I don't have all the answers, but I'm pretty sure these are the right questions. Finding ways to honestly connect marketing messages with interactive media creates openings for the sort of two-way, real-time relationships that marketers dream of, but a lot of these messages won't look much like 20th-century marketing. Media pundits still claim that Internet ads won't work until you can make someone laugh or cry. They're at least half wrong. Laughing and crying are fine – the Digital Age hasn't repealed all human emotional triggers – but in an info-rich empowered consumer world, creativity will require not just an emotional message, but a useful and informative one.

The best news, though, is that in this new world, profitability can take center stage. "I know half of my advertising is wasted," John Wanamaker famously said; "I just don't know which half." That sentiment will finally be obsolete in a digital media world where return on investment (ROI) is transparent and advertising is a profit center that's "always on." Knowing their margins on each consumer they reach will let advertisers break through the ceiling of the non-empowered world. Today's marketing industry spends $250 billion a year. How much will it spend once advertisers are sure that neither half is wasted, that they can do twice as much marketing that's twice as effective? The answer starts with the empowered consumer, includes search advertising and ends with an industry that's growing larger, not smaller. The flip side of fear is opportunity.

An example of this changed environment brought about by the Internet is the Google AdWords program, which uses keywords to target precisely ad delivery to Web users seeking information about a particular product or service. Like other programs described in this special section, AdWords is based on cost-per-click (CPC) pricing, so advertisers only pay when an ad is clicked on. Advertisers can take advantage of an extremely broad distribution network and choose the level of support and spending appropriate for their business. Such tools can empower consumers and allow advertising dollars to be spent more effectively.

**For Further Information**

Online: http://adwords.google.com
Email: Mike Nelsen at mnelson@google.com

# Click Fraud

## by Brendan Kitts, Benjamin LeBlanc, Ryan Meech and Parameshvyas Laxminarayan

*Brendan Kitts, Benjamin LeBlanc, Ryan Meech and Parameshvyas Laxminarayan, all of iProspect, 311 Arsenal Street, Watertown, MA. For questions, email: bkitts@excite.com.*

iProspect manages millions of dollars of advertising budgets on pay-per-click (PPC) auctions for many of the largest companies in the world. We have developed our own bidding agent and tracking system. We are Ambassadors for Yahoo!'s Paid Inclusion program. So you can imagine our surprise when we seemed to find an error in Yahoo!'s cost accounting.

Both the costs and clicks quoted to us by Yahoo! were lower than our independent tracking system was reporting. The under-charge was about 12.5%

Had Yahoo! made a mistake?

Little did we know, but Yahoo! had been efficiently removing huge numbers of clicks before they reached their customer reports. Think of this as like an Enron document shredding operation in reverse. They were removing fraud before it hit their advertisers. We had stumbled across a way of spying on Yahoo!'s Click Fraud Protection System in action.

Overture has created a truly revolutionary, market-driven, information retrieval system. The PPC model is fascinating because the relevance of paid search appears to be as good as classical information retrieval systems (see for example, Jansen et. al., 2005). We applaud Overture's efforts in fighting the newly emerging problem of click fraud. However, the revelation that as many as 1 in 10 clicks are fraudulent – even if they are being detected by the search engine – raises many difficult questions:

- Has all of the fraud been caught?
- What is it about paid search that makes it susceptible to fraud?
- What is the impact of the fraud?

- If fraud continues to grow, what is the future of the PPC model?

This article introduces readers to the problem of "click fraud," examines the scope of the problem and discusses methods for overcoming it.

## What is Click Fraud?

Click fraud, the intentional clicking on PPC advertisements, where the perpetrator has no intention of buying the products or services advertised, is one of the fastest growing problems on the Internet. Click fraud generally falls into two categories – clicking on competitors and network fraud.

Clicking on competitors occurs when a company purposely clicks on a competitor so as to cost them money, use up their daily budgets and force them off the auction.

John Carreras, president of Impact Displays, says that he knew he had a click fraud problem when he went to a major trade show. He returned to discover that his ad expenses had been 50% lower than normal. He surmised that his competitors were all at the trade show and weren't able to click on his ads (Eroshenko, 2004).

Olsen (2004) refers to a company executive who enjoys clicking on his competitor's ads. "It's an entertainment," he says. "Why do you run into a store without putting a quarter in the meter? You know it's wrong, but you do it."

Network fraud occurs when website owners click on their own banner advertisements in order to generate revenue from the search engine that is serving the banner advertisement. Most people committing network fraud are small-time operators. However, there are also some professionals.

Auctions Expert International LLC (Houston) allegedly ran an operation of up to 50 people to click on its own Google ads, which allowed it to generate about $50,000 in ad revenue (Blakely, 2004). The India Times reported that a "secret army" of housewives, graduates and working professionals in India were being paid up to $200 a month to click on Internet advertisements (Vidyasagar, 2004).

## An End to Internet Advertising?

How could a few clicks do any harm? The doomsday scenario goes something like this. Since ad-clicking is easy and lucrative, an increasing number of fraudsters begin to take advantage of the program. PPC auctions are eventually flooded with fraudulent clicks. Awash with clicks that cost advertisers but generate no purchases, advertisers are crippled by massive advertising costs with almost no return. They stop or reduce their participation in PPC. Search engines lose their fees and can no longer support their operations. The industry turns upon itself as advertisers sue the search engines for fraud. Like a massive star collapsing into a black hole, a mass of fraudulent clicks could cause the implosion of the industry.

High-flying Google executives are understandably concerned. George Reyes, chief financial officer of Google, says: "Click fraud is the biggest threat to the Internet economy" (Delaney, 2005). Stephen Messer, CEO of LinkShare, expresses similar sentiments: "Click fraud is rampant and staggering.... it could wipe out ROI in search marketing in 2005."

## How Prevalent Is Click Fraud?

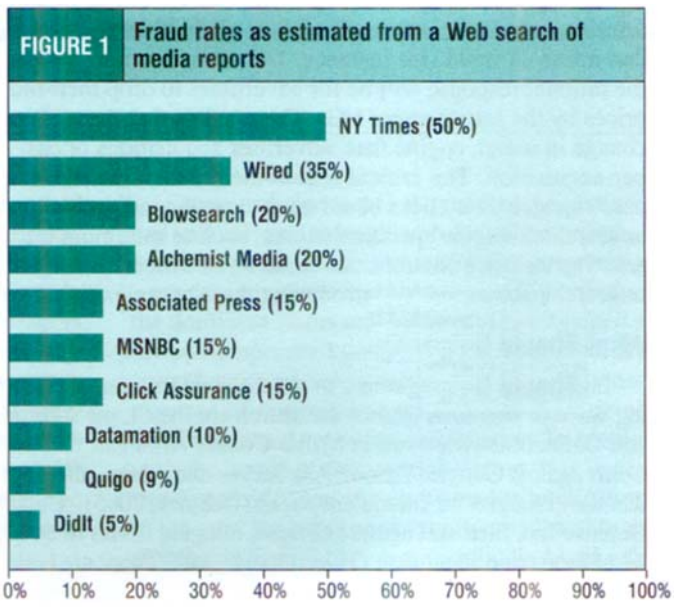
It seems that almost every story on click fraud quotes some expert with an estimate of click fraud in the industry. What are the facts? We developed three methods for estimating the level of click fraud in the industry.

***Statistical methods.*** Every website has an expected conversion (purchase) rate, $a$, that can be calculated by dividing conversions by their clicks.

Now let's consider the activity from one particular user, which we identify by their Internet Protocol (IP) address. A user clicking on an advertisement a large number of times and not converting (purchasing) is like flipping a coin and repeatedly having it come up tails every time. The probability of this sequence occurring at random can be calculated. A user who converts significantly less often than expected is regarded as probably fraudulent. (For more details on these methods, see Notes at end of article.)

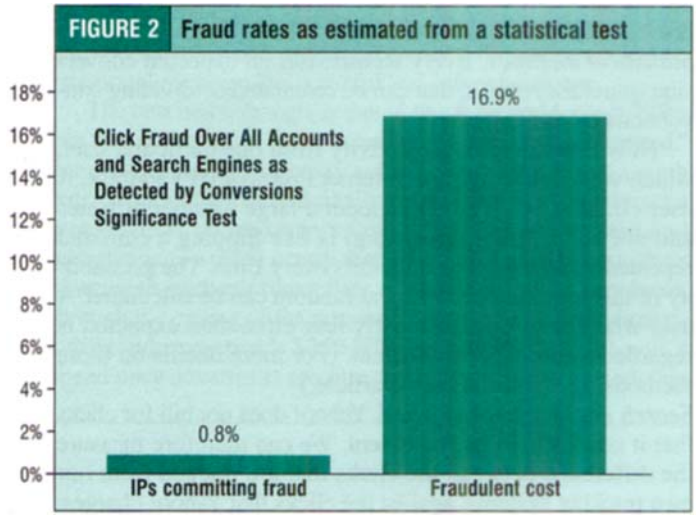
***Search engine removed fraud.*** Yahoo! does not bill for clicks that it considers to be fraudulent. We can therefore measure the difference between the clicks that are tracked from our own tracking systems, against the clicks that Yahoo! charges. This undercharge represents the amount of fraud that Yahoo! is detecting.

***Consensus estimate from popular media.*** In recent years, economists have gained a deeper appreciation for the wisdom of crowds. We ran a Web search in May 2005 and found every article we could on click fraud. Every time the story quoted an estimate of the rate of click fraud we recorded it. We then took the median value. The results are shown in Figure 1.



## Disposition of Fraud

The three methods estimated fraud across all industries at 17%, 12.5% and 15%, respectively (for methods, see the note at end of article.). The Internet protocol addresses (IPs) that we flagged as fraudulent through our statistical test comprised less than 1% of all IPs (Figure 2).




## It's Good To Be Rational

One would expect that if 15% of clicks were fraudulent, and the search engines were not offering rebates, then the search engine would generate 15% more revenue. However, a curious relationship called Ryan's theorem (Kitts, et. al., 2005) suggests that rational bidders may be completely unaffected by network click fraud.

If an advertiser is rational, its bid price for clicks should track the actual conversion value of the click. If there is a sudden influx of fraud (for instance, 1/G clicks are now valid), the rational response will be for advertisers to drop their bid prices by the same factor (1/G). The result is that there is no change in search engine fees, advertiser acquisitions or cost-per-acquisition. The critical requirement is that the bidders need to value their clicks based upon current conditions. As a note, search-engine specific features, such as minimum bids and discrete price controls, can break Ryan's theorem, but for brevity we have avoided introducing these complexities.

## What Should Be Done?

In 2005, we are beginning to see the crest from an oncoming wave of litigation against the search engines. Lane's Gifts and Collectibles filed suit in Miller County Arkansas Circuit Court against Google, Yahoo!, Ask Jeeves and others, alleging that they charged for fraudulent clicks (Delaney, 2005). Click Defense Inc. filed suit against Google, alleging losses of over $5 million from fraudulent clicks. Google and Yahoo! are both working overtime to hand out refund credits. Will any of these actions put an end to the problem of click fraud?

We have seen that rational bidding – accurate pricing – can protect advertisers from network fraud. Sadly, it cannot fix all sources of fraud. The most rational bidder in the world cannot survive if it has been targeted by competitor clicking. In order to eliminate all forms of fraud, two options seem promising.

The pay-per-purchase (PPP) model could be adopted. Under PPP, the search engines would only be paid after the advertiser achieves a conversion. PPP is undesirable because advertisers would report conversions, and so it opens the door to advertiser fraud. It would also be less lucrative for search engines, since a large number of irrational bidders who are not valuing their clicks properly today would suddenly become perfectly rational.

The second option involves no major change to the PPC model. Search engines could give advertisers the ability to block certain IP addresses from viewing their advertisement. A blocked searcher would still have access to the search engine's natural search results, as well as paid listings from other advertisers.

Whether customers are uninterested, fraudulent or like clicking on advertisements because they're not familiar with

## References


Blakely R. (2004, November 30), Google hits back at scam ad clickers. *Times online*. Available October15, 2005, at http://business.timesonline.co.uk/article/0, 9075-1381606,00.html

Delaney, K. (2005, April 6), Click fraud: Web outfits have a costly problem, marketers worry about bills inflated by people gaming the search ad-system. *Wall Street Journal*, A1.

Eroshenko, D. (2004, October 19). Click fraud: The state of the industry, *Pay Per Click Analyst*. Available October 15, 2005, at www.payperclickanalyst.com/content/templates/default.aspx?a=68&z=1

Jansen, B.J. & Renick, M. (2005). *An Examination of Sponsored Results for E-commerce Web Searching*. Technical Report, School of Information Sciences and Technology, The Pennsylvania State University. University Park, PA. Copy available upon request from jjansen@acm.org.

Kitts, B. Laxminarayan, P., LeBlanc, B. and Meech, R. (2005, June). A formal analysis of search auctions including predictions on click fraud and bidding tactics. *ACM Conference on E-Commerce – Workshop on Sponsored Search*, Vancouver, UK. June 2005. Available October 15, 2005, at http://research.yahoo.com/workshops/ssa2005/sched.html.

Olsen, S. (2004, July 19) Exposing click fraud. *CNET News.com*. Available October 15, 2005, at http://news.com.com/Exposing+click+fraud/2100-1024_3-5273078.html

Vidyasagar, N. (2004, May 3). India's secret army of online ad "clickers." *The Times of India*. Available October 16, 2005, at http://timesofindia.indiatimes.com/articleshow/msid-654822,curpg-1.cms

the Internet, the solution for advertisers is the same. They need to avoid showing their advertisements to those customers. Advertiser-initiated IP blocking would

- avoid search engines paying commissions to sites that are fraudulent;
- encourage websites displaying the advertisements to improve their quality;
- shift the fraud detection effort from one centralized authority, to thousands of interested advertisers. The information processing problem is even easier – while it is hard for a central authority to detect fraud, it is easy for advertisers to list their "non-converting" IP addresses.

The PPC model works because thousands of advertisers are targeting their advertisements to find converting customers and hide their listings from non-converting customers. Empower this network with the ability to block IPs, and sources of fraud should rapidly lose their traffic. At least, that's the theory.

### Note: Formulas for Calculating Fraud

***Statistical methods.*** Every website has an expected conversion (purchase) rate, $a$, that can be calculated by dividing conversions by their clicks. A user clicking on an advertisement a large number of times and not converting (purchasing) is like flipping a coin repeatedly and having it always come up tails. The probability of this occurring at random can be calculated using the binomial distribution, where $c_u$ are the number of clicks from user $u$, $A_u$ the number of conversions from the user, and $a$ is the conversion rate over all users. A user with a p-value less than a critical value of 0.01 will be regarded as probably fraudulent.

$$\text{fraud \%} = \frac{\sum_{u:p_u<0.01} c_u}{\sum_{i=1}^{N} c_i} \quad \text{where } p_u = bpdf(A_u, c_u, a) = \binom{c_u}{A_u} a^{A}(1-a)^{(c_u - A_u)}$$

***Search engine removed fraud.*** Yahoo! does not bill for clicks that it considers to be fraudulent. We can therefore measure the difference between the clicks that are tracked from our own tracking systems $c_i$, against the clicks that Yahoo! charges $c_i'$. This undercharge represents the amount of fraud that Yahoo! is detecting.

$$\text{fraud \%} = \frac{\sum_{j=1}^{M} c_j'}{\sum_{i=1}^{N} c_i}$$

The three statistics used in Figure 1 were (a) percentage of charges where the IP's clicks to conversions were $p<0.01$ under the null hypothesis, (b) percentage of rebated clicks and (c) media percentage estimate.

# The Value Implications of the Practice of Paid Search

by Michael Zimmer

*Michael Zimmer is a Ph.D. candidate in the department of culture & communication, New York University, The Steinhardt School of Education, 239 Greene Street, 7th Floor, New York, NY 10003. He can be reached by email at mtz206@nyu.edu.*

In his book *Technopoly*, Neil Postman remarked that "we are surrounded by the wondrous effects of machines and are encouraged to ignore the ideas embedded in them." It has been the goal of many scholars of technology to remove these blinders and critically explore the ideological biases embedded within our technologies and technical systems. Such scholars argue that technologies have, in varying degrees, certain social, political and epistemological biases; they tend to promote certain ideologies while obscuring others. Recently attention has been paid to how information technologies also have *ethical* and *value* biases.

Our knowledge tools – the particular techniques and technologies to assist with the collection, organization, classification and retrieval of information – are not immune to such ideological biases. Given the dominant status search engines have gained as the contemporary knowledge tool, it is crucial to consider the social, political and ethical consequences of our reliance on them for organizing, distributing and accessing information. In keeping with the theme of this special section, it is important to consider specifically the value implications of the growing practice of "paid search" within the search engine industry. For simplicity, I will refer