On Click Fraud

Simone Soubusta, Düsseldorf

Internet advertising is getting more and more important. As in every growing line of business, criminals try to profit from that growth. They do so by manipulating Pay-Per-Click (PPC) adverts with artificially generated clicks, a process which is known as Click Fraud (CF). CF is not difficult to commit and, consequently, it could threaten the existence of companies like Google. At present, there seems to be no way to either detect or prevent CF. It can, however, be altogether avoided by switching from PPC to other advertising schemes like Cost-Per-Action.

Über Klickbetrug

Internetwerbung wird immer wichtiger. Wie in jeder Wachstumsbranche versuchen auch hier Kriminelle von diesem Wachstum zu profitieren. Dies ist durch die Manipulation von Pay-Per-Click (PPC) Werbeanzeigen mithilfe künstlich produzierter Klicks möglich – eine Vorgehensweise, die als Klickbetrug (Click Fraud) bekannt ist. Klickbetrug ist nicht schwierig und könnte daher die Existenz von Firmen wie Google bedrohen. Zur Zeit scheint es keine Möglichkeit zu geben, Klickbetrug zu entdecken oder zu verhindern. Man kann Klickbetrug jedoch vermeiden, indem man PPC Werbung durch eine andere Berechnungsgrundlage, wie z. B. Cost-Per-Action, ersetzt.

1 Introduction

1.1 Internet Advertising

It is a truth universally acknowledged that the World Wide Web (WWW) has revolutionised our planet. Today the web allows us to communicate with people nearly anywhere in the world – and, indeed, with billions of people at once. It has changed the way we think, as well as the way we do business. But the web is not only a means of communication, it also contains a wealth of knowledge. Since the rise of the search engines, huge amounts of information have become readily accessible. The known web alone spans several billion pages.

Most of the information and services on the web come at no cost to the user. The publisher or webmaster, on the other hand, has to pay to provide them. In order to profit from their complimentary services, webmasters rent out space on their websites to advertisers. This is the main means of covering the costs of providing online services nowadays. The most famous example of this phenomenon is probably Google. The search engine accumulated considerable wealth by auctioning off advertising space next to their search results.

Advertising plays an important role for every company. “Most businesses operate with an advertising budget of 2-5 percent of their previous year’s gross sales” (Egelhoff, n.d.). Since its inception, advertising has thrived on mass media like newspapers and TV, where an ad could reach a wide audience. It comes as no surprise, then, that it has found the web a suitable playground.

1.2 Cyber Crime

The rise of the internet has also brought a completely new kind of crime: cyber crime. The most well-known type of cyber crime is *Phishing,* which means stealing sensitive information online. This can be anything from user name/password combinations to credit card information and the troublesome social security number in the U.S. The criminal gains this information by masquerading as an authority. In order to gain access to the PIN/TAN combination of a user’s bank account, for example, the criminal might write an e-mail to the client, in which he pretends to be the bank itself. He presents the user with a link that allegedly leads to the bank’s online banking login page, where the user is asked to log in for ‘maintenance reasons’ or something similar, but in reality the link leads to the phisher’s website (which imitates the look and feel of the bank’s site), allowing him to steal the information and, therefore, the user’s money. Two other types of cyber crime are *credit card fraud,* which the internet has made infinitely easier to commit, and *spam,* which doesn’t warrant further explanation.

However, there is another kind of cyber crime that the general public is not well acquainted with, but which is nonetheless detrimental to many people: so-called *click fraud,* which this paper will examine in detail.

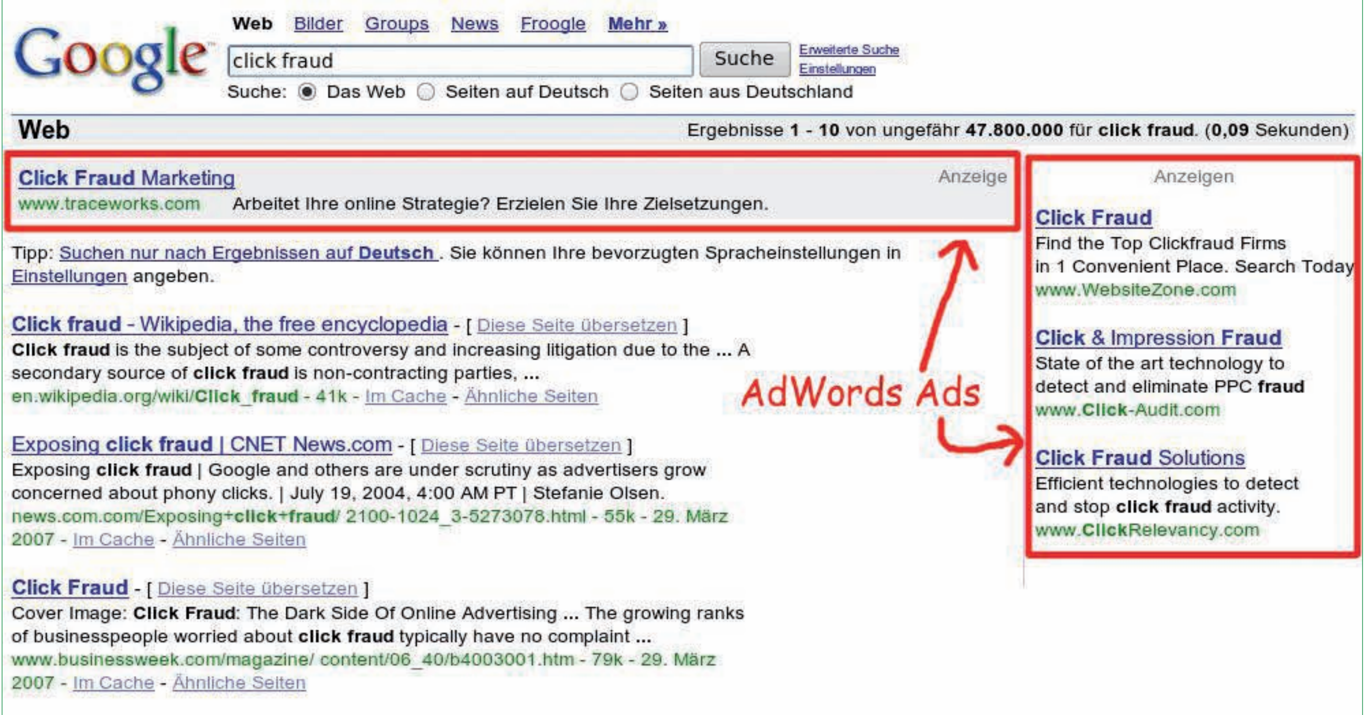
1.3 Structure of the Article

Chapter 2 introduces the reader to the concept of click fraud against the background of Pay-Per-Click advertising, followed by a short look at the legal ramifications. Chapter 3 expounds how click fraud is committed. An explanation of how a click can be faked is followed by two different techniques that are used to avoid detection: using a network of hijacked computers (a so-called botnet) to commit click fraud, and making visitors of a website unintentionally assist in the fraud. Chapter 4 presents different ways to detect click fraud, as well as ways to prevent click fraud altogether. Finally, chapter 5 concludes the article with a look at its result. (see Fig. 1)

2 What is Click Fraud?

In the past, online advertising used the Cost-Per-Impression model to charge for advertisements. The cost per impression is often measured in Cost Per Mile (CPM), that is, the cost of one thousand impressions of the ad. This advertising model was based on traditional TV and print advertising, where the advertiser is also charged on the basis of the number of times that the ad is viewed. Cost-Per-Impression is the preferred model of publishers (webmasters), because they are paid regardless of the effectiveness of an ad. However, search engines such as Google have given rise to the Pay-Per-Click (PPC) model of online advertising. In this type of arrangement, advertisers pay a certain amount of money to the publisher for every click on their ad (which leads to the advertiser’s website). That is to say, the costs are performance-dependent. Consequently, the model is favoured by advertisers, who get an interested visitor to their website for the money they pay.

In particular, Google’s AdWords and AdSense (see Fig. 1 and 2) services have added to the widespread use of PPC. With AdWords, advertisers bid in an auction for the placement of their ad next to the search results for a chosen keyword (or a combination of keywords). This form of advertising is also known as sponsored



Google

Web Bilder Groups News Froogle Mehr »

click fraud | Suche | Erweiterte Suche | Einstellungen

Suche: ● Das Web ○ Seiten auf Deutsch ○ Seiten aus Deutschland

Web | Ergebnisse **1** - **10** von ungefähr **47.800.000** für **click fraud**. (**0,09** Sekunden)

Click Fraud Marketing | Anzeige
www.traceworks.com Arbeitet Ihre online Strategie? Erzielen Sie Ihre Zielsetzungen.

Tipp: Suchen nur nach Ergebnissen auf **Deutsch**. Sie können Ihre bevorzugten Spracheinstellungen in Einstellungen angeben.

Click fraud - Wikipedia, the free encyclopedia - [Diese Seite übersetzen]
Click fraud is the subject of some controversy and increasing litigation due to the ... A secondary source of **click fraud** is non-contracting parties, ...
en.wikipedia.org/wiki/**Click_fraud** - 41k - Im Cache - Ähnliche Seiten

Exposing click fraud | CNET News.com - [Diese Seite übersetzen]
Exposing **click fraud** | Google and others are under scrutiny as advertisers grow concerned about phony clicks. | July 19, 2004, 4:00 AM PT | Stefanie Olsen.
news.com.com/Exposing+**click+fraud**/ 2100-1024_3-5273078.html - 55k - 29. März 2007 - Im Cache - Ähnliche Seiten

Click Fraud - [Diese Seite übersetzen]
Cover Image: **Click Fraud**: The Dark Side Of Online Advertising ... The growing ranks of businesspeople worried about **click fraud** typically have no complaint ...
www.businessweek.com/magazine/ content/06_40/b4003001.htm - 79k - 29. März 2007 - Im Cache - Ähnliche Seiten

Anzeigen

Click Fraud
Find the Top Clickfraud Firms in 1 Convenient Place. Search Today
www.WebsiteZone.com

Click & Impression Fraud
State of the art technology to detect and eliminate PPC **fraud**
www.**Click**-Audit.com

Click Fraud Solutions
Efficient technologies to detect and stop **click fraud** activity.
www.**Click**Relevancy.com

Figure 1. Google AdWords.

search. The obvious advantage for advertisers is that the user who sees the advertisement is already interested in the topic that is connected to the keyword (since he searched for it), hence there is a high chance that he is also interested in the product or service that is advertised next to it. A user who searches for the term *camera*, for example, might be shown an advertisement for a camera by Canon. In addition, the user also profits from this *targeted advertising*, since he is only shown ads for products he likely has an interest in.

AdSense, on the other hand, is an advertising network in which Google acts as a commissioner. Similar to AdWords, advertisers bid on certain keywords. Google then automatically analyses the content of websites that have signed up as publishers for AdSense and places the ad on those websites that contain content which is closely related to the keyword. Every time someone clicks on such an ad, the website owner gets a percentage of the proceeds which Google receives from the advertiser, and the rest is Google's commission.

From this point of departure, we can distinguish two types of click fraud: competitor click fraud and publisher click fraud.

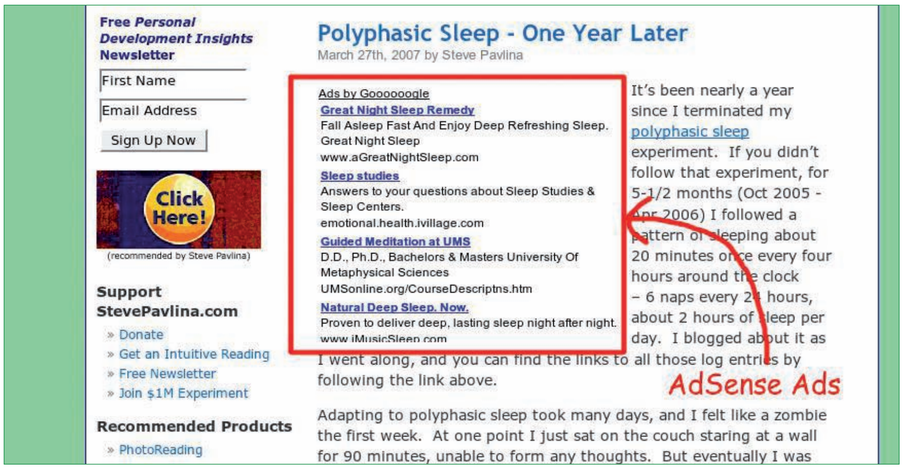
Competitor click fraud is not committed for any monetary gain (and, in fact, there is none), but rather in order to harm one's competitor. A prerequisite for the fraud is that the competitor in question has signed up as advertiser in a PPC scheme (e.g. Google AdWords). The fraudster, knowing that every click on the ads costs the competitor good money, clicks repeatedly on his ad to cause harm. He might perform this task himself or use more sophisticated means (see below).

Publisher click fraud is, as the name indicates, performed by the publisher of an ad, that is, by the owner of the website on which the ad (or, indeed, multiple ads) is displayed. It is committed solely for monetary gain. Since the publisher is paid for every click on an ad, it is in his interest to receive a high number of clicks. Thus, the fraudster artificially increases the number of clicks in order to make more money. Harmless dilettantes do this by clicking on the ads themselves or by asking their friends to do so. Professional fraudsters, on the other hand, either outsource the job to armies of 'clickers' in India (Vidyasagar, 2004) or China, or they automate the process entirely. This paper will concern itself primarily with the latter option.

According to Noogie C. Kaufmann from the University of Münster, *competitor click fraud*, as described above, is illegal under German law since it violates § 4 Nr. 10 UWG (purposeful hindrance of competitors) and § 826 BGB (immoral deliberate damnification) (Kaufmann, n.d.). Unfortunately, though, there hasn't been an original precedent yet to confirm this assumption. Nonetheless, competitor click fraud is a big problem for the harmed party, and several advertisers who participate in Google's AdWords program have sued Google "for their failure to prevent click fraud" (Hadjinian, 2006). Their point of attack is the contract between the two parties which states that only *actual clicks* will be charged, which enabled them to sue Google for breach of contract when they noticed that they were charged for fraudulent clicks as well as for actual ones (Hadjinian, 2006).

Publisher click fraud is, in most cases, simply a breach of contract. The Terms of Service of Google's AdSense program, for example, state: "You shall not, and shall not authorize or encourage any third party to: (i) directly or indirectly generate queries, Referral Events, or impressions of or clicks on any Ad, Link, Search Result, or Referral Button through any automated, deceptive, fraudulent or other invalid means, including but not limited to through repeated manual clicks, the use of robots or other automated query tools and/or computer generated search requests, and/or the unauthorized use of other search engine optimization services and/or software; [...] (iii) frame, minimize, remove or otherwise inhibit the full and complete display of any Web page accessed by an end user after clicking on any part of an Ad ("Advertiser Page"), any Search Results Page, or any Referral Page; (iv) redirect an end user away from any Advertiser Page, Search Results Page, or Referral Page; provide a version of the Advertiser Page, Search Results Page, or Referral Page that is different from the page an end user would access by going directly to the Advertiser Page, Search Results Page, or Referral Page; intersperse any content between the Ad and the Advertiser Page, between the page containing the Search Box and the Search Results Page, or between the Referral Button and the Referral Page; or otherwise provide anything other than a direct link from an Ad to an Advertiser Page, from the page containing the Search Box to the Search Results Page, or from the Referral Button to the Referral Page" (*Google AdSense Online Standard Terms and Conditions*, n.d.)

Based on these terms of service Google won a court battle against a participant of its AdSense program, Auctions Expert, in 2004 (Hadjinian, 2006). Although this case could be seen as establishing a precedent, it does not completely solve the publisher click fraud problem from a legal perspective, since the fraudster might not fall under the jurisdiction of the U.S. law.



Polyphasic Sleep - One Year Later

March 27th, 2007 by Steve Pavlina

Ads by Goooooogle

Great Night Sleep Remedy
Fall Asleep Fast And Enjoy Deep Refreshing Sleep.
Great Night Sleep
www.aGreatNightSleep.com

Sleep studies
Answers to your questions about Sleep Studies & Sleep Centers.
emotional.health.ivillage.com

Guided Meditation at UMS
D.D., Ph.D., Bachelors & Masters University Of Metaphysical Sciences
UMSonline.org/CourseDescriptns.htm

Natural Deep Sleep. Now.
Proven to deliver deep, lasting sleep night after night.
www.iMusicSleep.com

It's been nearly a year since I terminated my polyphasic sleep experiment. If you didn't follow that experiment, for 5-1/2 months (Oct 2005 - Apr 2006) I followed a pattern of sleeping about 20 minutes once every four hours around the clock – 6 naps every 24 hours, about 2 hours of sleep per day. I blogged about it as I went along, and you can find the links to all those log entries by following the link above.

Adapting to polyphasic sleep took many days, and I felt like a zombie the first week. At one point I just sat on the couch staring at a wall for 90 minutes, unable to form any thoughts. But eventually I was

Figure 2. Google AdSense.

3 Why Does Click Fraud Matter?

Google can easily be used to exemplify the explosiveness that click fraud presents today: This unquestionably famous billion-dollar concern makes 99 per cent of its turnover through Pay-Per-Click advertising. If click fraud is not countered in the near future, the backbone of its business threatens to collapse. Google's advertising revenues have risen from 6.07 billion US dollars in 2005 to 10.49 billion dollars in 2006. In both years (as well as in 2004) the advertising revenues accounted for 99 per cent of the overall revenues. In 2006, 60 per cent of the revenues (i.e. 6.29 billion dollars) were made through Google AdWords, a system that is susceptible to competitor click fraud, while the remaining 40 per cent (i.e. 4.2 billion dollars) were made through Google AdSense, a system that is susceptible to publisher click fraud (*Google Annual Report 2006*, 2007).

But Google is not the only one to profit from internet advertising. In the year 2005 the total revenue of internet advertising amounted up to 12.5 billion dollars. In 2006 it rose by 35 per cent to 16.9 billion dollars (*IAB Internet Advertising Revenue Report*, 2007). And the turnover in the first quarter of 2007 was an all-time high of 4.9 billion dollars (*Internet Advertising Revenues Soar*, 2007). Even though the numbers might look small compared to TV advertising revenues of 48.35 billion dollars in 2006 (*2006 TV Ad Revenue Figures,* 2007) and print advertising revenues of 46.6 billion dollars (*Mind the Gap*, 2007), the trend is clearly in favour of online advertising: TV advertising revenues recorded only a rise of 5.3 per cent and revenues of print advertising actually fell 1.7 %.

Although, as we will see later, it is not (yet) possible to accurately estimate the number fraudulent clicks, studies suggest that about 14 percent of all clicks aren't proper (Mills, 2006a; Mills 2006b). Consequently, companies like Google have two alternatives to avert the danger that click fraud presents for their business: Firstly, they can open up new business segments in order to put an end to their total dependency on their advertising customers. Secondly, they can stop the fraudsters before they cause any damage. However, in order to counter click fraud, one needs an understanding of how it works, which the following chapter will provide.

4 How it Works

4.1 Simulating a Click

At the heart of (automated) click fraud lies the simulation of a click on an advertisement. It follows, then, that in order to understand how click fraud works, we must first gain an understanding of the technology behind the advertisement. "Typical online advertisement services […] work by providing webmasters a snippet of JavaScript code to add to their pages. This code is executed by the web browser of a visitor to the site, and downloads ads from the advertiser's server at that time. The ad download triggers a rewrite of the frame in which the JavaScript appears, replacing it with the HTML code necessary to display the ads. When a user clicks an advertisement link, they "click through" the ad provider's server, giving the ad provider the opportunity to bill the client for the click. The user is then taken to the ad client's homepage" (Gandhi et al., 2006).

It is evident that a program which simulates clicks has to perform many of the same tasks as a browser. First, it has to execute the JavaScript code that retrieves the HTML code of the ad, as mentioned above. Then it has to parse this HTML code for links (indicated by a leading `href="` and a trailing `"`) and, lastly, it has to send an HTTP (Hypertext Transfer Protocol) request to the web server at the given URL (Uniform Resource Locator).

4.2 Distributed Click Fraud with Botnets

Simulating a click is, in effect, rather easy. However, detecting this kind of click fraud is just as easy. The reason for that is simple: when the program sends an HTTP request to the advertiser's (or commissioner's) server, the IP address of the computer making the request is transmitted in order to establish a connection between client and server. Thus, all it takes to detect the fraudster is for the advertiser to take a look at his log files, where a high number of requests from one IP address provides sufficient evidence of the fraud.

To avoid being discovered and to increase the efficiency of the fraud, the fraudster can distribute the program so that it does its work from all over the internet, with the help of a so called *botnet*. A botnet is a network of thousands or even millions (Keizer, 2005) of hacked computers, all of which do the bidding of the owner of the botnet.

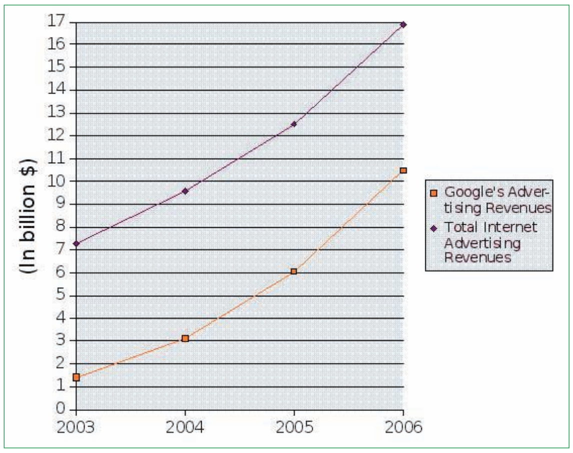


Figure 3. Web Advertising Revenues 2003–2006.
Sources: IAB Internet Advertising Revenue Report, 2007; Google Annual Reports.

4.2.1 Taking over a Computer

The targeted computer(s) can be compromised by exploiting security holes and vulnerabilities. The program which does the exploiting is commonly referred to as an exploit. Attackers either write these exploits themselves or, more commonly, use exploits for known security holes that are available on the internet. Most exploits utilise buffer overflows (Override, 2001).

“A buffer overflow occurs when a program or process tries to store more data in a buffer (temporary data storage area) than it was intended to hold. Since buffers are created to contain a finite amount of data, the extra information – which has to go somewhere – can overflow into adjacent buffers, corrupting or overwriting the valid data held in them. Although it may occur accidentally through programming error, buffer overflow is an increasingly common type of security attack on data integrity. In buffer overflow attacks, the extra data may contain codes designed to trigger specific actions […]” (*Buffer Overflow*, n.d.).

Once the attacker has decided upon a specific exploit, he begins scanning (IP) address blocks for systems which fulfil the requirements of the exploit, i.e. that are running a certain version of an operating system as well as the vulnerable program. One rather popular way to do these scans is to use the open source program Nmap, which “uses raw IP packets in novel ways to determine what hosts are available on the network, what services (application name and version) those hosts are offering, what operating systems (and OS versions) they are running, what type of packet filters/firewalls are in use” (*Nmap Security Scanner*, n.d.).

Once a vulnerable system has been detected, the attacker uses the exploit to gain remote access to that computer. He then deposits a program which will contact the botnet’s command and control centre, allowing the machine to be controlled from a central point.

4.2.2 Command & Control

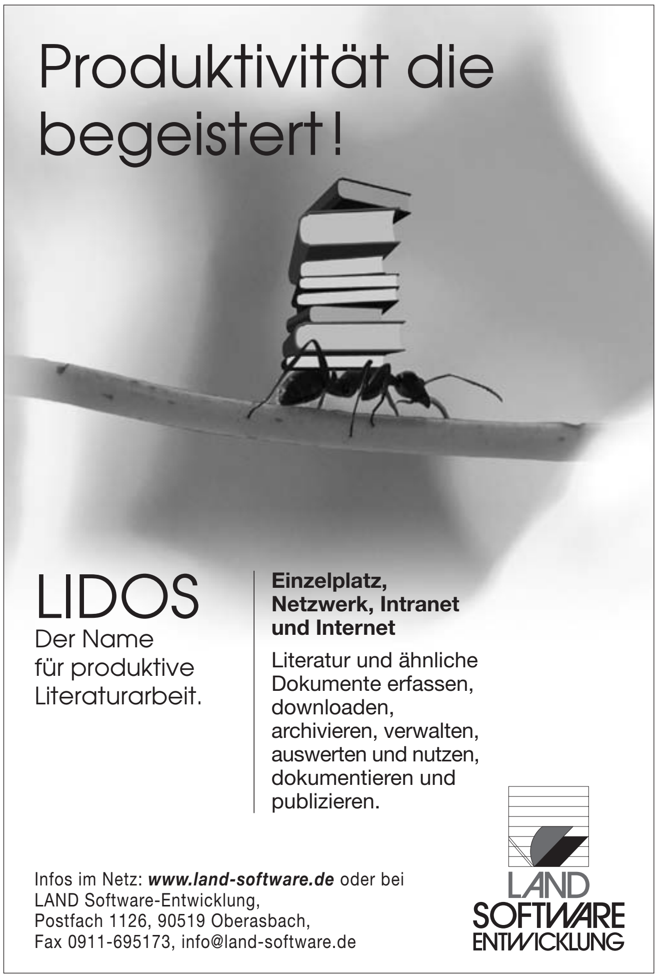
Most of the time, an internet relay chat (IRC) is used as command and control centre (C&C) for a botnet. An IRC consists of one or more servers which relay messages and/or commands to the connected clients. That way the botnet owner can centrally command the clients to download and execute a program which will commit click fraud on the owner’s website(s) (or, respectively, the competitor’s ads).

An alternative way of controlling the bots is to use a web interface to which the clients connect. However, this method requires the client to continually request updates from the C&C, whereas, when IRC is used, the server can just relay commands. Thus, using a web interface causes more traffic to and from the clients, increasing the chances of a bot being exposed and incriminating the botnet (Ianelli & Hackworth, 2005).

In general, a command and control centre will have to include a commonly-used protocol with little overhead and bandwidth usage, in order to remain inconspicuous.

4.3 Referrer Click Fraud

Anupam et al. introduce another, as-yet undetectable, method of committing click fraud: the deceitful publisher puts a script on his website that is automatically downloaded onto a visitor’s computer when said visitor goes to the publisher’s website. The script then imitates a click onto the advertisement leaving the visitor none the wiser. The log files of the advertiser (and, if applicable, of the commissioner) will thus show the visitor’s client ID and IP address, which are unique for everyone. This simple setup, however, is still detectable. All it takes in order to uncover the fraudulent publisher is for the advertiser to visit the publisher’s website and check afterwards whether his IP address and client ID have been registered as having performed a click. This can be remedied by a slight modification. The solution requires two versions of the publisher’s website: one which includes the aforementioned script and one which does not. Internet users who visit the websites are always presented with the latter version and, consequently, if the advertiser or the commissioner checks the website, he cannot detect any sign of fraudulent activity. The publisher then needs a second website, which is wholly unconnected to the first one, over which he has complete control. This can either be a site he owns or he could work together with the owner for the scheme to work. He puts a script onto this second website which automatically loads the first, fraudulent, version of the first website in the background whenever the second website is accessed (Anupam, Mayer, Nissim, Pinkas, & Reiter, 1999).



5 Proposed Solutions

In effect, click fraud costs advertisers and – as a result of lawsuits – advertising network operators (Hadjinian, 2006) millions of dollars a year. It should not come as a surprise, then, that a lot of people have worked (and are still working) on the problem, coming up with several more or less effective solutions. A few of these solutions shall be reviewed in this chapter.

5.1 Cost-Per-Action (CPA)

In the Cost-Per-Action model, advertisers don’t pay for clicks, but rather for specific actions that are performed on the advertiser’s page *after* the click. These actions might, for example, be making a purchase, filling out a form, or registering.

“Such systems are used by Amazon, for example, to sell books on web pages: a service provider, say Expedia, can list an Amazon ad for a travel guide with the understanding that, should a user purchase the product advertised, then the service provider will receive a payment” (Immorlica et al., 2005).

Although CPA does prevent publisher and competitor click fraud, it leaves room for advertiser fraud. Since the publisher (and the commissioner, if he exists) have no way to confirm whether or not a specific action has taken place they depend on the advertiser to report the customer's action truthfully. Additionally, the publisher has to rely on the advertiser's ability to produce both efficient advertisements and – if the action in question is making a purchase – worthwhile products on the target side in order to make a profit. This means that the publisher does not profit directly from the advertising space and traffic he provides anymore. Furthermore, a user might click on an ad on the publisher's website without making a purchase on the target site, only to return to the target site and make the purchase later – thus robbing the publisher of his well-deserved commission. For the advertiser, on the other hand, the CPA model is very advantageous: since he only has to pay when he actually does make a sale, there is virtually no risk on his side.

Security expert Bruce Schneier says: "It's a hard model to make work – [the publisher] would become more of a partner in the final sale instead of an indifferent displayer of advertising – but it's the right security response to click fraud: change the rules of the game so that click fraud doesn't matter" (Schneier, 2006).

However, it is still a model that is based on trust – the only difference being that, this time, it is the other party that has to trust – and as such it is a model that publishers might be wary of adopting. If they do, they will likely favour big companies like Amazon and Ebay over small and possibly unreliable ones.

5.2 Pay-Per-Impression

An alternative to the PPC model of advertising is the old Pay-Per-Impression model. It ''remains popular on major Internet portals, such as yahoo.com, msn.com, and aol.com'' (Edelman et al., 2005). In this model the advertiser pays a specific amount of money for each time that the ad is displayed. This is the most appealing form of advertising for publishers, since they aren't reliant on the effectiveness of the advertiser's ads, but are paid for the space and traffic they provide. Unfortunately, though, Pay-Per-Impression is not fraud resistant either. The technical methods that make click fraud possible can be easily adapted to so-called impression fraud. Instead of simulating a click, though, the script repeatedly requests the website on which the ad is displayed and consequently artificially increases the number of impressions.

5.3 Pay-Per-Percentage of Impressions

Pay-Per-Percentage of Impressions is an alternative to Pay-Per-Click that was suggested by J. Goodman (2005). In his paper he describes this model thusly: "In this system, an advertiser picks a keyword, e.g. 'cameras' and purchases, perhaps through bidding, a certain percentage of all impressions for that keyword. For instance, an advertiser might pay $1.00 to MSN Search. In return, the advertiser might receive 10% of all impressions for "camera" for 1 week. What does this mean? It means that for 1 week, one out of ten times that someone searches for the word 'camera', they will see the ad." The costs of advertising are thus fixed and do not depend on whether or not the ad is clicked. They do not even depend on the number of impressions: "if there are *R* real impressions over the week, and *F* fake impressions, that the advertiser will receive *.1 x R* real impressions and *.1 x F* fake ones" (Goodman, 2005).

Consequently, this system is not susceptible to competitor fraud of any kind (neither click nor impression fraud). However, evaluating how much a certain percentage of impressions on a specific website is worth remains problematic. If such an evaluation is based on the average number of impressions of a site it remains vulnerable to impression fraud.

Goodman (2005) himself admits that "the pay-per-percentage model is not appropriate for all kinds of affiliate advertising; in particular, it is most appropriate for high volume sites". He recommends using a rating company to estimate the traffic on these sites, basing the cost-per-percentage on the estimates.

For the publisher this model is just as advantageous as Pay-Per-Impression, since his payment depends solely on the services he provides. Although the advertiser is safe from competitor fraud, this scheme leaves him susceptible to publisher Impression fraud if he does not advertise on a trustworthy site. Furthermore, the model is worse for him than Pay-Per-Click because the price is not based on the effectiveness of the ad.

5.4 Duplicate Detection

Metwally, Agrawal and Abbadi have proposed detecting fraudulent clicks through a method they call "Duplicate Detection in Click Streams". In order to differentiate between authentic and fraudulent clicks, the advertising commissioner "tracks individual customers by setting cookies. Duplicate clicks within a short period of time, a day for example, raise suspicion on the commissioner's side" (Metwally et al., 2005a).

Duplicate Detection can, no doubt, detect *amateur* click fraud, where the fraudster operates from one or a handful of computers. However, it is clearly inadequate when it comes to detecting distributed click fraud – where millions of computers simulate clicks on an advertisement from all over the world – since it relies on cookies (text files which the commissioner stores and accesses on the user's computer) for detection. In the case of distributed click fraud, every computer in the attacker's vast network will have its own individual cookie. Moreover, an attacker who knows what he is doing will just delete the cookies after each 'click', leaving no duplicates to be detected.

5.5 Association Rules

Metwally, Agrawal and Abbadi have also proposed a solution to the *referrer click fraud* discussed in section 3.3. They propose encouraging ISPs (Internet Service Providers) to provide the data stream necessary to detect this kind of click fraud. This data stream would contain the HTTP requests to page P, which might or might not be fraudulent. They would devise an algorithm to detect associations between one or more sites that refer to P very frequently, and clicks on an ad on P. If strong associations are found, it is very probable that P is using one or more 'decoy' websites in order to commit undetected click fraud (Metwally et al., 2005b).

6 Conclusion

We have seen that it is frighteningly easy to commit click fraud. You can simply set up two websites to implement *referrer click fraud* or, with sufficient funds, buy control of a botnet. You could even relatively effortlessly build a botnet with enough time, starting small and expanding in imitation of a website that grows more and more popular with time.
Neither botnet click fraud nor referrer click fraud can be satisfactorily detected. Of the alternatives presented, only Pay-Per-Percentage of Impressions and Cost-Per-Action are noteworthy, since the Pay-Per-Impression model is as vulnerable as Pay-Per-Click, as well as being less effective. However, both alternatives are far from perfect: Pay-Per-Percentage of Impressions is only a solution to competitor click fraud, since a dishonest publisher can still artificially increase the number of impressions of the ad and, consequently, increase both his income and the costs for each advertiser, since the overall price of advertising on the page in question will increase relative to the number of impressions. Furthermore, if the click-through-rate of the ad is taken into account, he can simply increase that too by commit-

ting click fraud. Hence, we can see that the system is inherently vulnerable.

Cost-Per-Action, on the other hand, is the best alternative to Pay-Per-Click that has been put forward to date. The advantages of this model are quite obvious: not only is it much more effective for the advertiser, who only has to pay a sort of commission when he actually sells something, but it is also nearly impossible to fake *buying* something. (It is actually possible, but that is *credit card fraud* and is an entirely different problem that has little to do with advertising.) However, this model is not only less profitable for the (honest) publisher, but also susceptible to *advertiser* fraud: since the publisher has no way of knowing whether a click on his advertisement has actually led to a sale (or another significant action), it is quite easy for the advertiser to lie about the number of sales generated by an ad and shortchange the publisher. Nonetheless, since the advertisers are probably going to have the last word on the matter, it seems inescapable that Cost-Per-Action will at least play a big role in the future. Indeed, Google is already taking first steps into that direction (Parker, 2006).

References

2006 TV Ad Revenue Figures. (2007). http://tvb.org/rcentral/adrevenuetrack/revenue/2006/ad_figures_1.asp [07.07.2007].

Anupam, V., Mayer, A., Nissim, K., Pinkas, B., & Reiter, M. K. (1999). On the security of pay-per-click and other web advertising schemes. In WWW '99: Proceeding of the eighth international conference on world wide web (pp. 1091–1100). New York, NY, USA: Elsevier North-Holland, Inc.

Buffer overflow. (n.d.). http://searchsecurity.techtarget.com/sDefinition/0,290660,sid14gci549024,00.html [05.06.2007].

Edelman, B., Ostrovsky, M., & Schwarz, M. (2005). Internet advertising and the generalized second price auction: Selling billions of dollars worth of keywords. http://www.seo.com.ph/pdfguide/RP1917.pdf [06.06.2007].

Egelhoff, T. (n.d.). How to plan your advertising budget strategy. http://www.smalltownmarketing.com/adbudget.html [31.05.2007].

Gandhi, M., Jakobsson, M., & Ratkiewicz, J. (2006). Badvertisements: Stealthy click-fraud with unwitting accessories. Journal of Digital Forensic Practice, 1.

Goodman, J. (2005, June). Pay-per-percentage of impressions: An advertising method that is highly robust to fraud. Presented at the ACM E-Commerce Workshop on Sponsored Search Auctions, June, 2005.

Google AdSense Online Standard Terms and Conditions. (n.d.). https://www.google.com/adsense/static/en US/Terms.html [05.06.2007].

Google Annual Report 2005. (2006). http://investor.google.com/ pdf/2005_Google_AnnualReport.pdf [07.07.2007].

Google Annual Report 2006. (2007). http://investor.google.com/ pdf/2006_Google_AnnualReport.pdf [07.07.2007].

Hadjinian, D. L. (2006). Clicking away the competition: The legal ramifications of click fraud for companies that offer pay per click advertising services. Shidler J. L. Com. & Tech. at http://www.lctjournal.washington.edu/Vol3/a005Hadjinian.html, 3.

IAB Internet Advertising Revenue Report. (2007). http://www.iab.net/resources/adrevenue/pdf/IAB_PwC_2006_Final.pdf [07.07.2007].

Internet Advertising Revenues Soar Again, Near $5 Billion in Q1 07. (2007). http://www.iab.net/news/pr_2007_06_06.asp [07.07.2007].

Ianelli, N., & Hackworth, A. (2005). Botnets as a vehicle for online crime. Retrieved May 31, 2007, from www.cert.org/archive/pdf/Botnets.pdf.

Immorlica, N., Jain, K., Mahdian, M., & Talwar, K. (2005). Click fraud resistant methods for learning click-through rates. Lecture Notes in Computer Science, 3828.

Kaufmann, N. C. (n.d.). Click-spamming – ein Fall fuer das reformierte UWG? http://rsw.beck.de/rsw/shop/default.asp?sessionid=80CB3C7120844459E1FFCDA11953749&docid=138085&docClass=NEWS&from=mmr.root [10.05.2007].

Keizer, G. (2005). Dutch botnet suspects ran 1.5 million machines. http://www.techweb.com/wire/security/172303160 [10.05.2007].

Metwally, A., Agrawal, D., & Abbadi, A.E(2005a). Duplicate detection in click streams. In Proceedings of the 14th WWW International World Wide Web Conference, 12-21.

Metwally, A., Agrawal, D., & Abbadi, A E. (2005b). Using association rules for fraud detection in web advertising networks. In Vldb '05: Proceedings of the 31st international conference on very large data bases (pp. 169–180). VLDB Endowment.

*Metwally, A., Agrawal, D., &Abbadi, A.E.*2007). Detectives: detecting coalition hit inflation attacks in advertising networks streams. In Proceedings of the 16th international Conference on World Wide Web, 241-250.

Mills, E. (2006a). Study: Click fraud could threaten pay-per-click model. http://www.news.com/Study-Click-fraud-could-threaten-pay-per-click-model/2100-1024_3-6090939.html [05.10.2007].

Mills, E. (2006b). Click fraud increasing, study finds. http://www.news.com/Click-fraud-increasing,-study-finds/2100-1030_3-6095074.html [05.10.2007].

Mind the Gap. (2007). http://recoveringjournalist.typepad.com/recovering_journalist/2007/03/mind_the_gap.html [07.07.2007].

Nmap security scanner. (n.d.). http://insecure.org/nmap/ [05.06.2007].

Override, C. (2001). Exploits. http://www.gcf.de/papers/exploits.htm [01.06.2007].

Parker, P. (2006). Google tests cost-per-action. http://www.clickz.com/showPage.html?page=3615476 [25.06.2007].

Richardson, M., Dominowska, E., and Ragno, R. (2007). Predicting clicks: estimating the click-through rate for new ads. In Proceedings of the 16th international Conference on World Wide Web, 521-530.

Schneier, B. (2006). Google's click-fraud crackdown. http://www.wired.com/politics/security/commentary /securitymatters/2006/07/71370 [25.06.2007].

Vidyasagar, N. (2004, May 3). India's secret army of online ad 'clickers'. The Times of India.

Betrug, Werbung, Suchmaschine, Google, Benutzung, Fehler

Click fraud, Web advertising, Pay per Click, Google AdWords, Google AdSense, Competitor click fraud, Publisher click fraud, Botnet, Referrer click fraud, Pay per action, Advertiser fraud

The Author

Simone Soubusta, B.A.

has studied Information Science and Language Technology at Heinrich-Heine-University Düsseldorf. This article is a revised and shortened version of her bachelor's thesis.

simone@soubusta.de

24. Oberhofer Kolloquium vom 10. bis 12. April in Magdeburg

Im Jubiläumsjahr 2008 der DGI steht das traditionsreiche Oberhofer Kolloquium unter dem Rahmenthema „Informationskompetenz 2.0. Zukunft von qualifizierter Informationsvermittlung."

Die Veranstalter, DGI und VDI, freuen sich, dass der bekannte Informatiker und Medienkritiker Joseph Weizenbaum sein Kommen zugesagt hat. Auch andere Persönlichkeiten aus Informationswissenschaft und -praxis haben angekündigt, sich der Diskussion mit den Newcomern und Quereinsteigern im Bereich Informationsvermittlung stellen zu wollen. Das Tagungshotel bietet einen idealen Rahmen für das traditionelle Motto der Tagung: „Tagen und Wohnen unter einem Dach". Die Kommunikation unter den Teilnehmern, die informell auch in die Abendstunden fortgesetzt wird, gehört zu den wesentlichen Elementen des Kolloquiums. Bei den Vorträgen stehen berufsbezogene praktische Erfahrungen im Vordergrund. Alles über die Tagung und das detaillierte Programm finden Sie unter www.dgi-info.de/oberhofer.aspx.

Die Anzahl der Zimmer ist beschränkt, nehmen Sie Ihre Anmeldung möglichst rasch vor.