Paid Placement Strategies for Internet Search Engines

Hemant K. Bhargava
Smeal College of Business
Penn State University
342 Beam Building
University Park, PA 16802
bhargava@computer.org

Juan Feng
Smeal College of Business
Penn State University
118 Beam Building
University Park, PA 16802
jif1@psu.edu

ABSTRACT

Internet search engines and comparison shopping have recently begun implementing a paid placement strategy, where some content providers are given prominent positioning in return for a placement fee. This bias generates placement revenues but creates a disutility to users, thus reducing user-based revenues. We formulate the search engine design problem as a tradeoff between these two types of revenues. We demonstrate that the optimal placement strategy depends on the relative benefits (to providers) and disutilities (to users) of paid placement. We compute the optimal placement fee, characterize the optimal bias level, and analyze sensitivity of the placement strategy to various factors. In the optimal paid placement strategy, the placement revenues are set below the monopoly level due to its negative impact on advertising revenues. An increase in the search engine's quality of service allows it to improve profits from paid placement, moving it closer to the ideal. However, an increase in the value-per-user motivates the gatekeeper to increase market share by reducing further its reliance on paid placement and fraction of paying providers.

Categories and Subject Descriptors

H.3 [**Information Systems**]: Information Storage and Retrieval

General Terms

Economics

Keywords

Search engines, information gatekeepers, paid placement, bias, promotion

1. INTRODUCTION

The Internet and World Wide Web are home to vast repositories of information— from text to multimedia, from amateur opinions to expert thought, from voluntary contributions to commercial interests—on every conceivable topic. Lawrence & Giles [12] estimated the publicly indexable Web at 800 million pages, 6 terabytes of text data, on 2.8 million servers, as of February 1999. Internet search engines, which serve as a gateway to this information repository, have established a crucial role in today's important society. Most studies of Internet usage find that search engines play a vital role in information retrieval over the Internet. A recent *USA Today* (Dec. 11, 2000)[1] article states that 100 million queries are made on U.S. search engines each weekday, and a study of Web usage[2] by `Media Metrix` found that the top 3 search engines were each visited by 61%, 56% and 40% of tracked Internet users during the past month. The widespread use of search engines has facilitated technology transfer, so that search engine technologies are now licensed to business Web sites, used in digital library systems, etc. For the purpose of this paper, the term search engine encompasses various applications of these indexing-retrieval technologies, including traditional Web search engines (e.g., `Google`), metasearch engines (`Metacrawlers`), niche search engines (e.g., DEADLINER (Gruger etal (2000)) [10]), information portals (`Yahoo!`), and comparison shopping engines (`mySimon`).

Most search engines began as university projects that focused more on development and algorithms, and less on revenue generation. Even after transitioning into commercial entities, search engines tended to operate as a free resource to content providers and users alike. However, the recent drop in supply of cheap venture capital and sweat equity has forced commercial search engines to investigate mechanisms for generating revenue from content providers. These mechanisms—which we generically label as *paid placement*—include a fee for inclusion in the database, an increased relevance score in response to a query, or featured listings on the results pages. A paid placement strategy usually requires a minor modification of the ranking algorithm or to the display of results, either of which can be made at very low cost. Paid placement is widespread in search engines (e.g., `Google`), information portals (e.g., `Yahoo!`, and metasearch engines (`Metacrawler`). Nearly all major search engines and portals employ paid placement.[3] Table 1 presents data on the extent of paid placement for *metasearch* engines. And, as Figure 1 indicates, the major comparison shopping engines also employ paid placement.

The focus of this paper is on a search engine's strategy regarding revenues from content providers in its database, and how this objective conflicts with its other revenue sources which are a function of its user base, such as advertising and licensing revenues. We develop a mathematical model to analyze the dilemma that search engine faces in raising

[1]http://www.usatoday.com/life/cyber/tech/cti895.htm
[2]http://searchenginewatch.com/reports/mediametrix.html
[3]At `Search Engine Watch` http://searchenginewatch.com/-webmasters/paid.html

WWW2002, May 7–11, 2002, Honolulu, Hawaii, USA.
ACM 1-58113-449-5/02/0005.

revenues: it wants to charge content providers for priority placement, but this reduces the search engine's credibility, hence its market share and potential user-based revenues. Specifically, we determine the optimal paid placement policy, i.e., the optimal placement fee and the resulting percentage of sites that choose paid placement. Our longer term interest is to determine the optimal bias-level that would give a search engine the best balance between revenues from content providers and revenues based on its user base.

The revenue problem is a critical one for search engines, since it impacts both current performance and future development and improvements. In spite of many years of research on information retrieval, search engines are still far from perfect in terms of the usual metrics of relevance and recall. Hence, there remains considerable research and commercial interest in refining the indexing and ranking algorithms, and user interfaces, employed by search engines. Recent research examines a variety of topics, including Web page ranking algorithms evaluation and comparison of alternative ranking algorithms (Singhal & Kaszkiel [14]), contextual and topic-based search (e.g., Bharat & Henzinger [2]), design and evaluation of metasearch engines (e.g., Dreilinger & Howe [6]), metasearch using full-text analysis of Web pages (e.g., Lawrence & Giles [11]), and visualization of results (Hearst [9]). Since further research and development is expensive, commercial search engines need to find new revenue sources in order to balance these costs. Paid placement offers an intriguing possibility: placement revenues in one period can support research and development aimed at improving indexing and retrieval algorithms, database index, or user interface. Hence the negative impact (on users) of paid placement could be reversed by using placement revenues to improve search engine quality.

The rest of this paper is organized as follows. In §2, we develop our model of the search engine's revenue problem, considering network effects, the effect of paid placement, and third-party revenues. We characterize the optimal paid placement strategy in §3. In §4 we discuss the sensitivity of the paid placement strategy to various controllable parameters such as the extent of bias and search engine quality, and other factors such as perceived disutility and the advertising rate. We conclude with a summary of our results and possible application of our work to other forms of Internet-based information intermediaries.

2. MODEL OF SEARCH ENGINES' REVENUE PROBLEM

Consider a search engine which offers indexing and retrieval services over a virtual library of aggregated content from multiple content providers, and presents an ordered list of results in response to a user's query term. In developing an economic model of search engines, we consider three types of entities: users of search engines, content providers, and third-parties such as advertisers and licensing firms. The search engine serves a market of users who are heterogeneous in their preferences. Let Θ represent user types in a descending order of users valuation for the searching service, and let q represent quality of the search engine as perceived by users, which may be a composite measure of its database of content providers, user interface, and indexing and retrieval algorithms. We write $U(\theta, q)$ to denote the value to a type-θ user for a given quality q. For convenience we assume that θ is uniformly distributed in the interval $[1, 2]$. We consider q to be exogenously specified in the period of interest. The search engine benefits content providers by directing users to their sites. Content providers are also heterogeneous in their profit expectations. Let Φ represent a descending order of provider types, where (for convenience) ϕ is uniformly distributed in $[1, 2]$.

2.1 Network Externalities

We conceptualize the search engine as offering network benefits to both users and content providers i.e., the overall value of the search engine to users increases in the total number of content providers), and the value to content providers increases in the number of users. Considering first the users' valuation function, this means that $U(\theta, q)$ is an increasing function in q since q represents the index of content providers. For convenience we assume that this function is $U(\theta, q) = \frac{q}{\theta}$. Similarly, the value to content providers $V(\Phi, \mathcal{M})$ is an increasing function of the search engine's market coverage or user base $\mathcal{M}$. We write the provider valuation function as $V(\phi) = \frac{\mathcal{M}}{\phi}$.

The search engine's generates revenues on the basis of its user base, including revenues from third-party firms such as advertisers and fees for licensing their information retrieval technologies. Advertisers are interested in exposure to users of the search engine, hence advertising revenues are a function of the search engine's user base. Firms that license the search engine technology pay a "per-click-through" fee, hence these revenues are also a function of the user base. For simplicity in exposition, we denote all user-based revenues as advertising revenues. Let a represent the advertising value per user, so that the search engines revenues equal

$$\pi_1 = a \cdot \mathcal{M}$$

2.2 Effect of Paid Placement

Internet search engines execute a user's search query on a database index, and typically return a set of results ranked according to their *relevance score*. Given these results, the user selects specific content providers in the list for further transactions. It is well known that the ranking of a result term is strongly correlated with the probability that the user will follow up on the result term (McLuhan ([13]). Commercial content providers are interested in *clickthroughs* and *conversion rates*—i.e., the likelihood that a search engine user will enter into a commercial transaction with the content provider. For this reason, content providers have an incentive towards *paid placement*—to pay the search engine in order to be included, ranked highly, or prominently featured in the search result.[4] In practice, this may mean a higher relevance score, a featured listing, or perhaps even a guaranteed retrieval for certain search terms. Figure 1 displays a screen shot from a comparison shopping engine that includes paid placement (featured listing), normal placement, and advertising.

What is the impact of paid placement on a search engine's perceived quality? Here we assume that search engine can not hide the fact that they perform paid placement, because this can not exist in equilibrium in the long run (bidders can

[4]The *USA Today*, cited earlier, quotes the CEO of `GoTo.com` as saying that commerce-related requests make up nearly 50% of search engine queries, and that over 32,000 businesses pay to be listed on `GoTo`.

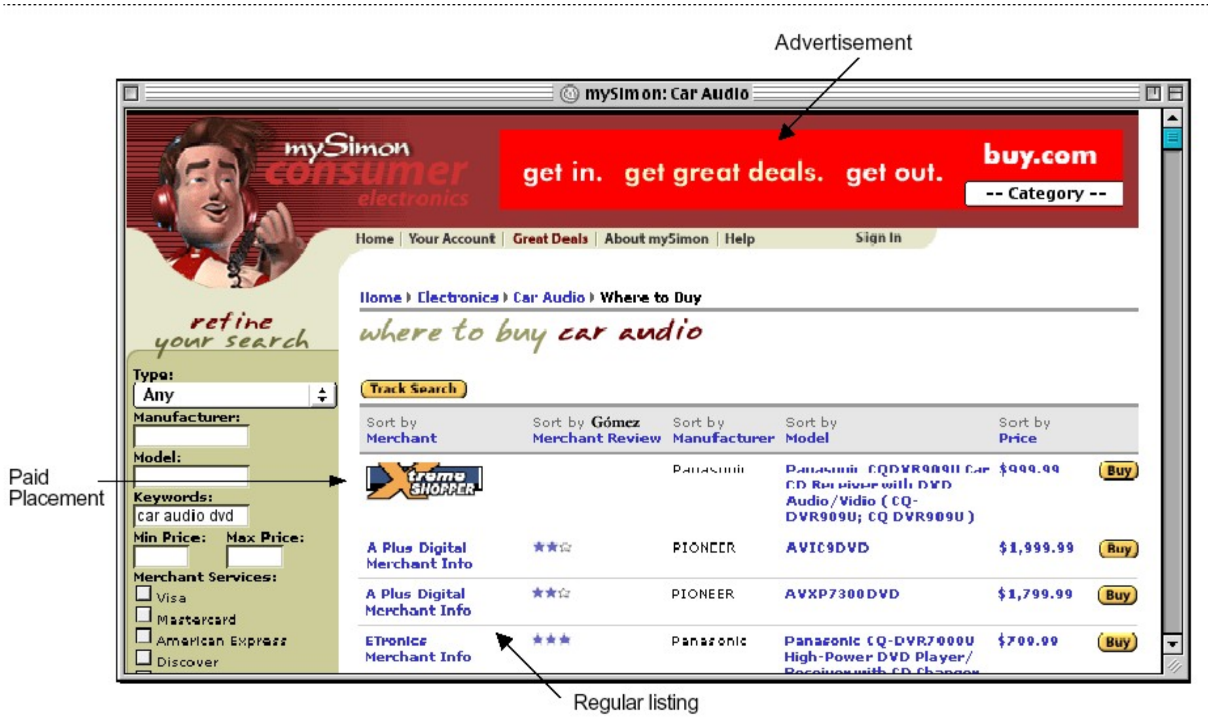

Figure 1: Paid placement, regular listings, and advertising in a comparison shopping engine. The top listing and graphical icon increase the likelihood that the paid placement listing will be followed up.

perceive finally), or may cause some serious legal issues.[5] Articles in the business press and data from commercial research firms suggest that paid placement strategies have a negative impact on a search engine's perceived quality and credibility. Goodman [8] argues that search engines must act as "referees—fair arbiters of relevance" or they will lose market share. Since loss of market share causes a fall in advertising revenue, search engines must trade-off potential revenues from paid placement with those from advertising.

To model the effect of paid placement, suppose that the search engine offers priority placement to content providers who pay a placement fee γ. Let x represent the percentage of free listings, hence $1-x$ is the fraction of providers who choose paid placement. The search engine can bias its relevance scoring algorithm and displays in many ways. Let β be the extent of bias chosen by the search engine. We represent the positive effect of the bias β on content providers with the function $\lambda(\beta)$, and the negative effect on users (after normalization, without loss of generality) as β itself. Assume $\frac{\partial\lambda}{\partial\beta} > 0$, and $\frac{\partial^2\lambda}{\partial\beta^2} \leq 0$. Hence β represents the users' perceived disutility of paid placement. In the current model, β is treated as exogenous. However, our future work aims to endogenize β and determine the optimal bias level.

Due to the negative impact on users, we rewrite the utility function as $U(\theta, q, \beta) = \frac{q(1-\beta(1-x))}{\theta}$. Providers who choose paid placement get an additional benefit $\lambda(\beta)$, so that provider ϕ's valuation increases to $V(\phi)(1+\lambda(\beta))$. The search engine now gets additional revenues

$$\pi_2 = \gamma \cdot (1-x)$$

2.3 Search Engine's Profit Function

To compute the search engine's profits, we first determine the fraction of users that will visit the search engine and the fraction of content providers that will choose paid placement. Let c be the threshold value desired by users before they use the search engine. This may represent an opportunity cost or effort in using the search engine, or the value provided by a competing search engine. Hence the users who use the search engine's service are $\{\theta : U(\theta, q, \beta) \geq c\}$. Since θ is uniformly distributed in $[1, 2]$, the search engine's market coverage is

$$\mathcal{M} = \frac{q(1-\beta(1-x))}{c} - 1 \tag{1}$$

To make the problem meaningful, we require that at least the highest valuation user ($\theta = 1$) will use the search engine when there is no paid placement, i.e., that $q \geq c$.

The search engine benefits content providers by directing users to their sites. Content providers are heterogeneous in their profit expectations, which is a function of the search engine's market coverage $\mathcal{M}$. Providers have a choice between regular placement (which provides a value $V(\phi)$ at no cost) and paid placement (which provides value $V(\phi)(1+\lambda(\beta))$ at cost γ). Rational providers will choose paid placement if and only if $V(\phi)(1+\lambda(\beta)) - \gamma \geq V(\phi)$ i.e., $\phi \leq \frac{\mathcal{M}\cdot\lambda(\beta)}{\gamma}$. Hence the fraction of providers who choose paid placement

[5] For example, Commercial Alert filed its complaint with the FTC on Monday, claiming that AltaVista, AOL Time Warner, Direct Hit, iWon, LookSmart, Microsoft and Terra Lycos are violating US law by inserting paid listings within their search engine results "without clear and conspicuous disclosure that the ads are ads." http://searchenginewatch.com/sereport/01/07-ftc.html

is

$$(1-x) = \frac{\mathcal{M} \cdot \lambda(\beta)}{\gamma} - 1 \quad (2)$$

The search engine obtains revenues from two sources, third party firms and paid placement. The first type of revenue $\pi_1 = a \cdot \mathcal{M}$ is a function of the search engine's market coverage, $\mathcal{M}$, and profit rate brought by each user, a. If $\mathcal{M}$ is interpreted as the number of queries to the search engine, a may be considered as the rate per impression. Hence

$$\pi_1 = a\left(\frac{q(1-\beta(1-x))-c}{c}\right)$$

The search engine's placement revenue π_2 is $\gamma(1-x)$. Substituting for $\mathcal{M}$ and rearranging terms, we get

$$\pi_2 = \lambda(\beta)\left(\frac{q(1-\beta(1-x))-c}{c}\right) - \lambda(\beta)\left(\frac{q(1-\beta(1-x))-c}{c(2-x)}\right)$$

The search engine's total profits are $\pi = \pi_1 + \pi_2$, and it aims to choose the optimal fraction of paid placement $1-x$ (alternately, the optimal degree of independence, x) in order to maximize π.

2.4 Literature Review

In related work, Bhargava & Choudhary [3] and Corbett & Karmarkar [4] study the case where an intermediary has the option to charge subscription fees for customers and listing fees for suppliers. However, Bhargava & Choudhary [3] consider only a one-sided network benefit, and their model does not incorporate advertising revenue. Corbett & Karmarkar [4] model two-sided network benefits, but assume homogeneous content providers and do not incorporate advertising revenue. Baye & Morgan [1] applied a game theoretic model to study a similar question. None of these papers consider the possibility that the gatekeeper may bias its outputs due to payments from content providers. Dewan et al. [5] study the problem faced by a content web sites to balance content and advertising by an infinite horizon control program. Gabszewics et al. [7] analyze the case of two TV channels competing in both the audience market and the advertising market. But in both these models, advertising is the only revenue source, hence no trade-off between different resources is considered.

3. OPTIMAL PLACEMENT STRATEGY

Solving first-order conditions for the search engine's profit function, we see that the optimal degree of independence x^* is given by

$$(2-x^*)^2 = \frac{\lambda(\beta)}{(\lambda(\beta)+a)\beta}\left(\beta + \left(1-\frac{c}{q}\right)\right) \quad (3)$$

It can be verified that when $\beta > 0$, the profit function is concave and that the optimal x^* given by Eq. 3 satisfies second-order conditions for optimality.[6] Since we require that $x^* \in [0,1]$, Eq. 3 gives the optimal solution when $1 \leq \frac{\lambda(\beta)}{(\lambda(\beta)+a)\beta}\left(\beta + (1-\frac{c}{q})\right) \leq 4$.

LEMMA 1. *For any function $\lambda(\beta)$ this inequality yields an interval $[\underline{\beta}, \overline{\beta}]$ such that for all β within this interval, Eq. 3 determines the optimal x^*.*

[6]The case where $\beta < 0$ is not of interest here, but it is easy to show that $x^* = 0$, i.e., the search engine can obtain full paid placement.

To further develop the results, we first state the following property.

LEMMA 2. *Let ε represent the elasticity of λ with respect to β, hence $\varepsilon = \frac{\partial\lambda}{\lambda} / \frac{\partial\beta}{\beta}$. Then for every given β and $\lambda(\beta)$, there exists a $\hat{\varepsilon}(\beta)$ such that if $\varepsilon < \hat{\varepsilon}(\beta)$, then $[\underline{\beta}, \overline{\beta}] > 0$, and if $\varepsilon > \hat{\varepsilon}(\beta)$, then $[\underline{\beta}, \overline{\beta}] < 0$. Specifically, this value $\hat{\varepsilon}(\beta)$ equals $\frac{1-\frac{c}{q}}{a(2-\frac{c}{q})}$.*

PROOF. From the first order condition we have $(2-x^*)^2 = \frac{\lambda(\beta)}{(\lambda(\beta)+a)\beta}\left(\beta + (1-\frac{c}{q})\right)$. We know that x is decreasing with the right hand side value. Take derivative of the right hand side with respective to β, we have the sign of the first order condition is the same as the sign of $(2-\frac{c}{q}) \cdot a\beta \cdot \frac{\partial\lambda}{\beta} - (1-\frac{c}{q})\lambda(\beta)$. Define the elasticity of λ with respect to β, ε, as $\frac{\frac{\partial\lambda}{\lambda}}{\frac{\partial\lambda}{\beta}}$. This is the percentage of changes in λ induced by a small percentage change in β. When $\varepsilon \leq \frac{1-\frac{c}{q}}{a(2-\frac{c}{q})}$, the right hand side is decreasing in β, and it should be between $[1,4]$, there exists an interval $[\underline{\beta}, \overline{\beta}]$ of β to satisfy this condition. Since when $\beta = 0$, the right hand side value is infinity, we know that 0 is to the left of the interval $[\underline{\beta}, \overline{\beta}]$. So when ε is small enough and when β is between this range , $x^*$is increasing in β. On the contrary, if $\varepsilon \geq \frac{1-\frac{c}{q}}{a(2-\frac{c}{q})}$, $x^*$is decreasing in β. In the same way we know that 0 is to the right of the interval $[\underline{\beta}, \overline{\beta}]$, which is the simple case when there is no disutility toward paid placement strategy. $\square$

To understand this result, consider ε, the responsiveness of listing company's benefit to the users disutilities from paid placement strategy. When the responsiveness is relatively low, the search engine gets limited increase in revenue due to the limited increase in benefit of the listing companies by hurting users utilities, so the search engine should be careful about his paid placement strategy.

From the second part of Lemma 2, the case where responsiveness is high (i.e., $\varepsilon > \hat{\varepsilon}(\beta)$) has no interior optimal solution, and $x^* = 0$, i.e., the search engine can take full advantage of paid placement. This is because for the search engine, the benefit for the listing companies is far greater than the disutilities the users suffer, so he can make every listing company pay and increase his revenue at the same time. Hence the search engine's optimal placement policy is specified as below.

LEMMA 3. *When $\varepsilon < \hat{\varepsilon}(\beta)$, then for all β in $[\underline{\beta}, \overline{\beta}]$, the optimal placement fee is $\gamma^* = \frac{\lambda(\beta)}{2-x^*}\frac{q(1-\beta(1-x^*))-c}{c}$, and the optimal fraction of paid placement is given by $(2-x^*)^2 = \frac{\lambda(\beta)}{(\lambda(\beta)+a)\beta}\left(\beta + (1-\frac{c}{q})\right)$. When the relative disutility to users is below $\underline{\beta}$, the optimal fraction of paid placement is $1-x^* = 1$, whereas when the disutility is very high ($\beta > \overline{\beta}$), the search engine should not offer paid placement.*

This can be easily understood by considering the relationship between $\lambda(\beta)$ and β. If ε is large, i.e., the content providers benefit much more than the users suffer, the search engine makes as many providers pay as possible. Otherwise, the search engine must trade-off placement revenue with advertising and restrict the fraction of paid placements.

4. ANALYSIS

In this section we interpret the optimal placement strategy and examine the sensitivity of the placement strategy to exogenous factors and factors that the information search engine can control. We examine the case where $\varepsilon < \varepsilon(\hat{\beta})$, omitting the less interesting case of high ε which always results in the boundary solution $x^* = 0$.

Consider the effect of users' disutility (β) for paid placement. The level of bias is controllable by search engine. Hence it's useful to examine the sensitivity of placement strategy and fees to β. If users were indifferent to paid placement ($\beta = 0$), the search engine would maximize its placement revenues without regard for the effect on users. A disutility β forces the search engine to tradeoff between placement and advertising revenues. Proposition 1 elaborates on this.

PROPOSITION 1. *There exists a threshold $\beta^* \in [0, \overline{\beta}]$ such that when β is below β^*, the search engine can improve profits by increasing its bias, and when $\beta > \beta^*$, an increase in the bias causes the search engine's profits to decrease.*

PROOF. From Lemma 3, when the responsiveness of λ with respect to β is small ($\varepsilon \le \frac{1-\frac{c}{q}}{a(2-\frac{c}{q})}$) then $\beta \le \underline{\beta}$ and $x^* = 0$, hence the search engine can employ its ideal placement strategy. When $\beta > \overline{\beta}$, then from Eq. 3 it follows that $\frac{\partial x^*}{\partial \beta} > 0$. Hence the fraction of paid placements $1 - x^*$ falls. Further, we see that when β is close to 0,$x^* = 0$, and

$$\frac{\partial \pi^*}{\partial \beta} = -\frac{aq}{c} + \frac{1}{2c}\left[\frac{\partial \lambda}{\partial \beta}(q(1-\beta) - c) - \lambda q\right]$$

By definition of $\lambda(\beta)$, $\frac{\partial \lambda}{\partial \beta} \to \infty$ when $\beta \to 0$, which makes $\frac{\partial x^*}{\partial \beta} > 0$ when $\beta \to 0$, hence the fraction of paid placements $1 - x^*$ falls. For example, consider $\lambda(\beta) = k \cdot \beta^m$ (where $m < 1$): we see that $\frac{\lambda}{\beta} \to \infty$ when $\beta \to 0$, hence the first order condition is positive. That means, when β is small enough, the search engine can increase the bias a little bit to increase its profit. $\square$

An implication of this result is that there exists an optimal bias β^* for some functional forms. The next result describes how the search engine's optimal fraction of paid placement changes with the extent of bias.

PROPOSITION 2. *When $\beta \le \underline{\beta}$, then an increase in β leaves x^* unchanged at zero. When $\beta > \underline{\beta}$, an increase in β causes a decrease in the optimal fraction of paid placements.*

This result can be easily seen from Eq. 3 and Lemma 3. The change of the search engine's revenue will be determined by proposition 1.

One of the controllable factors of the information search engine is its quality of service, which is determined by the size of the database, the algorithms, and the user interface. Hence the search engine can improve quality via investments in these areas. Intuitively, the search engine could give up some placement revenues, improve q, attract more customers and get more advertisement revenues. What is the tradeoff between the search engine's quality and the placement revenue?

PROPOSITION 3. *An increase in the search engine's quality q allows it to increase the fraction of paid placement (level of independence x^* goes down), increasing its placement revenues and total profits. The search engine's market coverage increases as well, hence an increase in q increases surplus for all players.*

PROOF. In Eq. 3, note that an increase in q increases the RHS, thus increasing the fraction $(1 - x^*)$. We see that

$$\frac{\partial \pi_2^*}{\partial q} = \frac{d\pi_2}{dq} + \frac{d\pi_2}{dx^*}\frac{dx^*}{dq}$$

which is positive since $\frac{d\pi_2}{dq} > 0$ and the remaining two terms are negative.

To analyze the change in market coverage $\mathcal{M}$ as q increases, note that

$$\frac{\partial \mathcal{M}^*}{\partial q} = \frac{1}{2c}\left[2(1+\beta) - \frac{\lambda(\beta)}{(\lambda(\beta)+a)(2-x)}\left(2(1+\beta) - \frac{c}{q}\right)\right]$$

which is positive since $\frac{\lambda(\beta)}{(\lambda(\beta)+a)(2-x)} < 1$. Hence the search engine's market coverage and revenues from the third party $\pi_1 = a \cdot \mathcal{M}$ increase, hence its total profit increases. Since both fractions $\mathcal{M}$ and $1-x$ increase with q, the total surplus for both users and content providers also increases. $\square$

In general, to increase placement revenues, the search engine must increase its fraction of paid placement, but this increases users' disutility and reduces demand and advertising revenues. An increase in q, however, compensates for the increased disutility from increased paid placement. Hence the search engine is able to increase its placement revenues and yet increase total profits.

Finally, we consider the impact of per user profit. How do changes in a affect the search engine's paid placement strategy?

PROPOSITION 4. *An increase in the per user profit a allows the search engine to increase its degree of independence, so that the fraction of paid placements $1-x^*$ decreases while the placement fee γ increases. As a result, the search engine increase its market coverage $\mathcal{M}$ and total profits π.*

PROOF. From Eq. 3, an increase in a decreases the RHS, thus increasing x^*. Since $1-x^*$ decreases, we can infer from Eq. 2 that γ increases. That $\mathcal{M}$ increases is obvious from Eq. 1. The search engine's revenues brought by users and total profits increase. $\square$

To understand this result, consider the search engine's tradeoff between its two revenue sources. As a increases, the potential for advertising revenue increases, hence a partial sacrifice of revenues brought by users imposes a greater cost on the getekeeper. Therefore, it reduces its level of paid placement in order to provide greater utility to users, and captures a greater percentage of potential advertising revenues.

5. CONCLUSION

This article considers a paid placement strategy for search engines. On the one hand, paid placement appears to be a financial necessity, embraced by most major Web search engines. On the other, paid placement can hurt the search engine's market share and its potential for revenues brought by users. We have developed a mathematical model for optimal design of a paid placement strategy, examined this tradeoff and analyzed sensitivity of the placement strategy to users' perceived disutility, the service quality of the gatekeeper, and the advertising rate.

Our preliminary results are as follows. We show that the negative impact of paid placement on users causes the search engine to set paid placements at a below-ideal level. However, when disutility for paid placement is quite low (though not zero), the search engine can maintain its ideal placement revenues. We find that an increase in the search engine's quality of service allows it to improve its utilization of paid placement, moving it closer to the ideal; this also increases surplus for all players. However, an increase in the advertising rate motivates the search engine to increase market share by reducing further its reliance on paid placement and fraction of paying providers. As consumers get a better understanding of the factors underlying paid placement, the search engine would likely need to spend heavily on marketing campaigns in order to minimize users' perceived disutility for paid placement.

While this research is set in the context of Internet search engines, our model and results apply more generally to many other contexts that share similar characteristics as search engines. This broader category is often called information gatekeepers, that intermediate between a set of users (or buyers, or consumers) and a set of products (or content providers, or vendors). Baye & Morgan [1] argue that modern markets for information tend to be dominated by "information gatekeepers" that specialize in collating, aggregating, and searching massive amounts of information available on the Web – and can often charge consumers, advertisers, and information providers, for their ability to acquire and transmit information. Wise & Morrison [15] emphasize the increasing role of information gatekeepers in today's economy, noting that in business-to-business markets, "value has shifted from the product itself to information about the product." Specific categories of information gatekeepers to which our work applies include recommender systems (e.g., at `Amazon.com`), comparison shopping services (e.g., `mySimon.com`), e-marketplaces and exchanges (e.g., `FreeMarkets`), and more traditional information gatekeepers such as investment advisors and television networks. Like search engines, many information gatekeepers generate user-based revenues, but also seek to obtain revenues from their provider-base by offering some form of preferential placement. For example, some Internet booksellers are influenced by advertising fees in determining their *bestseller* lists. Similarly, certain Internet exchanges provide preferential service (such as real time notification or favorable recommendation to buyers) to some clients in return for higher fees.

We are pursuing extensions of this work, including a formal derivation of the optimal bias, generalization of demand assumptions, and elimination of free placement by the gatekeeper. Our models can be extended to examine conditions under which the information gatekeeper will begin to charge users, and specifically the case where the gatekeeper differentiates between users by offering two versions: a fee-based premium service with no bias in the query results, and a free basic version with paid placement bias. The fee-based premium version will bring additional user revenues to the search engine, however it may reduce placement revenues because paid placement becomes less attractive to content providers. In addition, the search engine's market coverage and placement fee may change as well, and the models can be used to determine if it is optimal for the gatekeeper to offer differentiated service. Similar models can be developed to examine the impact of differentiation based on advertising. Some search engines have already began to offer fee-based premium search services that contain no advertising. If this is the trend, it may eventually change people's view of Internet search engines as a free resource for fair information.

6. APPENDIX

Meta Search	Paid Links	Total Links	% Paid
Dogpile	30	35	86
qbsearch	66	98	67
MetaCrawler	13	25	52
Mamma	6	15	40
Search.com	10	29	34
ProFusion	2	14	14
Ixquick	1	10	10
Vivisimo	0	20	0

Table 1: `Paid placement in metasearch engines. what percentage of the links shown on the first page of results from a meta search service were paid listings. The table shows what percentage of the links shown on the first page of results from a meta search service were paid listings. The search query was "canada," done using each meta search service's default settings. http://www.searchenginewatch.com/-sereport/01/05-metasearch.html`

7. REFERENCES

[1] M. R. Baye and J. Morgan. Information gatekeepers on the internet and the competitiveness of homogeneous product markets. *American Economic Review*, 91(3):454–474, June 2001.

[2] K. Bharat and M. Henzinger. Improved algorithms for topic distillation in a hyperlinked environment. In *SIGIR Conference on Research and Development in Information Retrieval*, pages 104–111, 1998.

[3] H. K. Bhargava and V. Choudhary. Economics of electronic intermediary with aggregation. Technical report, Carnegie Mellon University, 2001.

[4] C. J. Corbett and U. S. Karmarkar. Optimal pricing strategies for an information intermediary. Technical report, UCLA Andersen Graduate School of Management, Working Paper # 99-001, 1999. http://www.anderson.ucla.edu/acad_unit/dotm/information/papers/99-001.p
df.

[5] R. M. Dewan, M. L. Freimer, and J. Zhang. Managing web sites for profitability: Balancing content and advertising. Technical report, University of Rochester, 2001. http://rdewan.ssb.rochester.edu/re.htm.

[6] D. Dreilinger and A. E. Howe. Experiences with selecting search engines using metasearch. *ACM*

Transactions on Information Systems (TOIS), 15(3):195–222, July 1997.

[7] J. Gabszewicz, D. Laussel, and N. Sonnac. Tv-broadcasting competition and advertising. Technical report, 1999. http://www.core.ucl.ac.be/services/psfiles/dp00/dp2000-6.pdf.

[8] A. Goodman. Why search engines must continue to be referees. *www.traffick.com*, November 2000.

[9] M. A. Hearst. TileBars: Visualization of term distribution information in full text information access. In *Conference Proceedings on Human factors in computing systems*, pages 59 – 66. ACM Press/Addison-Wesley Publishing Co, 1995.

[10] A. Kruger, C. L. Giles, F. M. Coetzee, E. G. G. W. Flake, S. Lawrence, C. Omlin, A. Agah, J. Callan, and E. Rundensteiner. Deadliner: Building a new niche search engine. In *Proceedings of the Ninth International Conference on Information and Knowledge Management*. ACM, 2000.

[11] S. Lawrence and C. Giles. Context and page analysis for improved web search. *IEEE Internet Computing*, 2(4):38–46, July/August 1998.

[12] S. Lawrence and C. L. Giles. Accessibility of information on the web. *Nature*, 400(6740):107–109, 1999.

[13] R. McLuhan. Search for a top ranking. *Marketing*, page 47, October 2000.

[14] A. Singhal and M. Kaszkiel. A case study in web search using trec algorithms. In *Series Proceedings of the The tenth international World Wide Web conference on World Wide Web*, pages 708– 716. ACM Press, 2001.

[15] R. Wise and D. Morrison. Beyond the exchange: The future of b2b. *Harvard Business Review*, 78(6):86–96, November-December 2000.