The Role of Search Engine Optimization in Search Rankings

Zsolt Katona (with Ron Berman)

University of California
Berkeley
Haas School of Business

TELECOM ParisTech

September 15, 2010
Introduction

- **Search Advertising**
  - Web sites bid for sponsored links on the “results” page of a search engine.
  - Sponsored links are assigned to sites based on their (pay-per-click) bids for specific keywords.
  - Over 11% of advertising money is spent online, over 40% of that on search advertising

- **Search Engine Optimization**
  - Websites can also affect their position on the organic list.
  - “Black hat” vs. “White hat”
  - SEO is over a billion $ business
BMW 5er Limousine - Automobile Leidenschaft.

Details

Produktinformationen

Limousine | Touring
Coupé | Cabrio

BMW 5er

Limousine | Touring

BMW 3er

BMW 1er

BMW 6er

Coupé | Cabrio

BMW 7er

BMW X3

BMW X5

BMW Z4

BMW M

Neue Automobile direkt

Fahrzeug konfigurieren
Infomaterial bestellen
Händler suchen
Probefahrt vereinbaren
Finanzierung berechnen

Weitere Specials

BMW 1er
BMW 5er Limousine
BMW 6er Coupé
BMW 6er Cabrio
BMW 7er
BMW X3
BMW X5
BMW Z4
BMW M5
BMW M6
BMW Z4 M

Gewährleistungen

Informationen zu den Gewährleistungen beim Kauf

The Role of Search Engine Optimization in Search Rankings
BMW Neuwagen


Ein BMW Neuwagen gesucht?

Google Bombs

The Role of Search Engine Optimization in Search Rankings

Biography of President George W. Bush
Biography of the president from the official White House web site.
www.whitehouse.gov/president/gwbbio.html - 29k - Cached - Similar pages
Past Presidents - Kids Only - Current News - President
More results from www.whitehouse.gov »

Welcome to MichaelMoore.com!
Official site of the gadfly of corporations, creator of the film Roger and Me and the television show The Awful Truth. Includes mailing list, message board, ...
www.michaelmoore.com/ - 35k - Sep 1, 2005 - Cached - Similar pages

BBC NEWS | Americas | 'Miserable failure' links to Bush
Web users manipulate a popular search engine so an unflattering description leads to the president's page.
news.bbc.co.uk/2/hi/americas/3298443.stm - 31k - Cached - Similar pages

Google's (and Inktomi's) Miserable Failure
A search for miserable failure on Google brings up the official George W. Bush biography from the US White House web site. Dismissed by Google as not a ...
searchenginewatch.com/sereport/article.php/3296101 - 45k - Sep 1, 2005 - Cached - Similar pages
Google Bombs

The Role of Search Engine Optimization in Search Rankings

THE ADMINISTRATION - President Barack Obama
WhiteHouse.gov is the official web site for the White House and President Barack Obama, the 44th President of the United States.
www.whitehouse.gov/administration/President_Obama/- Cached - Similar -  

Google search for "worst failure ever" shows results for WhiteHouse.gov.
Common Black Hat SEO techniques

- Content spam
  - Cloaking
  - Keyword stuffing
  - Hidden or invisible unrelated text
  - Doorway pages
  - Scraper sites

- Link spam
  - Link farms
  - Hidden links
  - Spam blogs

- UGC spam
  - Comment spam
  - Wiki spam
Research Questions

- What are sites’ incentives to invest in SEO?
- How does SEO affect the results?
- Do (which) sites profit from it?
- What is the relationship with sponsored links and search engine revenues?
- How does searcher behavior matter?
- Conventional wisdom: SEO is good for low quality sites, bad for SE and consumers.
Related Literature

- **Search Advertising**
  - Chen and He (2006)
  - Jeziorski and Segal (2009)
  - Goldfarb and Tucker (2008), Rutz and Bucklin (2007)
  - Yao and Mela (2009), Ghose and Yang (2009)

- **Role of organic links**
  - Katona and Sarvary (2010), White (2009), Xu, Chen and Whiston (2009),
  - Yang and Ghose (2010)

- **Non-desirable behaviors**
  - Wilbur and Zhu (2009a, 2009b)

- **All-pay auctions**
  - Hillman and Riley (1987)
Model

- Players: Search Engine (SE), $n$ Web sites
- Site $i$ has relevance (quality) $q_i$ with $q_1 \geq q_2 \geq \ldots \geq q_n$.
- Relevance: consumers are satisfied with probability $q_i$ if clicking on link $i$.
- SE displays $k$ organic links and $l$ sponsored links.
- SE places the best links on the organic side and maximizes revenue on the sponsored side. Total traffic ($T$) is an increasing function of expected consumer utility $f(U)$.
- Consumers:
  - with probability $\psi$, click on the best link (conscientious clicking)
  - otherwise, click on the $i$th organic link ($i$) with prob $\gamma \beta_i$ and on the $j$th sponsored link with prob $(1 - \gamma) \beta_j$ (random clicking)
- Site $i$ gets utility $R_i(t)$ from $t$ visitors, where $r_i() = R_i'(t)$ is decreasing.
SE cannot measure $q_i$ perfectly,
error term: $\varepsilon_i$ (i.i.d., centered around 0)

First, SE assigns (public) scores $s_i^S = q_i + \sigma \varepsilon_i$.

After observing these, sites can invest in SEO: $b_i$.

Final scores: $s_i^F = s_i^S + \alpha b_i$.

Effectiveness of SEO: $\alpha$ (cost: $\frac{1}{\alpha}$).

SE orders sites according to $s_i^F$. 
Simple case: $n = 2, k = 1, l = 0$

- Two sites: $q_1 \geq q_2$, $v_i = R_i(f(q_i))$, one link, $\varepsilon_i = \pm 1$.
- When there is no SEO ($\alpha = 0$), $a_i^S = a_i^F = q_i \pm \sigma$.
- If $\sigma > (q_1 - q_2)/2$ then 1 wins with prob 3/4.
- Probability of desired outcome $P(\alpha) = Pr(1 \text{ wins}|\alpha)$, that is, $P(0) = 3/4$.
- Expected traffic $ET(\alpha) = P(\alpha)(f(q_1) - f(q_2)) + f(q_2)$.
- What is the optimal $\alpha$?
- How does efficiency change with $q_1, q_2, v_1, v_2$?
Relation to all-pay auctions (contests)

- Players compete for an item, highest bid wins, full information. Their investment is sunk.
- Two players $v_1 > v_2$: No pure strategy eq.
- Mixed strategies: both players mix between 0 and $v_2$.
- 1 wins with probability $1 - \frac{v_2}{2v_1}$.
- Expected payoffs $v_1 - v_2$ and 0.

With headstart $h$:

- Player wins if bid + headstart is higher than others’.
- If $v_1 + h > v_2$, player 1 has a higher chance to win.
- Expected payoffs $\min(v_1 - v_2 + h, v_1)$ and 0.
Results

- There always exists an \( \hat{\alpha} > 0 \) such that \( P(\hat{\alpha}) \geq P(0) = \frac{3}{4} \).
- If \( v_1 > (3/2)v_2 \) then for any \( \sigma > (q_1 - q_2)/2 \) there is an \( \hat{\alpha} > 0 \) such that \( P(\hat{\alpha}) > P(0) = \frac{3}{4} \).
- If the most relevant site values traffic highly, some positive level of SEO always improves the efficiency

\[
R_1(f(q_1)) > \frac{3}{2} R_2(f(q_2))
\]

- In any other case, low levels of SEO do not hurt.
The Role of Search Engine Optimization in Search Rankings
Comparative statics

- $P(\alpha, \sigma, v_1, v_2, q_1, q_2)$ is increasing in $v_1, q_1$ decreasing in $v_2, q_2$.
- The more relevance is aligned with site valuation, the more efficient the ranking.
- Intuition: the auction is a mechanism that favors players with high valuation. If these coincide with the most relevant sites, then the ranking is improved.
- If the non-relevant site has a high valuation $P(\alpha)$ can be reduced to 0.
- Spammer sites can take advantage (low relevance, high valuation).
- In reality, some search engines accept paid links in the organic list.
How does the optimal SEO effectiveness level change with the variance of the measurement error?

Optimal SEO for search engine: \( \hat{A} \).

Result: If \( \nu_1 > (3/2)\nu_2 \) then \( \hat{A} = \hat{A}(\sigma) \) is increasing in \( \sigma \).

Intuition: a more effective correction mechanism may be useful if there is more error.

Implication: Investment in search algorithms and against SEO are complements
The Role of Search Engine Optimization in Search Rankings
The Role of Search Engine Optimization in Search Rankings
So far...

- We have shown that SEO might be good for search engine and consumers, if high quality sites value consumer high enough.

What about sites? Do they profit from SEO?

- Site 2 (lower quality) always loses with more SEO if $v_1 > v_2$.
- Site 1 is always strictly better off if $v_1 > 4v_2$ or $\sigma < \frac{v_1}{v_2} \frac{q_1-q_2}{2}$.
- If valuations are close, SEO makes sites fight hard for the top spot.
Timing: After the SEO process, second-price auction for the sponsored link.

Sites bid given their organic traffic. If marginal valuation for traffic is constant, they are independent.

If a site loses on the organic side, it compensates on the sponsored side.

If $R_1(f(q_1)) > (3/2)R_2(f(q_2))$ and $\gamma < 1$ is high enough, there is $\hat{\alpha}$, s.t. $ET(\hat{\alpha}) > ET(0)$. 
What about SE profits?

- Controversy: would search engines want to provide low quality sponsored links to gain sponsored revenue?

\[ E_{\pi}(\hat{\alpha}) > E_{\pi}(0) \text{ iff } R_1(f(q_1)) - R_1(\gamma f(q_1)) \geq R_2(f(q_2)) - R_2(\gamma f(q_2)). \]

- Higher traffic results in higher profits if
  - \( |r'_1| \) is small: Revenues don’t decline rapidly with traffic,
  - \( q_1/q_2 \) is close to 1: The difference between sites is not big, or
  - \( f'(\cdot) \) is not too high: consumers are not too sensitive to expected quality.
Conscientious search ($\psi > 0$)

\[ r_1(t) = 1 - t, \quad r_2(t) = r(1 - t). \]

- Consumers can tell which is the best search result without trying.
- \( ET \) is increased by SEO when \( r \) is low and \( \gamma \) is high or when \( r \) is high and \( \gamma \) is low (increasingly the later as \( \psi \) increases)
- \( E\pi \) is increases with \( ET \) when \( \gamma < \hat{\gamma} \), where \( \hat{\gamma} \) is decreasing in \( r \) and \( \psi \).
SEO benefits ($\psi = 0$)
SEO benefits ($\psi = 0.15$)
The Role of Search Engine Optimization in Search Rankings

SEO benefits ($\psi = 0.30$)
SEO benefits ($\psi = 0.45$)
SEO benefits ($\psi = 0.6$)

The Role of Search Engine Optimization in Search Rankings
SEO benefits ($\psi = 0.75$)
General model

- Multiple organic links
- Multiple players
- Error has a finite support
- $\bar{e} < \frac{q_1 - q_k}{k}$.
- Results are similar: some positive level of SEO can improve the organic ranking.
Conclusion

- We model the economic incentives behind SEO.

- We show that some positive level of SEO can be beneficial to
  - the search engine by increasing traffic and even profits
  - consumers
  - but generally not to sites, who have to defend their position

- SEO and sponsored links
  - Sites can go for either or both.
  - Higher quality organic links results in higher sponsored revenues when consumers trust the SE, valuations are different, there is enough click on sponsored links.
Future directions

- Different assumptions on timing and information
- Content investment / white hat SEO
- More active search engine
- Dynamics
- Heterogeneous consumers