now we examine how the performance of the different mechanisms is influenced by δ, the difference in the attention that a certain listing item can get in different positions. From (1),



(a) negative correlation (b) positive correlation

Figure 3: Revenue vs. Attention Decay Parameter δ, under Negative and Positive Correlation Between Relevance and Willingness to Pay.

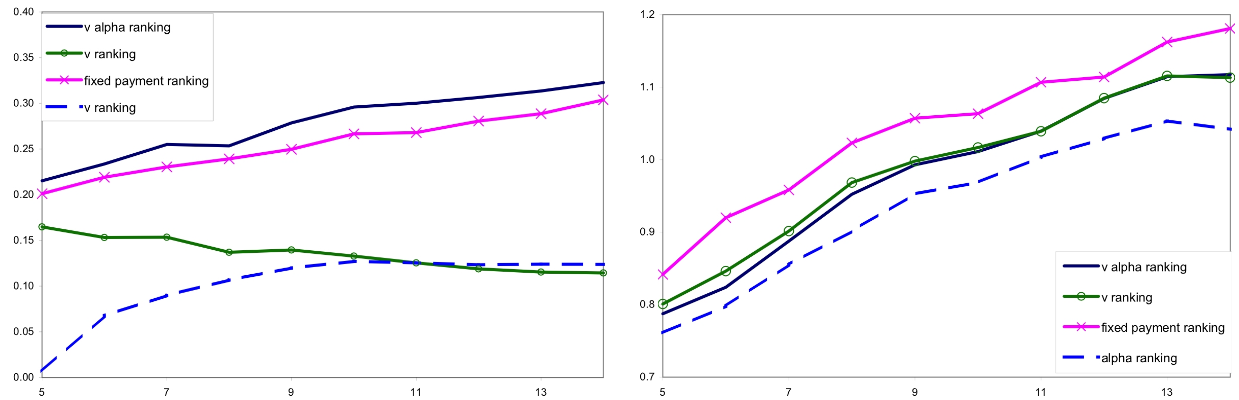
we see that revenue decreases as attention decay increases, since $P_i = \frac{p_i \alpha_j}{\delta^{j-1}}$. Figures 3a and 3b show the performance of each mechanism with respect to δ, when the correlation between paid listing firm's relevance and WTP is strongly negative (left subfigure) and strongly positive (right subfigure), respectively.

Finding 3 *The revenue generated from sponsored search decreases as δ increases.*

This result is intuitive. As δ increases, the lower rank positions become less attractive so will generate lower revenues. We note that the search engine can exert some control over the decay factor, e.g., by designing a better user interface that maintains attention over a larger subset of paid listings. Still, as the limits of improved user interfaces are reached, a fundamental limitation on human attention is unavoidable. Figure 3 also indicates that the posted-price mechanism converges to the performance of the $v \times \alpha$ mechanism when there is significant attention decay.

4.4. Impact of Demand for Sponsored Search

The search engine's potential for placement revenues depends on the overall demand for sponsored search. Intuitively, when more providers compete for these slots, this will increase the market clearing price and the search engine's revenues. Our analysis extends this intuition by revealing that this relationship is also affected by the correlation between the paid listings' relevance score and willingness to pay. Figure 4 demonstrates that placement revenues increase with the level of demand when WTP and relevance are positively correlated, but that an increase in demand does not yield greater revenues (except in the case of $v \times \alpha$ ranking, where revenue increases modestly with demand) when the correlation is negative.



(a) negative correlation

(b) positive correlation

Figure 4: Revenue vs. the Demand for Paid Listings s, under Negative and Positive Correlation Between Relevance and Willingness to Pay.

Finding 4 *When the correlation between the paid listings' relevance and WTP is highly positive (cov = 0.15, correlation=0.9 in this case), the increase in the demand for sponsored search increases the search engine's revenue. When the advertisers' relevance and WTP are very negatively correlated (correlation* $= -0.9$ *in this case), there is no obvious increasing trend with increasing the demand.*

It might appear counter-intuitive that an increase in demand for sponsored slots fails to increase revenues. However, when relevance and WTP are negatively correlated, the v ranking and posted-price mechanisms systematically favor those advertisers with lower relevance since these are the advertisers with the k highest bids. An increase in the number of advertisers competing for the k slots amplifies this effect. The $v \times \alpha$-ranking mechanism, on the other hand, cancels out the negative correlation to some extent since it incorporates relevance into the selection of advertisers.

4.5. Impact of Supply of Paid Listings

The search engine controls the number of paid slots it makes available. Intuitively, the more paid listings a search engine has, the more revenue it can generate from advertisers. However this expanded enrollment will most likely cause the search engine to enroll some listings with less relevance, thus negatively affecting the overall quality of the search engine as perceived by users. This in turn will reduce the total traffic at the search engine and the clickthrough rates, thereby lowering revenue from sponsored search. Hence, our intuition would indicate that increasing the number of paid slots will benefit the search engine up to a point, but that further increase in slots will cause reduction in overall placement revenue.

To verify this intuition, we extend our experiment, varying the amount of paid links (k) the search engine decides to enroll. In this round, we assume there are 20 potential advertisers who are interested in purchasing paid slots, and the search engine varies the number of paid slots from 1 to 19. Again we present the expected revenue over 200 runs for each value of k. Results are shown in Figure 5.

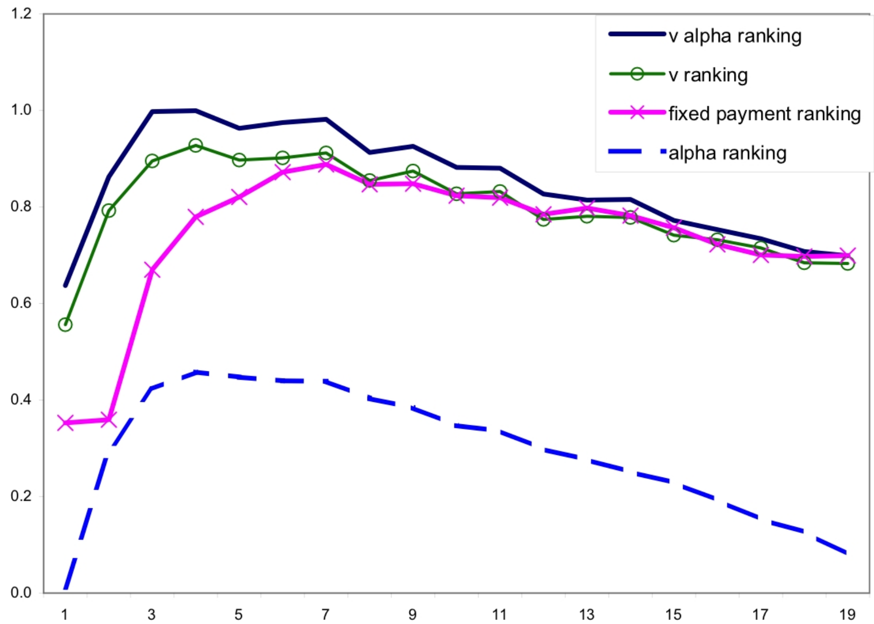


Figure 5: Revenue vs. Number of Paid Positions Available (k), under Positive Correlation Between Relevance and Willingness to Pay.

Finding 5 *When WTP and relevance have a high positive correlation, the search engine's expected revenue is approximately concave in the number of paid links it enrolled.*

Interestingly, the optimum values displayed in Figure 5 roughly correspond to typical industry practice of displaying up to five sponsored listings alongside search results.

4.6. The Effect of Editorial Control on the Quality of Paid Listings

Our analysis indicates that when relevance and bid prices are negatively correlated, the v ranking mechanism tends to award sponsored slots to advertisers who have low relevance, thus yielding lower revenue overall. However, in practice, paid placement companies exert some editorial control to maintain relevance. `Overture` in particular expends considerable resources in human editorial screening to filter out irrelevant or objectionable listings. Along these lines, we implement a *modified* v ranking mechanism that removes listings below a certain threshold relevance. The filtering step helps in two ways: (1) it tends to select the better-performing advertisers when the correlation is negative, and (2) it controls the overall quality of the paid links, yielding higher traffic.

We ran a parallel set of simulations to test the performance of modified v ranking. In implementing the mechanism, we assume that the editorial control policy is perfectly able to identify and eliminate listings below a specified threshold level. Our implementation works as follows. Advertisers below the specified threshold are discarded, and then the highest k (or fewer) bids are selected from the remaining advertisers. Again every data point presented in the simulation represents an average of 200 runs. Figure 6 shows the performance of the modified v ranking in comparison to the $v \times \alpha$ ranking mechanism, with different choice of threshold levels.

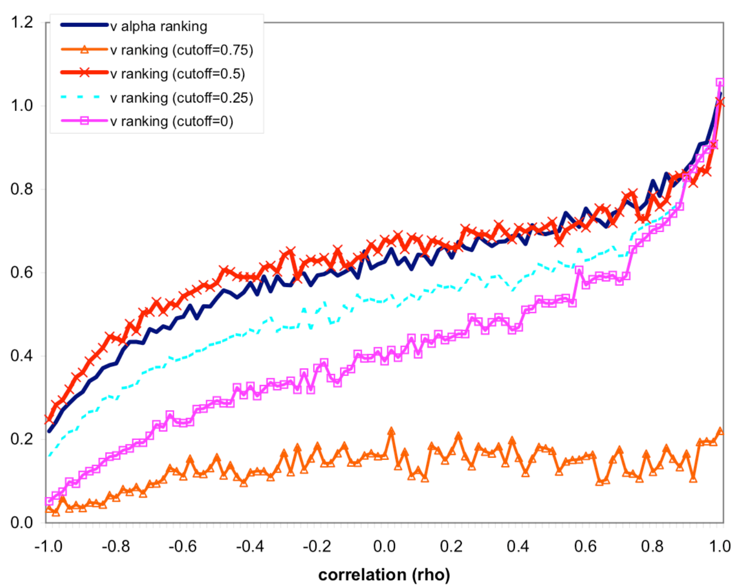


Figure 6: The Performance of the Modified v Ranking with Different Cut-Off Values.

Finding 6 *The modified v ranking can improve the performance of the original v ranking mechanism. With a good choice of the threshold value, the modified v ranking mechanism performs better than $v \times \alpha$ ranking in regions of both positive and negative correlation.*

In our simulation, modified v ranking, with a threshold value of 0.5, performed slightly better than $v \times \alpha$ ranking when the correlation ρ is negative or positive but small, and performed equivalently when the correlation is positive and large; Tables 3 and 4 (in the Online Supplement) give the t-test results. This result is explained as follows: with the added control over quality, the modified v ranking picks up more relevant paid links and improves the search engine's quality perception, thereby generating greater traffic to the search engine; the $v \times \alpha$ ranking, on the other hand, can on occasion award slots to links with low relevance.

The simulation also demonstrates that the threshold value must be chosen carefully. If it's set to be too low (e.g., 0.25 in this case), the modified mechanism has limited ability to eliminate the less relevant listings; however if it's too high (e.g., 0.75), few listings will survive the screening so the search engine will be forced to leave more slots unfilled and will generate lower revenue. In reality, any editorial control policy will be less than perfect, but the degree of accuracy in filtering low relevance links can be improved by making further investment in the editorial process. The search-engine firm can therefore choose a suitable level of investment by trading the costs with the improvement in revenue.

5. Ranking Dynamics

The previous section analyzed the steady-state performance of various ranking mechanisms, which requires knowledge of the listing's relevance. Since the search engine does not have this information, it approximates the relevance by observing the clickthrough rate for each listing. This approximation improves over time as the search engine gathers more data, hence the ranking produced by each mechanism needs to be revised over time until the steady-state solution is obtained. This section studies the ranking dynamics and compares different approaches for dynamic ranking. Accounting for measurement error in clickthrough rates makes sense only for the mechanisms that depend on clickthrough rate, namely $v \times \alpha$ ranking and α ranking. We focus our analysis on the $v \times \alpha$-ranking mechanism, because the α ranking is dominated by all other methods.

Since $v \times \alpha$ ranking orders listings based on the product of bid price and number of clickthroughs, the ranking will change over time as the observed clickthrough history itself evolves. Items that receive more clicks get promoted in the list. We implement dynamic ranking by dynamically computing each listing's *clickscore*, and then ranking the listings by the product of bid price and clickscore. We compare two different approaches for score revision. First, in the basic $v \times \alpha$ ranking, each click increases the listing's clickscore by 1, regardless of whether the click was earned at a high rank or a low rank. In this case, each click has an equal impact on the likelihood of promotion. We call this the *unweighted* revision mechanism. Second, we propose a new rule where clicks on low-ranked items have greater impact on promotion. We implement this rule by revising the listing's clickscore as follows: each click received by the listing increases the listing's click-score by δ^i, where i is the listing's position when it generated the click. Intuitively, a lower-ranked item is more likely to move up in the list under this *weighted* revision mechanism than under the *unweighted* revision mechanism.

We performed computational simulations to examine the dynamics of each revision mechanism, with the following settings. There are five sponsored slots available and there are five content providers. In each simulation run, each provider's relevance score and willingness to pay are i.i.d. random variables that follow a truncated normal distribution in $[0, 1]$. The decay factor δ is 2. In each period of a simulation, a provider's probability q of getting a clickthrough is $\frac{\alpha_j}{\delta^i}$, a function of its true relevance score α and its position i in the result page. Since there is no information about clickthrough rates in the beginning, each simulation run consists of a trial period where every item is given a chance to be on the top as well as every other position: the clickthrough they earned is recorded, and the weighted and unweighed products (scores) are calculated. At the end of the trial period, the items are presented and ranked according to their scores. In the remaining periods, the ranks are revised according to the weighted and unweighted revision rules. Once a lower-ranked item achieves a larger score, it can be promoted to a higher position. We executed 200 runs for each mechanism. For each run, we record the number of periods to converge to the optimal allocation, and terminate it after 1000 periods if there is no convergence. In each period, the average *distance* between the optimal allocation and the current allocation is also calculated as the the sum of the absolute value of the difference between the current rank and equilibrium rank of each listing. More specifically, let R_0 be the vector of r_i's in the optimal case (where $r : I \to J$ is the ranking function allocating position i to listing company j), and R_t be the

Table 2: Number of Runs That Converged within 1000 Periods

trial period	1	3	5	7	9	11	13	20
weighted	88	96	112	116	125	125	133	130
unweighted	42	73	73	78	92	92	96	99

Table 3: Number of Periods at which the Mechanism Converges

trial period	1	3	5	7	9	11	13	20
weighted	595.87	558.71	476.83	470.89	417.12	423.06	382.62	380.32
unweighted	804.87	650.25	643.76	616.17	560.69	558.59	533.80	519.01

vector showing the ranking in period t. Then the distance between the current ranking and the optimal ranking is $\sum_{m=1}^{m=k} |r_0(m) - r_t(m)|$.

5.1. Impact of Trial Period

Due to the interdependence between the number of clickthroughs and placement rank, each simulation begins with a trial period that measures the clickthroughs over several trials. Each trial consists of five allocations, chosen such that every listing appears once in every position, offering each listing a fair chance to be tested for its true relevance. The next two findings state the performance of the two revision rules in terms of how frequently each converges to the optimal ranking, and how quickly it gets there. Table 2 shows the number of runs that converged to the equilibrium when the paid listings' relevance and WTP were independent (correlation equals zero). Table 3 shows the speed of convergence for each of the two mechanisms.

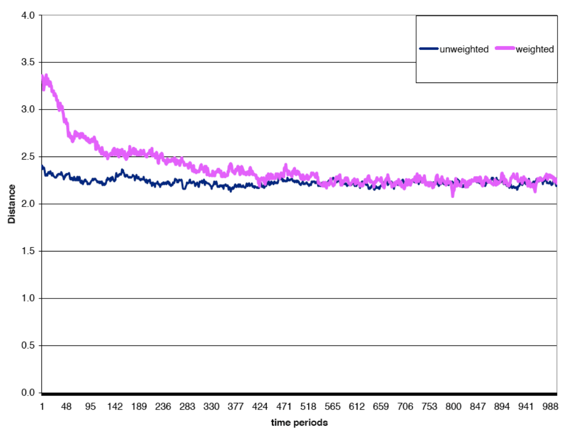
Finding 7 *An increase in the length of the trial period increases the speed of convergence, but at a decreasing rate.*

Finding 8 *The weighted mechanism exhibits better convergence to the optimal state than the unweighted mechanism, both in terms of the number of runs that converge, and the speed of convergence.*

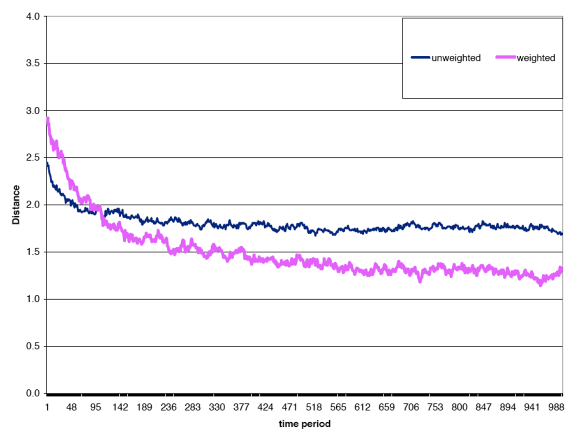
Convergence to the optimal ranking is important because any deviation leads to lower revenue. Speed of convergence is important because demand for sponsored slots—for specific terms—is highly variable, thus if a mechanism takes too long to converge it will implement the optimal policy for only a very short period. The number of trials has a significant effect on the performance of each mechanism, since trials are costly (they involve allocations that generate poor revenues) but generate useful information for approximating relevance. Thus a longer trial period produces a better approximation of relevance (thereby resulting in greater revenues in future periods) at the expense of lower revenues during the trial periods, a typical exploration-exploitation tradeoff. The weighted mechanism performs better because it is better able to recognize errors in ranking—it promotes more quickly those listings that receive clicks at lower ranks.

5.2. Impact of Correlation

Now we examine the impact of correlation between WTP and relevance on the convergence properties of the two mechanisms. Figure 7 shows the average distance between the current run and the equilibrium state in each period. The results are displayed for the two cases of large negative correlation ($\rho = -0.9$, or covariance equals -0.15) in Figure 7a and large positive correlation ($\rho = -0.9$) in figure 7b. The trial period in each simulation consisted of five trials.



(a) negative correlation



(b) positive correlation

Figure 7: Distance to Equilibrium vs. Time for the Weighted and Unweighted Ranking Mechanisms, under Negative (left) and Positive (right) Correlation Between Relevance and Willingness to Pay

Finding 9 *Both mechanisms converge more closely to the equilibrium as the correlation between WTP and relevance increases. The weighted mechanism always converges faster* *than the unweighted mechanism. The difference between the two mechanisms increases with increase in correlation.*

The reason that performance improves over time is that the search engine develops a better approximation of the true relevance by observing the actual clickthroughs. The intuition behind the better performance of the weighted ranking mechanism is that lower-ranked items are less likely, *ceteris paribus*, to be selected by a user (due to attention decay at lower positions). Hence if some link generates more clicks despite this handicap, this signals a greater error in its current position. The weighted ranking rule accounts for this fact, while the unweighted rule ignores the rank at which clicks were received. The difference between the two mechanims' performance is statistically significant at the 0.05 level. Tables 5 and 6 in the Online Supplement display the t-test statistics for $\rho = -0.9$ and $\rho = 0.9$, respectively.

6. Conclusions

This paper analyzed the implementation of sponsored search strategies for Web search engines. Via computational simulations, we compare four alternative mechanisms for allocating sponsored slots, including two stylized versions of mechanisms employed by the two leading services: `Yahoo/Overture`'s v ranking mechanism and `Google`'s $v \times \alpha$ ranking mechanism. We find that $v \times \alpha$ ranking performs equally or better than other mechanisms in almost all cases, while v ranking works comparably in the expected case of positive correlation between v and α. Editorial filtering can improve the performance of v ranking significantly. Placement revenues decrease when users' attention is significantly lower for lower-ranked listings, emphasizing the need to develop better user interfaces and control features. The search engine must carefully choose the total number of paid slots to make available, due to the tradeoff between direct revenue increases and indirect revenue losses due to consumer defection. While the $v \times \alpha$ ranking dominates in the computations, it is of little use unless good estimates of relevance scores are available. Search engines can improve these estimates by observing clickthroughs, at different ranks, so the accuracy of learning is critical to performance. We examined the dynamic behavior of the $v \times \alpha$-ranking mechanism, which promotes (or demotes) a provider's rank based on the number of clicks it receives, and proposed an alternative mechanism where the reward for a click is larger when it is received at a lower rank. We found that this weighted mechanism converges faster and is more stable than the unweighted mechanism for rank revision. Based on personal communication, we believe that `Google` may employ a similar weighting procedure in practice.

Sponsored search in search engines is a thriving and growing industry. The average price paid per click on `Overture`'s network in 2003 was roughly US$0.40. Total industry revenue was approximately US$0.9 billion in 2002 and US$2.5 billion in 2003. This practice is widely credited for the revitalization of the search-engine business. This paper has taken early steps in studying the implementation of sponsored search in search engines. Much more work remains to be done, with respect to understanding user attitudes toward different forms of paid placement, the impact of various user interfaces on users' willingness to accept and browse sponsored slots, and on the design of optimal mechanisms for allocating these slots. For example, laboratory experiments can be done to determine how advertisers' bidding strategies are affected by the various mechanisms in practice.

Acknowledgements

We thank Sandip Debnath, Dan Fain, Gary Flake, Lee Giles, and Steve Lawrence.

References

Asdemir, K., N. Kumar, V. Jacob. 2002. Internet advertising pricing models. Tech. rep., School of Management, The University of Texas at Dallas, Richardson, TX.

Bhargava, H. K., J. Feng. 2002. Preferential placement in internet search engines. D. Lassner, ed., *Proceedings of the Eleventh World Wide Web Conference*. Honolulu, HI, 117–123.

Breese, J. S., D. Heckerman, C. Kadie. 1998. Empirical analysis of predictive algorithms for collaborative filtering. S. M. Gregory F. Cooper, ed., *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*. University of Wisconsin Business School, Madison, Wisconsin, 43–52.

Hoffman, D., T. Novak. 2002. How to acquire customers on the web. *Harvard Business Review* **78** 179–188.

Hu, Y. J. 2003. Performance-based pricing models in online advertising. Tech. rep., Sloan School of Management, MIT, Cambridge, MA.

Kumar, S., M. Dawande, V. Mookerjee. 2003. Revenue maximization in web advertisement. D. Dey, R. Krishnan, eds., *Proceedings of The Thirteenth Annual Workshop On Information Technologies And Systems (WITS'03)*. Seattle, WA, 121–126.

Lawrence, S., C. L. Giles. 1999. Accessibility of information on the web. *Nature* **400** 107–109.

Rolland, E., R. A. Patterson. 2003. Classification in online pay-for-performance advertising. D. Dey, R. Krishnan, eds., *Proceedings of The Thirteenth Annual Workshop On Information Technologies And Systems (WITS'03)*. Seattle, WA, 219–224.

Tantono, W., S. Selby, J. Bagner, C. Sonu. 2002. Ftc advises search engines to disclose paid placements. *Intellectual Property & Technology Law Journal* **14** 21–22.

Weber, T. A., Z. E. Zheng. 2003. A model of search intermediaries and paid referrals. Tech. rep., 02-12-01, The Wharton School, University of Pennsylvania.