The Lane's Gifts v. Google Report

Alexander Tuzhilin

Table of Contents

1. Dr. Tuzhilin's Background	1
2. Materials Reviewed	2
3. Google Personnel Interviewed	3
4. Development of the Internet	4
5. Growth of Search Engines and Google's History	5
6. Development of the Pay-per-Click Advertising Model	6
7. Google's Pay-per-Click Advertising Model	9
8. Invalid Clicks and Google's Definition	15
9. Google's Approach to Detecting Invalid Clicks	22
10. Conclusions	48

Executive Summary

I have been asked to evaluate Google's invalid click detection efforts and to conclude whether these efforts are reasonable or not. As a part of this evaluation, I have visited Google's campus three times, examined various internal documents, interviewed several Google's employees, have seen different demos of their invalid click inspection system, and examined internal reports and charts showing various aspects of performance of Google's invalid click detection system. Based on all these studied materials and the information narrated to me by Google's employees, I conclude that Google's efforts to combat click fraud are reasonable. In the rest of this report, I elaborate on this point.

1. Dr. Tuzhilin's Background

I have recently been appointed as a Professor of Information Systems at the Stern School of Business at New York University (NYU), having previously served as an Associate Professor at the Stern School. I received my Ph.D. in Computer Science from the Courant Institute of Mathematical Sciences, NYU in 1989, M.S. in Engineering Economics from the School of Engineering at Stanford University in 1981, and B.A. in Mathematics from NYU in 1980.

My current research interests include knowledge discovery in databases (data mining), personalization, Customer Relationship Management (CRM) and Internet marketing. My prior research was done in the areas of temporal databases, query-driven simulations and the development of specification languages for modeling business processes. I have coauthored over 70 papers on these topics published in major Computer Science and Information Systems journals, conferences and other outlets. I currently serve on the Editorial Boards of the IEEE Transactions on Knowledge and Data Engineering, the Data Mining and Knowledge Discovery Journal, the INFORMS Journal on Computing, and the Electronic Commerce Research Journal. I have also co-chaired the Program Committees of the IEEE International Conference on Data Mining (ICDM) in 2003 and the 2005 International Workshop on Customer Relationship Management that brought together researchers from the data mining and marketing communities to explore and promote an interdisciplinary focus on CRM. I have also served on numerous program and organizing committees of major conferences in the fields of Data Mining and Information Systems. I have also had visiting academic appointments at the Wharton School of University of Pennsylvania, Computer Science Department of Columbia University, and Ecole Nationale Superieure des Telecommunications in Paris, France.

On the industrial side, I worked as a developer at Information Builders, Inc. in New York for two years and consulted for various companies, including Lucent's Bell Laboratories on a data mining project and Click Forensics on a click fraud detection project.

Additional information about my background can be found in my CV in the Appendix.

2. Materials Reviewed

During this project, I reviewed the following materials:

1. Internal documents provided to me by Google, including the following documents:

    - Type of data collected and statistics/signals used for the detection of invalid clicks
    - Description of the filtering methods
    - Description of the log generation and log transformation/aggregation system used for the analysis and detection of invalid clicks.
    - Description of the AdSense auto-termination system
    - Description of the duplicate AdSense account detection system
    - Description of the ad conversion system
    - Description of the AdSense publisher investigation, flagging and termination systems
    - Description of various Click Quality investigative processes, including the rules on when and how to terminate the publishers
    - Description of the advertiser credit processes and systems

- Description of the inquiry handling processes and guidelines
- Description of the attack simulation system
- Description of the alerting system
- History of the doubleclicking action
- Overview of the Click Quality team's high-priority projects
- Investigative reports generated by 3 different inspection systems that investigated three different cases of invalid clicking activities. One was an attack on an advertiser by an automated system, another one was an attack on a publisher by an automated system, and the third one was a general investigation of certain suspicious clicking activities. These reports were generated as a part of giving me demos on how Google's inspection systems worked and how manual offline investigations are typically conducted by Google personnel.
- Different internal reports and charts showing various aspects of performance of Google's invalid click detection systems.

2. Demos of various invalid click detection and inspection systems developed by the Click Quality team. Of course, these demos were provided only for the Click Quality systems that can be demoed (e.g., have appropriate User Interfaces).

3. Interviews with Google personnel, as described in the next section.

   This report is based on this reviewed information and on the information narrated to me by Google personnel during the interviews.

3. Google Personnel Interviewed

All the invalid click detection activities are performed by the Click Quality team at Google. The Click Quality team consists of the following two subgroups

- *Engineering*
  Responsible for the design and development of online filters and other invalid click detection software. It consists primarily of engineers and currently has about a dozen staff members on the team.
- *Spam Operations*
  Responsible primarily for the offline operations, inspections of invalid clicking activities including investigations of customer's inquiries. The group currently has about two dozens staff members on the team.

In addition, several other groups at Google, including Web spam, Ads quality, Publications quality and others interact with the Click Quality team and provide their expertise on the issues that are related to invalid clicks (e.g., Web spam and click fraud have some issues in common). Overall, the Click Quality team can draw upon the knowledge and expertise of a few dozens of other people on these teams, whenever required.

The two groups, although located in different parts of the Google campus, interact closely with each other.

In addition, the Product Manager of the Trust and Safety Group works closely with the Click Quality team on more business oriented and public relations issues pertaining to invalid click detection.

During this project, I visited Google campus three times and interviewed over a dozen of the Click Quality team members from the Spam Operations and the Engineering groups, as well as the Product Manager of the Trust and Safety Group. I found the members of both groups to be well-qualified and highly competent to perform their jobs. Most of them have relevant prior backgrounds and strong credentials.

Before focusing on the Pay-per-Click advertising model and Google's efforts to combat invalid clicks, I first provide some background materials on the Internet and the growth of the search engines to put these main topics into perspective.

4. Development of the Internet

The Internet is a worldwide system of interconnected computer networks that transmit data using packet switching methods of the Internet Protocol (IP). Computing devises attached to the Internet can exchange data of various types, from emails to text documents to video and audio files, over the pathways connecting computer networks. These documents are partitioned into pieces, called packets, by the Internet Protocol and travel over the pathways in a flexible manner determined by routers and other devices controlling the Internet traffic. These packets are assembled back in the proper order at the destination site using the well-developed principles of the Internet Protocol.

Internet was developed long time ago. The predecessor of the Internet (called the ARPANET) was developed in late 1960's and early 1970's. The first wide area Internet network was operational by January 1983 when the National Science Foundation constructed a network connecting various universities. The Internet was opened to commercial interests in 1985.

Prior to the 1990's, Internet was predominately used by the people with strong technical skills because most of the Internet applications at that time required such skills, and only relatively few people had these skills in those days. This situation changed dramatically and the Internet became much more accessible to the general public after the invention of the World Wide Web (WWW) by Tim Berners-Lee in 1989.

WWW is a globally connected network of Web servers and browsers that allows transferring different types of Web pages and other documents containing text, images, audio, video and other multimedia resources *over* the Internet using a special type of protocol developed specifically for the Web (the so-called HTTP protocol). Each resource on the WWW (such as a Web page) has a unique global identifier (Uniform Resource Identifier (or Locator) – URI (URL)), so that each such resource can be found and accessed. Web pages are created using special *markup languages*, such as HTML or XML that contain commands telling the browser how to display information contained in these pages. The markup languages also contain commands for *linking* the page to other pages, thus creating a *hypertext* environment that lets the Web user navigate from one Web page to another using these links (clicking on them) and thus letting the users to "surf" the Web.

The development of the World Wide Web, Web documents and Web browsers for displaying these documents in a user-friendly fashion, made Internet much more userfriendly. This opened Internet to the less technologically savvy general public that simply wanted to display, access and exchange various types of information without resorting to complicated technical means that were needed before to achieve these goals. By developing the Web and thus making the tasks of displaying, accessing and exchanging information over the Internet much simpler, spawned the development of various types of *websites* that collect, organize and provide systematic access to Web documents. The number of these websites experienced explosive growth in the 1990's and continued to grow rapidly worldwide up until now.

Massive volumes of Web documents were created over a short period of time since the invention of the WWW. To deal with this information overload, it was necessary to search and find relevant documents among millions (and later billions) of Web pages spread all over the world among numerous websites. This gave rise to the creation and growth of search engines designed to search and find relevant information in the massive volumes of Web documents.

5. Growth of Search Engines and Google's History

A search engine finds information requested by the user that is located somewhere on the World Wide Web or other places, including proprietary networks and sites, and on a personal computer. The user formulates a search query, and the search engine looks for documents and other content satisfying the search criteria of the query. Typically, these search queries contain a list of keywords or phrases and retrieve documents that match these queries. Although the search can be done in various environments, including corporate intranets, the majority of the search has been done on the Web for different kinds of documents and information available on the Web. Since searching these documents directly on the Web is prohibitively time consuming, all the search engines use indexes to provide efficient retrieval of the searched information. These indexes are maintained regularly in order to keep them current.

The history of search engines goes back to Archie and Gopher, two tools designed in 1990 – 1991 for searching files located at the publicly accessible FTP sites over the Internet (and not over the WWW which did not exist at that time). The early commercial search engines for the Web documents were Lycos, Infoseek, AltaVista and Excite, which were launched around 1994 – 1995.

Google co-founders started working on developing Google search engine in 1997 and Google Inc. was founded in September 1998. The beta label came off the Google website in September 1999. The co-founders have developed innovative patented search technologies based on the PageRank concept that turned out to be highly effective in generating good search results. Google popularity grew rapidly, and the company was handling more than 100 million search queries a day by the end of 2000. Around that time, Google started launching various additional offerings, such as Google Toolbar, and this trend continued since then. Currently, Google supports a couple of dozens of such offerings publicly available on the Google's website.

Currently, the main competing search engines for Google include (a) Yahoo! that acquired Inktomi search engine in 2002 and also Overture which owned AltaVista, and (b) Microsoft which launched its own independent MSN Search engine in early 2005. Google is currently the market leader in the search engine field, accounting for over 50% of all the Web search queries.

Google realized the power of the keyword-based targeted advertising back in 2000 when it launched its initial version of AdWords, which was quite different from its current version and even from the version launched in February 2002. The Pay-per-Click overhauled version of AdWords was launched in February 2002. It was followed by the AdSense program in March 2003.

The AdWords and AdSense programs will be described later in Section 7 in the context of Google's overall Pay-per-Click advertising model. However, before doing this, I will first present a general overview of the Pay-per-Click advertising model in Section 6.

6. Development of the Pay-per-Click Advertising Model

The idea of delivering *targeted ads* to an internet user has been around for a long time. For example, such companies as DoubleClick have been involved in this effort since the 90's. The key question in this problem is: what is the *basis* for targeting these ads? The ads can be targeted based on:

1. personal characteristics of a web page visitor known to the party delivering an ad 2. keywords of a search query launched by the user

3. content of a web page visited by the user.

The first source of targeting, based on personal characteristics of a web page visitor, has been adopted by various companies in the personalization and Customer Relationship Management area. The two other sources of targeting are adopted by the search engines, including Google.

The second issue dealing with the delivery of targeted ads is the payment model. When the ads are delivered to the user, *for what exactly* should advertisers pay and when? The alternative choices for charging an advertiser are:

- when the ad is being shown to the user
- when the ad is being clicked by the user
- when the ad has "influenced" the user in the sense that its presentation lead to a *conversion event*, such as the actual purchase of the product advertised in the ad or other related conversion events, such as placing the related product into the user's shopping basket.

From the advertiser's point of view, the weakest form of delivery is when an ad is only shown to the user because the user may not even look at it and may simply ignore the ad. Clicking on an ad indicates some interest in the product or service being advertised. Finally, the most powerful user reaction to an ad is the conversion event when the user actually *acts* in response to the ad, with the most powerful type of action being actual purchase of the advertised product or service. For these reasons, advertisers value these three activities differently and, generally, are willing to pay more money per conversion event than per clicking event and than per ad viewing event (however, there are also some exceptions to this observation, which I will not cover in this report because they have only tangential relevance).

The two key measures of how effective an advertisement is are

- *Click-Through Rate (CTR):* it specifies on how many ads X, out of the total number of ads Y shown to the visitors, the visitors actually clicked; in other words, $CTR = X/Y$. CTR measures how often visitors click on the ad.
- *Conversion Rate:* it specifies the percentage of visitors who took the conversion action. Conversion rate gives a sense of how often visitors actually act on a given ad, which is a better measure of ad's effectiveness than the CTR measure.

Conversion actions are actually very relevant to click fraud because proper conversion actions following clicking activities, such as a purchase of an advertised product, are really good indicators that the clicks are valid. However, less direct conversion actions, such as putting a product into a shopping cart, may still not be indicative of a valid click since it can be a part of a *conversion fraud* (an unethical user may do it on purpose without a true intent to purchase the product, but just simply to confuse an invalid click detection system).

The three situations described above give rise to the following three different internet advertising payment methods:

- *CPM – Cost per Mille* – an advertiser pays per one thousand impressions of the ad ("Mille" stands for "thousand" in Latin); an alternative term used in the industry for this payment model is *CPI (Cost per Impression)*.
- *CPC – Cost per Click* (a. k. a. *Pay per Click or PPC;* we will use these terms interchangeably) – an advertiser pays only when a visitor clicks on the ad, as is clearly stated in the name of this payment model.
- *CPA – Cost per Action* – an advertiser only pays when a certain conversion action takes place, such as a product being purchased, an advertised item was placed into a shopping cart, or a certain form being filled. This is the best option for an advertiser to pay for the ads from the advertisers' point of view since it gives the best indication among the three alternatives that the ad actually "worked" (as I said before, however, there are certain exceptions to this general observation).

Early forms of internet advertising models were mainly CPM-based. For example, Google initially based the AdWords program only on the CPM model between 2000 and February 2002.

However, the CPC model is more attractive for many (but not all) advertisers than the CPM model, and it replaced the CPM as a predominant internet advertising payment model. For example, this is certainly the case for Google since most of its advertisers currently use the CPC model.

The origins of the CPC model go back to mid-90's when different payment models were debated in the internet marketing community. The first major commercial keywordbased CPC model was introduced by Overture (previously known as GoTo.com, now part of Yahoo!) that has developed certain patented technologies for implementing this model that go back to 1999. Google introduced its keyword- and CPC-based AdWords program in February 2002. Besides Google and Yahoo!, Microsoft has also recently deployed the CPC payment model through its adCenter program. Also, several other online advertising programs use the CPC/PPC payment model.

If one combines a particular ad payment method with a particular targeting method, this combination determines a specific targeted ad delivery model. For Google and Yahoo! the two main models are the keyword-based PPC and the content-based PPC models.

Although currently popular, the CPC/PPC model has two fundamental problems:

- Although correlated, good click-through rates (CTRs) are still not indicative of good conversion rates, since it is still not clear if a visitor would buy an advertised product once he or she clicked on the ad. In this respect, the CPA-based models provide better solutions for the advertisers (but not necessarily for the search engines), since they are more indicative that their ads are "working."

- It does not offer any "built-in" fundamental protection mechanisms against the click fraud since it is very hard to specify which clicks are valid vs. invalid in general, as will be explained in Section 8 (it can be done relatively easily in some special cases, but not in general). For this reason, major search engines launched extensive invalid click detection programs and still face problems combating click fraud.

In response to these two problems and for various other business reasons, Google is currently testing a CPA payment model, according to some reports in the media. Some analysts believe that the conversion-based CPA model is more robust for the advertisers and also less prone to click fraud. Therefore, they believe that the future of the online advertising payments lies with the CPA model. Although this is only a belief that is not supported by strong evidence yet, Google is getting ready for the next stage of the online advertising "marathon."

7. Google's Pay-per-Click Advertising Model

As stated in Section 6, Google introduced the CPC/PPC model in addition to the previously deployed CPM model for the AdWords program in February 2002. The PPC model is widely adopted by Google now and its two main programs, AdWords and AdSense, are based on it. These two programs are described below, including how the PPC advertising model is used in them.

7.1. The AdWords Program

AdWords is a program allowing advertisers to purchase CPC-based advertising that targets the ads based on the keywords specified in users' search queries. An advertiser chooses the keywords for which the ad will be shown on Google's web page (Google.com) or some other "network partner" pages, such as AOL and EarthLink (to be discussed below in Section 7.4), and specifies the maximum amount the advertiser is willing to pay for each click on this ad associated with this keyword. For example, an accounting firm signs with Google AdWords program and is willing to pay up to $10/click for showing its ad (a link to its home page combined with a short text message) on Google.com when the user types the query "tax return" on Google.

When a user issues a search query on Google.com or a network partner site, ads for relevant words are shown along with search results on the site on the right side of the Web page as "sponsored links" and also above the main search results.

The ordering of the paid listings *on the side* of the page is determined according to the Ad Rank for the candidate ads that is defined as

Ad Rank = CPC x QualityScore,

where QualityScore is a measure identifying the "quality" of the keyword/ad pair. It depends on several factors, one of the main ones being the clickthrough rate (CTR) on the ad. In other words, the more the advertiser is willing to pay (CPC) and the higher the clickthrough rate on the ad (CTR), the higher the position of the ad in the listing is. There exists the whole science and art of how to improve the Ad Rank of advertisers' ads, collectively known as Ad Optimization, so that the ad would be placed higher in the list by Google. Various tips on how to improve the results are presented on Google's website at https://adwords.google.com/support/bin/static.py?page=tips.html&hl=en_US. The *topof-the-page* placement rank is also determined by the above Ad Rank formula; however, the value of the QualityScore for the top-of-the-page placement is computed somewhat differently than for the side ads.

The actual amount of money paid when the user clicks on an ad is determined by the lowest cost needed to maintain the clicked ad's position on the results page and is usually less than the maximal CPC specified by the advertiser. Although the algorithm is known, the advertiser does not know a priori how much the click on the ad will actually cost because this depends on the actions of other bidders which are unknown to the advertiser beforehand. However, it is lower than the maximal CPC that the advertiser is willing to pay.

An advertiser has a certain budget associated with a keyword, which is allocated for a specified time period, e.g. for a day. For example, the accounting firm wants to spend no more than $100/day for all the clicks on the ad for the keyword "tax return." Each click on the ad decreases the budget by the amount paid for the ad, until it finally reaches zero during that time period (note that more money is added to the budget during the next time period, e.g., the next day). If the balance reaches zero, the ad stops showing until the end of the time period (actually, the situation is somewhat more complex because Google has developed a mechanism to extend the ad exposure over the whole time period, but do it over short time intervals with long blackout periods; however, in the first approximation, we can assume that the ad stops showing when the balance reaches zero). For example, if the budget for the keyword "tax return" reached zero by the mid-day, then no ads for the accounting firm are shown for the "tax return" query for the rest of the day (modulo the previous remark). However, the ad is resumed the next day, assuming that the accounting firm has signed up with Google for the next day.

This is one of the motivations for the click fraud with the purpose to hurt other advertisers. If an advertiser or its partner can deplete the budget of a competitor by repeatedly clicking on the ad, the competitor's ad is not being shown for the rest of the time period, and the advertiser's ad has less competition and should appear higher in the paid ads list. Moreover, the advertiser may also end up paying less for his/her ad since there is less competition among the advertisers. Therefore, unethical advertisers or their partners not only hurt their competitors financially by repeatedly clicking on their ads, they also knock them out of the auction competition for the rest of the day by depleting their advertising budgets and thus improving their positions in the sponsored link lists and also paying less for their own ads.

When search queries are launched on the network partners' websites, such as AOL or EarthLink, the PPC model works the same way as on Google.com with two caveats: (a) the ads are displayed somewhat differently on these websites than on Google.com and (b) Google shares parts of its advertising revenues with these partners.

AdWords based on the CPC/PPC advertising model described above was launched in February 2002. It changed Google's business model and was responsible for generating major revenue streams for the company.

7.2. The AdSense Program

Google AdSense is a program for the website owners (known as *publishers*) to display Google's ads on their websites and earn money from Google as a result. To participate in this program, website publishers need to register with Google and be accepted into the program by Google. These ads shown on the publishers' websites are administered by Google and generate revenue on either per-click or per-thousand-ads-displayed basis. Since we are interested in click fraud, we will limit our considerations only to clicks and to the PPC payment method.

AdSense was launched in March 2003 and constituted the second major milestone in Google's PPC advertising model that generated significant additional revenues for the company.

There are two ways for publishers to participate in the AdSense program:

- *AdSense for Search (AFS):* publishers allow Google to place its ads on their websites when the user does keyword-based searches on their sites. In other words, as a result of a search, relevant ads are displayed as links sponsored by Google, and these links are produced using the same methods as on Google.com. Examples of such publishers include AOL and EarthLink. Moreover, the search results pages containing the ads are customizable to fit with the publisher's site theme, and may have a different "flavor" than the ads on Google.com.
- *AdSense for Content (AFC):* the system that automatically delivers targeted ads to the publisher's web pages that the user is visiting. These ads are based on the content of the visited pages, geographical location and some other factors. These ads are usually preceded by statement "Ads by Google." Google has developed methods for matching the ads to the content of the pages that also take into account the CPC values when selecting the best ads to place on the page. The whole idea is to display ads that are relevant to the users and to what the users are looking for on the site so that they would click on the displayed ads. This is also combined with financial considerations (the CPC factor) to maximize the expected revenues for Google from displaying the ad.

In both the AFS and the AFC cases, the publishers and Google are being paid by the advertisers on the PPC basis. Google does not disclose how it shares the clicking revenues with the publishers. What the publishers can see though, are the detailed online reports helping the publishers to track their earnings. These reports contain several statistics of clicking activities on the ads displayed on publisher's website. These statistics help the publisher to get an idea of how well his or her website is performing in the AdSense program and how much the publisher is expected to earn over time.

As we can see from this description, there is a direct incentive for the publishers to attract traffic to their websites and encourage the visitors to click on Google's ads on the site to maximize their own AdSense income. They can do this in three ways:

- Build a valuable content on the site that attracts the most highly paid ads.
- Use a wide range of traffic generating techniques, including online advertising.
- Encourage clicks on ads using *legitimate* means (Google has a list of prohibited activities for the publishers, such as explicit requests to click on Google's ads, that can lead to terminations of their accounts).

Unfortunately, overzealous and unethical users can "stretch" or directly abuse this system in the effort to maximize their revenues from the AdSense program. This leads to the invalid clicks problem discussed in the next section.

It is interesting to note that AdWords and AdSense have different motivations for the unethical users to abuse the programs. Unethical users on AdWords constitute advertisers or their partners whose motivation is to hurt other advertisers. In contrast to this, the main motivation of the AdSense unethical publishers is to enrich themselves through certain prohibited means. Therefore, motivations of these two groups of unethical users are significantly different.

Although both motivations are important and should be addressed in the most serious manner, greedy motivations of unethical AdSense publishers constitute more serious problem for Google than the desire to hurt the competitors by unethical advertisers or their partners. This results in a significantly greater *percentage* of invalid clicks being generated by unethical AdSense publishers than by unethical AdWords advertisers (however, it is not clear if this statement is still true in terms of *absolute numbers* of invalid clicks generated by these two sources because of different volumes of clicks for the two programs).

7.3 The Google Network

Initially, Google's sponsored links were displayed only on Google.com. However, over the years, Google built and expanded its partner's network to include various websites into, the so-called, *Google Network.* With this network of partners, Google ads can be placed not only on Google.com but also on the partners' websites either using the searchbased or the content-based methods described in Section 7.2. Google provides tools for advertisers to express preferences on which types of sites in the Network they prefer their ads to appear.

Based on how these ads are placed, Google Network can be categorized into the following types of websites:

- *Google.com*: the flagship and the original site in the Network against which all other Network sites are compared.
- *AdSense for Content (AFC)* sites*:* web publishers' sites where content-based ads are served as described in Section 7.2. These publishers are divided into ○ *Direct Publishers:* the most important and trusted publishers, such as New York Times, with whom Google has special relationships. Because of the brand names and reputations of these publishers, very little invalid clicking activities occur on these websites. Even when invalid clicking activities occur, they usually arise because of some technical problems and "miscommunications" between Google's and publisher's software systems. These problems are usually quickly detected and resolved, and the resulting invalid clicks are credited back to advertisers.
    - *Online Publishers:* smaller "self-service" publishers, such as various bloggers who joined the AdSense program. Most of the invalid clicking activities are associated with these publishers.
- *AdSense for Search (AFS)* sites: search sites displaying Google's ads based on the searches done by the site visitors, as described in Section 7.2. These sites are also divided into ○ *Direct:* the most important and trusted search sites, such as AOL and EarthLink, with whom Google also has special relationships.
    - *Online:* other search sites.

  Most of the search sites are Direct with whom Google has special relationships.

This network of partner sites is constantly evolving as new partners are added and old ones either leave or are terminated by Google. All the partner sites in the network are periodically reviewed and monitored to detect possible problems and assure advertisers that their ads are placed only on the sites that passed certain quality control standards.

Among the five types of sites in the Google network, the one category that is intrinsically prone to invalid clicking activities is the AFC Online category. Examples of these publishers include various bloggers and "homegrown" web masters with unknown or unclear reputation in the field.

7.4 What Google Knows about Clicking Activities

In order to manage the AdSense and AdWords programs, properly charge advertisers for the PPC revenue model, share revenues with publishers and detect invalid clicks, Google collects various types of information about querying and clicking activities, including certain types of "post-clicking" data about conversion actions on the advertiser's website where the visitor is taken following the click. All this data accumulated by Google is extracted from various sources and contains comprehensive information about visitor's activities on the Google Network.

As stated before, the conversion data – the "post-clicking" data about conversion actions on the advertiser's website – constitutes an important piece of this collected data. In particular, if the advertiser formally agrees to provide this information, Google collects data on whether or not the user visited certain designated pages on the advertised website that the advertiser marked as "conversion" pages, such as the checkout page and certain form filling pages. This conversion data is limited to what the advertiser decided to provide to Google and is not as rich as the clickstream data collected by advertisers themselves on their websites. Also, many advertisers decide to opt out from providing this conversion data. In this case, Google does not have any conversion information and therefore does not know what happened after a visitor clicked on the ad. Nevertheless, this post-clicking conversion data is important for Google even in its limited form because it conveys some intentions of the visitors on the advertised website and provides good insights into whether or not the visitor is seriously considering purchasing the advertised product or service.

This "raw" clicking data described above is subsequently cleaned, preprocessed and stored in various internal logs by Google for different types of subsequent analysis conducted on this data.

One inherent weakness of Google's (or any other search engine) data collection effort that is important for detecting invalid clicks, is inability to get full access to *all* the clicking activities of the visitors of the advertised website. In other words, the conversion data that Google collects provides only a partial picture of all the post-clicking activities of the visitor on the advertised website. This data is important for detecting invalid clicks since better invalid click detection methods can be developed using this data. Unfortunately, Google (and other search engines) does not have full access to this data, unless the advertised website decides to provide its clickstream data to Google, which many websites are reluctant to do. However, this is not Google's fault – this is an inherent limitation of the types of data available to Google.

However, this lack of full conversion data available to Google is compensated by various types of querying and clicking data that Google can collect, whereas advertisers and third-party vendors cannot. Therefore, there exists a tradeoff between the types of data relevant for detecting invalid clicks that is available to Google, advertisers and the thirdparty vendors. None of these three groups have the most comprehensive set of data pertinent to detecting invalid clicks, and each of them needs to settle for the invalid click detection methods possible only with the data that they have.

7.5 The Advertisers' Dilemma or What Knowledge Google Shares with Advertisers about Clicks

When advertisers are billed by Google, they receive reports describing the clicking and billing activities. These reports can be customized by the advertisers who can select various clicking statistics that they want to see in these reports. These reports were much simpler initially; but Google enhanced its reporting functionality over the last few years, and the customers can see a wide range of clicking statistics in these reports now.

One problem with these reports, however, is that these statistics are aggregated by Google over some time period. The smallest unit of analysis is one day. For example, the number of invalid clicks on an ad detected by Google (or any other related statistic) can *only* be reported on a daily basis (although there are certain alternative methods of obtaining aggregation granularity that is smaller than a day). In other words, advertisers cannot know if a particular click on a particular ad was marked as valid or invalid by Google, and Google refuses to provide this information to advertisers.

This is a source of contention and dispute between Google and the advertisers, and one can understand both parties in this dispute. On one hand, the advertiser has the right to know why a particular click was marked as valid by Google (when the advertiser thinks that it is invalid) because the advertiser pays for this click. On the other hand, if Google discloses this information, it opens itself to click fraud on a massive scale because, by doing so, it provides certain hints about how its invalid click detection methods work. This means that unethical users will immediately take advantage of this information to conduct more sophisticated fraudulent activities undetectable by Google's methods.

This conflicting dilemma between advertisers' right to know and Google's inability to provide the appropriate information to advertisers because of the security concerns is part of the Fundamental Problem of the PPC advertising model to be discussed in the next section.

More recently, Google tried to bridge this gap between Google and the advertisers by explaining to advertisers a little more about Google's invalid click detection efforts. However, these activities, although indicative of Google's desire to work closer with the advertisers, are too small to be of any major consequence. Therefore, the gap described above and the Fundamental Problem of the PPC model still remains pretty much open.

8. Invalid Clicks and Google's Definition

8.1. Conceptual Definitions of Invalid Clicks

There are numerous definitions of fraudulent and invalid clicks. One such definition, taken from Wikipedia (http://en.wikipedia.org/wiki/Invalid_click), is

> *"Click fraud* occurs in pay per click online advertising when a person, automated script or computer program imitates a legitimate user of a web browser clicking on an ad, for the purpose of generating an improper charge per click."

Google does not like the concept of "fraudulent" clicks and uses the term "invalid" (or "spam") click instead. Google provides the following definition of invalid clicks (https://www.google.com/support/adsense/bin/answer.py?answer=32740&topic=8526):

> "Clicks … generated through prohibited means, and intended to artificially increase click … counts on a publisher [or advertiser – AST] account"

Google has also used other definitions of invalid clicks in the past, such as

> *Click spam* [*invalid click* – AST] is any kind of click received from a Cost-PerClick (CPC) advertising engine that is generated artificially though human or technological means with the sole purpose of creating a debiting click, resulting in zero possibility for a conversion to occur

All these related definitions emphasize the following points:

- Invalid clicks can be generated either by humans or technological means, including various types of deceptive software programs, such as scripts or bots.
- When evaluating validity of a click, it is necessary to understand the *intent* of clicking on the ad by the user and to determine if there is *any* possibility of conversion or the intent is *only* to generate a charge for the click.
- Existence of *prohibited means*, such as deceptive software or a publisher clicking on the ads placed on that publisher's web site (Google explicitly prohibits this type of activity in the Terms and Conditions statement for the publishers when they sign with Google's AdSense program).

These definitions point to the problems associated with the whole effort of identifying invalid clicks. First of all, to determine if a certain click is invalid, it is necessary to understand the *intent* of generating the click: was the click generated "artificially" (improperly) or not and what does exactly "artificial" mean in this case. In certain cases the intent can clearly be determined. Positive intent can clearly be determined in such cases as when the click is eventually converted into a purchase of the advertised product or into another conversion event. Some of the negative intents can also be clearly determined. For example, Google lists several "prohibited means" (such as the ones stated in the AdSense Program Policies (https://www.google.com/adsense/policies?sourceid=asos&subid=ww-ww-et-HC_entry&medium=link) and also discussed on the AdSense page "What can I do to ensure that my account won't be disabled" (https://www.google.com/support/adsense/bin/answer.py?answer=23921&ctx=sibling)). Any click generated using these "prohibited means" is, by definition, invalid, and some of them can be detected with near-100% certainty. For example, clicks using certain types of software bots or clicks on Google's ads on the publisher's own web site constitute examples of such "prohibited means" and can be detected using technological means and marked as "invalid".

Unfortunately, in several cases it is hard or even impossible to determine the true intent of a click using any technological means. For example, a person might have clicked on an ad, looked at it, went somewhere else but then decided to have another look at the ad shortly thereafter to make sure that he/she got all the necessary information from the ad. Is this second click invalid? To make things even more complicated, the second click may not be strictly necessary since the person remembers the content of the ad reasonably well (hence there is no real need for the second click). However, the person may not really like or care about the advertiser and decides to make this second click anyway (to make sure that he/she did not miss anything in the ad and his/her information is indeed correct) without any concerns that the advertiser may end up paying for this second click (since the person really does not care about the advertiser and his/her own interests of not missing anything in the ad overweigh the concerns of hurting the advertiser). Therefore, in some cases the true intent of a click can be identified only after examining deep psychological processes, subtle nuances of human behavior and other considerations in the mind of the clicking person. Moreover, to mark such clicks as valid or invalid, these deep psychological processes and subtle nuances of human behavior need to be operationalized and identified through various technological means, including software filters. Therefore, it is simply impossible to identify true clicking intent for certain types of clicking activities and, therefore, classify these clicks as valid or invalid.

Furthermore, whether a particular click is valid or invalid sometimes depends on the parameters of the click. For example, consider the case of a doubleclick, i.e., two clicks on the same ad impression, where the second click follows the first one within time period *p*. Is the second click in a doubleclick, valid or invalid? The answer depends on the time difference *p* between two clicks. If *p* is "relatively large," e.g., 10 seconds, then the second click on the same impression *can* be valid because the visitor may click on an impression, click on the Back button of the browser and come back to the same ad impression again and wanted to have another look at the ad (for example, doing comparison shopping). However, as will be argued below, if *p* is really small, e.g. ¼ of a second, then this click can be defined as invalid (again, based on the nuances of the definition of "invalid clicks" to be discussed below). This puts us in a very uncomfortable situation of defining validity of a click based on specific values of its parameters. For example, what should the delineating value of parameter *p* be in the above example to define the second click as invalid, e.g. should it be 0.5 second, 1 second, 1.1 seconds?

In summary, between the obviously clear cases of valid and invalid clicks, lies the whole spectrum of highly complicated cases when the clicking intent is far from clear and depends on a whole range of complicated factors, including the parameter values of the click. Therefore, this intent (and thus the validity of a click based on the above definitions) *cannot* be operationalized and detected by technological means with any reasonable measure of certainty.

All the definitions of invalid clicks presented above allude to the *malicious intent* to make the advertiser pay for the click, and the absence or presence of this malicious intent differentiates fraudulent from invalid clicks. If the clicks are generated "artificially" with no possibility of conversion and only with the result of generating a charge for the click, then these clicks are *invalid.* If, in addition to this, there is also a malicious intent to hurt an advertiser or another stakeholder, these clicks are *fraudulent*. Note that "invalid clicks" is a strictly more general concept then "fraudulent" clicks because (a) the latter are invalid clicks made with a malicious intent, (b) there exist inadvertent clicking activities with no possibility of conversion that do not have a malicious intent. An example of an invalid click that is not fraudulent is the second immediate click in a doubleclick made by a person out of an old habit (e.g., he/she may usually doubleclick on all the applications, including Word, Excel and Web applications, since older versions of Windows required doubleclicks in many cases). Since this second click is made only out of an old habit, it is inadvertent and does not have intent to hurt the advertiser. Moreover, it is invalid because it does not increase the probability of a conversion: if time between two clicks on the same ad impression is too short, the visitor cannot change his or her mind whether to convert within this short time period or not. Therefore, this click is invalid but not fraudulent. Because the concept of an invalid click is broader than that of a fraudulent click, Google prefers to use the term *invalid clicks* or *spam clicks.*

These discussions have the following consequences: all the three definitions above, including two Google's definitions,

- need to be adjusted accordingly to incorporate the differences between fraudulent and invalid clicks
- are impossible to operationalize in the sense that a set of procedures (algorithms) can be developed that would detect valid and invalid clicks *always* according to the above conceptual definitions of invalid clicks.

The last statement has one important implication: given a particular click in a log file, *it is impossible to say with certainty if this click is valid or not in* all *the cases.* This means that

- It is *impossible* to measure the true rates of invalid clicking activities, and all the reports published in the business press are only *guesstimates* at best.
- The invalid click detection methods need to be developed *without* a proper operationalizable conceptual definition of invalid clicks.

The important word above is *all the cases* since in some cases it can be stated with certainty if a particular click is valid or not. For example, it is easy to detect a doubleclick using relatively simple technological means, assuming that the doubleclick is invalid.

The invalid clicks can come from the following sources:

1. individuals deploying automated clicking programs or software applications (called bots) specifically designed to click on ads
2. an individual employing low-cost workers or incentivizing others to click on the advertising links
3. publishers manually clicking on the ads on their pages
4. publishers manipulating web pages in such a way that user interactions with the web site result in inadvertent clicks
5. publishers subscribing to paid traffic websites that artificially bring extra traffic to the site, including extra clicking on the ads
6. advertisers manually clicking on the ads of their competitors
7. publishers being sabotaged by their competitors or other ill-wishers
8. various types of unintentional clicks, such as doubleclicks or customers getting confused and unintentionally clicking on the ad without a malicious intent.
9. technical problems, system implementation errors and coordination activities between Google.com and its affiliates resulting in double-counting errors
10. multiple accounts of AdSense publishers: some AdSense publishers illegally open "new" accounts under different names and using false identities; all the clicks originated from these illegal accounts are considered invalid.

Some of these invalid clicks are clearly fraudulent, while others are just invalid. Some of them are generated as a part of the AdSense while others of the AdWords program. Some of them are easy to detect, while others are very hard. The goal of the Click Quality team is to identify *all* these invalid clicks *regardless* of its nature and origin and make sure that advertisers do not pay for these invalid clicks.

This is a formidable task for many reasons, one of the main reasons being that the conceptual definitions of invalid clicks, as presented above, are impossible to operationalize in the sense that invalid click detection methods can be developed that would algorithmically identify invalid and only invalid clicks satisfying these definitions. Since it is impossible to have a working conceptual definition of invalid clicks, an alternative approach would be to provide an operational definition that can be technologically enforced. Such definitions are presented in the next section.

8.2 Operational Definitions of Invalid Clicks

An operational definition does not really say *what* invalid clicks are but specifies methods for identifying invalid clicks, thus emphasizing the *how* of invalid click detection rather than the "what" of the conceptual definition. In other words, clicks satisfying certain identification procedures are, by definition, invalid.

There are the following operational approaches to identifying invalid clicks:

- *Anomaly-based* (or *Deviation-from-the-norm-based*). According to this approach, one may not know what invalid clicks are. However, one can know what constitutes "normal" clicking activities, *assuming* that abnormal activities are relatively infrequent and do not distort the statistics of the normal activities. Then invalid clicks are those that *significantly deviate* (mainly in the statistical sense) from the established norms. For example, if a normal average clicking frequency on an ad is 4 clicks per week and if someone clicks on it 100 times per week, then this is an abnormally large clicking activity. The main challenges of this approach are how to (a) identify what the "normal" clicking activities are and (b) define what "deviation from the norm" is.

- *Rule-based*. In this approach, one specifies a set of rules identifying invalid clicking activities; alternatively, one can also identify a set of other rules identifying valid clicking activities. Each rule has one or several conditions in its antecedent and is of the form "IF Condition1 AND Condition2 AND … AND ConditionK hold THEN Click X is Invalid (or respectively Valid)." An example of such a rule is "IF Doubleclick occurred THEN the second click is Invalid." These rules are specified by invalid click detection experts based on their experiences. Therefore, these experts define what valid and invalid clicks are (note that this can be done for both valid *and* invalid clicks). These experts can be either local experts from Google or some global standardization committees that collectively develop rule-based standards of invalid clicks.

  The main challenge with this approach is to demonstrate that these conditions are "reasonable" in the sense that they are consistent among themselves and with the conceptual definition(s) specified in Section 8.1 in the following sense. If a rule of the type described above says that click X is valid (i.e., it satisfies the conditions of the rule) then it is necessary to demonstrate that it is possible to generate click X using valid (non-prohibited) means and that a non-zero probability of conversion can occur under these conditions. A similar check should be done for the rules stating when click X is invalid. For example, consider a doubleclick. Should the experts introduce the rule stating that a doubleclick is valid or not? In order to do this, it should be demonstrated that the corresponding rule is in agreement with the conceptual definition(s) of invalid clicks stated in Section 8.1. According to these conceptual definitions, if time p between the clicks is too short (e.g., less than a second) then the second click cannot affect the visitor's intention to convert that is over and above the intention associated with the first click. Therefore, the second click in the doubleclick *should* be treated as an invalid click based on the conceptual definition(s) from Section 8.1. Therefore, the *only* feasible rule-based operational definition is "IF Doubleclick(X) and p(X) is "small" (e.g., less than a second) THEN X is invalid." It turns out that Google had a history associated with the definition of a doubleclick: at some point doubleclick was considered to be a valid click and advertisers were charged for it, while subsequently Google reconsidered and treated doubleclick as invalid. This issue is discussed further in Section 9.

- *Classifier-based*. Using various data mining methods, one can build a statistical (data mining) model based on the past data that can classify new clicks into valid or invalid and also assign some degree of certainty (probability) to this classification. According to this approach, although one may not know what invalid clicks are, one can simply learn to recognize them with a certain degree of certainty based on the prior experiences of studying past clicking activities and knowing from exogenous sources which ones are truly valid and invalid. One fundamental assumption in this approach is that the past clicking behavior is indicative of the future behavior. The main problems with this approach are: (a) it is a truly operational approach: an invalid click is the one identified by the classifier, as opposed to being defined in conceptual terms based on some "higher" knowledge; (b) one needs to identify a sizable number of past clicks that are known to be truly valid and invalid, which may be an issue in some cases, as discussed above.

Google uses the first two operational approaches (anomaly- and rule-based) to define and identify invalid clicks, as will be discussed in Section 9. Google also uses a third one; but only in a couple of relatively minor cases.

One problem associated with these operational definitions is that they cannot be fully released to the general public because unethical users will immediately take advantage from knowing these definitions, which may lead to a massive click fraud. However, if it is not known to the public what valid and invalid clicks are, how would the advertisers know for what exactly they are being charged? This is the essence of the Fundamental Problem of the PPC model to be discussed in the next section.

8.3 Conclusions about Definitions of Invalid Clicks

Based on the discussions in Sections 8.1 and 8.2, we conclude that there is a *fundamental problem* associated with the definition of invalid clicks for the Pay-per-Click model. This problem can be summarized as follows:

- There is no conceptual definition of invalid clicks that can be operationalized in the sense defined above.
- An operational definition cannot be fully disclosed to the general public because of the concerns that unethical users will take advantage of it, which may lead to a massive click fraud. However, if it is not disclosed, advertisers cannot verify or even dispute why they have been charged for certain clicks.

This problem lies at the heart of the click fraud debate and constitutes the main problem of the CPC model: it is inherently vulnerable to click fraud. For this reason, we will refer to it as the *Fundamental Problem* of invalid (fraudulent) clicks.

Two possible solutions to this Fundamental Problem are:

- *The "trust us" approach* of the search engines. The search engines can assure advertisers that they are doing everything possible to protect them against the click fraud. This is not easy because of the inherent conflict of interest between the two parties: the money from invalid clicks directly contribute to the bottom lines of the

search engines. Nevertheless, it may be possible for the search engines to solve this trust problem by developing lasting relationships with the advertisers. However, the discussion of how this can be done lies outside of the scope of this report.
- *Third-party auditors*. Independent third-party vendors, who have no financial conflicts of interest, can work with advertisers and audit their clickstream files to detect invalid clicks.

These two approaches would still constitute only a partial solution to the Fundamental Problem because there is no conceptual definition of invalid clicks that can be operationalized.

9. Google's Approach to Detecting Invalid Clicks

The mission statement of the Click Quality team (as taken verbatim from one of their internal documents) states:

> Protect Google's advertising network and provide excellent customer service to clients. We do that by:
> - Vigilantly monitoring invalid clicks/impressions and removing its source
> - Reviewing all client requests and responding in a timely manner
> - Developing and improving systems that remove invalid clicks/impressions and properly credit clients for invalid traffic
> - Educating clients and employees on invalid clicks/impressions.

The Click Quality team tries to put this mission statement into practice by raising the quality of invalid click detection methods to the levels where committing click fraud against Google becomes hard and unrewarding in the sense that the cost of committing fraud (e.g., publishers being caught and terminated) significantly exceeds its benefits (earning extra money or hurting competitors). If Google can achieve this, then rational spammers will go from the Google Network to some other "weaker links" in search of easier targets.

Google tries to achieve these strategic objectives in two ways:
- *Prevention.* Discouraging invalid clicking activities on its Network by making life of unethical users more difficult and less rewarding
- *Detection.* Detecting and removing invalid clicks and the perpetrators.

In addition to launching an extensive effort to detect and remove invalid clicks, Google also tries to build other mechanisms for *preventing* invalid clicking that reduce inappropriate activities on the Google Network even *before* invalid clicks are made. Some of these preventive activities include:
- Making hard to create duplicate accounts and open new accounts after the old ones are terminated

- Making hard to register using false identities
- Development of certain mechanisms that automatically discount fraudulent activities, i.e., advertisers pay less for invalid clicks since certain invalid clicking patterns would automatically reduce costs that advertisers pay for these clicks.

In the rest of this section, I will focus on the second task of detecting and removing invalid clicks. The process of invalid click detection can be characterized by the following dimensions, capturing different aspects of this process:

- *Online filtering vs. Offline monitoring and analysis:* are there some time constraints on how fast the invalid click detection should be done? In case of the online filtering, it is crucial to detect invalid clicks fast, ideally in real-time, while in the offline case there is no "serious" time constraint on the speed of the detection process.
- *Automated vs. Manual detection:* were invalid clicks detected by a specialpurpose software or by a human expert?
- *Proactive vs. Reactive detection:* has the detection of invalid clicks occurred before or after the advertiser's complaint?
- *Where were invalid clicks made?* Were invalid clicks associated with the AdSense or AdWords programs? On which part of the Google Network were they made?

The process of detecting and removing invalid clicks consists of the following stages:

- *Pre-filtering*: removal of the most obvious invalid clicks, such as "testing" and "meaningless" clicks (to be described below) before they are even seen by the filters.
- *Online Filtering*: several online filters monitor various logs for certain conditions and detect the clicks in these logs satisfying these conditions; such clicks are marked as "invalid" and are subsequently removed.
- *Post-filtering*: offline detection and removal of invalid clicks that managed to pass the online filtering stage. This stage consists of two sub-stages:
    - *Automated monitoring* for certain additional and more comprehensive conditions than in the online filtering stage.
    - *Manual reviews* of potentially invalid clicking activities by the Operations group of the Click Quality team. These examinations are performed either
- *Proactively:* after the filtering and automated monitoring stages but *before* the customers complain about invalid clicks. This gives Google the ability to either not charge advertisers for invalid clicks if they are detected before the customers are billed or give proactive credits to their accounts for these detected invalid clicks.
- *Reactively*: examination of potentially invalid clicking activities after the customers complained about certain clicking activities and charges. This is not truly a detection process, but is rather a postfactum investigation of potentially inappropriate activities.

In the rest of this section, I describe different stages of the process presented above, starting with the pre-filtering stage.

Pre-Filtering. Certain clicks are removed immediately from the logs before they are even "seen" by the online filters. This is done in order for these clicks not to be a part of the various statistics pertaining to the performance of the filters (and thus do not distort the filter performance results). Two main categories of such pre-filtered clicks are "*test*" clicks (when a click comes from the Google IP, i.e., is generated by one of the Google employees for testing purposes). The second category constitutes "*meaningless*" clicks, clicks that were improperly recorded in the log files and whose records, therefore, have some technical problems rendering these clicks either "unreadable" or meaningless. Needless to say, advertisers are never charged for such clicks, since they are removed even before the filtering process starts.

After this first preliminary stage, the next three "lines of defense" against invalid clicks include online filtering, automated offline detection and manual offline detection, *in that order*. We describe each of these stages of defense in the next three sections.

9.1 Online Filtering

9.1.1 Review of Google's Approach. Google deploys several filters to detect and remove invalid clicks. These filters are rule-based, using the terminology of Section 8.2, and monitor various logs for certain conditions and check if the clicks in these logs satisfy these conditions. As in the case of the rule-based methods described in Section 8.2, if a click or a group of clicks satisfies these conditions, then these clicks are identified and marked as *invalid* and advertisers are not charged for them. One example of such a filter is the doubleclick rule stating that when a double click occurs on an ad, then mark the second click as being invalid. Moreover, some of the filters are not only rule-based, but also anomaly-based because the conditions of some of these rule-based filters check for certain anomalous behaviors.

The filtering process is done *online*, meaning that the detection of an invalid click should take place within a short time window since that click occurred. For this reason and because of the never-seizing arrivals of new clicks, the detection process should be efficient and scalable to very large volumes of clicks occurring on the Google Network. This process can be compared to the speed with which customers are served in queues in stores and other facilities: if the arrival rates of new customers exceed the speed with which the customers are served, the queues can grow indefinitely. Therefore, as in the case of the store queues, it is necessary to avoid processing bottlenecks in the online filters. This requirement imposes certain constraints on which methods Google can and cannot deploy for the invalid click detection purposes since the exceedingly slow filtering methods would simply lead to runaway processing delays.

Currently, Google deploys several online filters and prioritizes them by specifying the *order* in which they are used in checking invalid clicks. The invalid clicks are removed only at the end of the filtering process. Therefore, each filter "sees" every click. However, each invalid click is associated with the first filter in the packing order that detected it. It turns out that the vast majority of invalid clicks are detected by the first few most powerful filters (in the order of their prioritization), and the last few filters in the packing order detect only a small portion of invalid clicks that have not been yet detected by the previously applied filters.

When the PPC-based AdWords program was launched in February 2002, Google had only three filters, and the number and the quality of the filters steadily grew over the years. The Click Quality team constantly works on the development of new and improvement of the current set of filters using the following *feedback* process:

1. Monitor the performance of the current generation of the online filters. The invalid clicks not detected during the filtering process can still be identified "downstream" during other detection stages, including offline automated monitoring and offline manual inspection stages.
2. Examine the reasons why the current set of filters missed the invalid clicks caught downstream in the automated and manual offline detection stages. After understanding these reasons, determine whether they are actionable and could lead to the revisions of the current set of filters in order to improve the overall performance of the filtering system. Note that not all the reasons why the filters missed certain invalid clicks can be fixed by developing new or modifying existing filters. This is the case because it may be very difficult to express the filtering conditions for some of these situations. The Click Quality team looks at all the detected problems, studies them carefully, and tries to formulate these new filtering conditions or adjust the conditions in old filters, whenever possible.
3. Use the knowledge obtained in Step 2 for revising existing filters or adding new filters in order to eliminate the reasons for missing these types of invalid clicks or preventing these or similar types of attacks in the future. These revisions can be of the following type:
   (a) modify parameters of a filter
   (b) add new conditions to a filter
   (c) introduce a new filter
   (d) remove an old underperforming filter.

This monitoring-feedback-revision is an ongoing process executed in a *feedback loop*. It gives Google an opportunity to progressively improve performance of its filters over time and fix any problems missed by filters as they emerge.

The *reactive (post-factum)* improvement process of Google's filters described above is complemented by a *proactive* process of developing new filters before the actual problems occur. However, it is becoming progressively more difficult to develop new filtering ideas proactively because all the "low hanging fruits" of straightforward filtering approaches have been examined and introduced by now, and one needs to work significantly harder to develop new filters proactively.

When new filters are developed, they first undergo extensive testing before being moved into production to see how well they perform in practice. After the Click Quality team observes their performance and is convinced that the new or modified filters should be used in practice, these filters are deployed in the production mode. It turns out that only few new filters provide sufficient additional benefits in terms of detecting additional invalid clicks over and above of what the existing set of filters does already that warrant their deployment. Even those recently deployed filters provide only incremental improvements over the existing set of filters. For example, Google recently introduced a new filter that discarded x% of invalid clicks per day at the point in the ranking order where it was placed by the Click Quality engineers. If it were applied first, it would have discarded y% of invalid clicks. The ratio of x/y fluctuated between 2%-3% demonstrating that most of the invalid clicks detected by this new filter were actually detected by the previously introduced filters. This means that this new filter provided only incremental improvements over the existing set of filters. Nevertheless, Google engineers still decided to deploy it in production because they felt that it was still an important filter. Similarly, another filter also recently proposed by one of the Click Quality engineers was not moved into production because it did not contribute much over and above the existing base of filters in terms of catching new invalid clicks.

These last observations are significant since they demonstrate that the current set of Google filters is fairly stable and only requires periodic "tuning" and "maintenance" rather than a radical re-engineering, even when major fraudulent attacks are launched against the Google Network. It also demonstrates that various recent efforts of the Click Quality team to improve performance of their filters produce only incremental improvements. Thus, the Click Quality team currently reached a stability point since additional efforts to enhance filters produce only marginal improvements.

Having said this, the Click Quality team also realizes that this is only a *local* stability point in the sense that major future modifications in clicking patterns of online users and new types of fraudulent attacks against Google can lead to radically new types of invalid clicks that the current set of filters can miss. Therefore, the Click Quality team is working on the next generation of more powerful filters that will monitor a broader set of signals and more complex monitoring conditions. These new filters will require a more powerful computing infrastructure than is currently available, and the Click Quality team also participates in developing this infrastructure. Their overall goal is to make click spam hard and unrewarding for the unethical users thus making it uneconomical for them and turning many of them away from Google and the Google Network.

The reactive improvement process of Google's filters (new filters are introduced, then problems with these filters missing new attacks are detected and analyzed, and corrective actions are taken to fix these problems by improving the filters) would have been unacceptable in several other types of "detection" applications, such as fraud, virus and terrorism detection applications dealing with *irreversible* types of damages where only proactive detection methods are acceptable. This reactive approach adopted by Google, although not ideal, is nevertheless *reasonable* for invalid click detection because remedial actions are possible: once Google realizes that their filters missed invalid clicks, Google simply gives credits to the advertisers for these missed clicks and tries to fix the filters. This approach remedies the problem while producing only limited "side-effects" (such as additional concerns on the part of advertisers and the necessity for them to request refunds).

9.1.2 Performance of Online Filters. I spent a considerable time trying to understand how well Google's online filters perform, including understanding of various measures determining performance of Google's filters. In data mining and related disciplines, there exist many measures determining performance of data mining models. One of the most popular ones is the confusion matrix that is defined as follows.

A *true* click is either valid or invalid, assuming that we know the "absolute truth" about validity of all the clicks (which is not the case for Google, as discussed in Section 8). Also, Google filters can *label* a click as either valid or invalid. These two dimensions (the actual click vs. click labeling by filters), give rise to the following *confusion matrix:*

		Click classified by filters as	
		Invalid	*Valid*
Actual click	*Invalid*	True Positive (TP)	False Negative (FN)
	Valid	False Positive (FP)	True Negative (TN)

where

True Positive (TP) is an invalid click that is correctly identified as invalid
True Negative (TN) is a valid click that is correctly identified as valid
False Positive (FP) is a valid click that is incorrectly identified as invalid
False Negative (FN) is an invalid click that is incorrectly identified as valid

Given the total number of clicks N, we can identify the number of TP, TN, FP and FN clicks. Note that $TP + TN + FP + FN = N$. Then the *accuracy rate* of a filter is equal to $(TP + TN)/N$ and the *error rate* to $(FP + FN)/N$. In addition to these measures, there are several other measures that can be used for determining performance of the filters.

All these measures would have been ideal for determining performance of online filters since these are hard objective measures. Unfortunately, as explained in Section 8.1, Google does not have full knowledge of which clicks are actually valid and invalid, and it is impossible to identify performance rates of the filters without this knowledge.

Still, the Click Quality team could have conducted some studies trying to obtain this knowledge for certain samples of clicks. I have discussed these possibilities with some members of the Click Quality team. Their arguments were that it is extremely difficult to obtain this knowledge in a systematic and unbiased manner for Google (or any other search engine). For this reason, Google does not have this information about actual validity of various clicks and, therefore, cannot use the standard TP, FP, TN, FN and other measures described above to determine performance of their online filters.

I understand difficulties of obtaining systematic and unbiased samples of valid and invalid clicks for Google and the arguments made by some of the Click Quality team members. I still believe that it is possible to generate these samples and determine the appropriate error rates, although I agree that it is a difficult and a non-trivial task. I also understand that this may open Google to various criticisms regarding methodologies of generating these samples and computing performance measures for their filters. Given their list of priorities for managing their invalid click detection efforts and potential set of problems when trying to generate samples of actual valid and invalid clicks, I find their decision of not to pursue this effort now to be *reasonable*, although I don't fully agree with the Click Quality team on this point.

In the absence of hard direct statistical measures of how well Google filters perform, including rates of invalid clicks on the Google Network, the only resort for the Click Quality team to determine how well their filters work is to provide *indirect* evidence that Google filters perform reasonably well. Two main pieces of such evidence for the *filters* are:

1. Newly introduced and revised filters detect only few additional invalid clicks. As explained in Section 9.1.1, a recently introduced filter managed to detect only 2%-3% of its invalid clicks not detected by other filters already. Similarly, some newly introduced filters were not even moved into production because they hardly caught any new clicks.

2. The offline invalid click detection methods, to be described in Section 9.2, detect relatively few invalid clicks in comparison to the filters. Therefore, the online filters capture a very significant percentage of invalid clicks detected by Google. This observation does not provide irrefutable evidence that the filters work well since the previous observation can simply be attributed to the poor performance of the offline methods. However, the Click Quality team put much thought into developing reasonable offline methods. Therefore, the low ratio of the offline to the online detections provides some evidence that the online filters perform reasonably well.

In addition to these two points, the Click Quality team provided me with four additional pieces of evidence indicative of reasonable performance of invalid click detection methods. Since these pieces of evidence are applicable to the whole invalid click detection system and not just to filters, I will present them in Section 9.5 when discussing and assessing the overall performance of the invalid click detection system.

9.1.3 Simplicity of Google's Filters and the Long Tail Phenomenon. The structure of most of Google's filters, with a few exceptions, is surprisingly simple. I was initially puzzled and thought that Google did not do a reasonable job in developing better and more sophisticated filters. I was initially certain that these simple filters should miss many types of more complicated attacks. However, the evidence reported in the previous two sections indicates that these simple filters perform reasonably well. Therefore, I further examined this phenomenon and concluded that this reasonable performance is due to the following factors:

1. *Combination of filters.* Google provides *several* filters that are applied one after another. If one filter misses an invalid click, one of the "downstream" filters may detect this click and filter it out. This phenomenon of several individually simple objects *collectively* performing surprisingly well is a well-known phenomenon in science and technology. I believe that this is also the case for Google filters.
2. *Extra complexity of some of the filters.* As explained before, a few filters do have a somewhat more complex structure (although most of them don't), and this helps in detecting certain types of invalid clicks.
3. *Simplicity of most of the attacks*. Although some of the coordinated attacks can be quite sophisticated, the majority of the invalid clicks usually come from relatively simple sources and less experienced perpetrators. This is also a known phenomenon in some other professions, such as medicine, where the majority of patients' medical problems are relatively simple (such as common colds) and can be managed reasonably well by less experienced doctors, while really complicated cases arise significantly less often than these few simple and standard problems. I expect that a similar situation occurs with invalid clicks where simple Google filters detect the majority of less sophisticated attacks. Still, there are certain types of attacks that Google filters will miss; but these attacks should be quite sophisticated and would require significant ingenuity to launch. Therefore, there cannot be too many of these, unless perpetrators become much more imaginative.
4. *The Long Tail of invalid clicks*. (First of all, I would like to put a *disclaimer* that this point (#4) constitutes only my attempt to explain the performance of Google filters, and is based exclusively on my ideas and hypotheses. None of this information was provided to me by Google. Therefore, I take full responsibility for all the arguments in this report pertaining to the Long Tail concept. These arguments should be construed as "working hypotheses" and not as "hard facts.") If we plot the frequency of inappropriate activities (including fraudulent activities) on the Y-axis and rank these activities in the order of their frequency on the X-axis, then we can expect to get a distribution as shown in Figure 1 that follows the so-called Zipf Law stating that the frequency of the inappropriate activities should be inversely proportional to the ranks of these activities (*disclaimer:* this statement is purely hypothetical and constitutes only my attempt to explain the phenomenon; it is not based on any actual scientific evidence provided to me by Google or derived from any other sources). This Zipf distribution is characterized by massive amount of invalid clicks arising from a relatively few types of inappropriate activities with the smallest ranks (i.e., most frequently occurring inappropriate activities) and are

followed by the *Long Tail* of relatively few idiosyncratic types of activities that happen only infrequently. My explanation of the reasons why simple Google filters perform reasonably well is that most of the invalid clicks that Google filters out come from the Left Part of the Zipf's distribution, while the unfiltered clicks belong to the Long Tail of Figure 1. Since the Left Part consists of predominately simple inappropriate activities, this explains why a collection of simple Google filters should be able to filter out most of the invalid clicks.

These four reasons constitute my explanation why the collection of simple Google filters performs reasonably well.

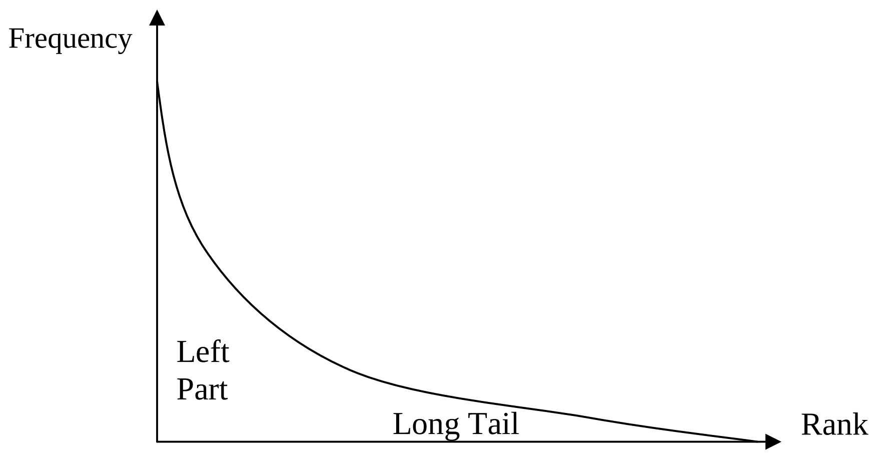

Figure 1: The Zipf's Distribution and the Long Tail of Invalid Clicks.

Despite its current reasonable performance, this situation may change significantly in the future if new attacks will shift towards the Long Tail of the Zipf distribution by becoming more sophisticated and diverse. This means that their effects will be more prominent in comparison to the current situation and that the current set of simple filters deployed by Google may not be sufficient in the future. Google engineers recognize that they should remain vigilant against new possible types of attacks and are currently working on the Next Generation filters to address this problem and to stay "ahead of the curve" in the never-ending battle of detecting new types of invalid clicks.

9.1.4 Are Google's Filters Biased? Since Google does not charge advertisers for invalid clicks, this means that it loses money by filtering out these clicks. Thus, there is a financial incentive for Google not to forgo some of these revenues and simply be "easy" on filtering out invalid clicks. Therefore, it is important to know if any business considerations entered into the filter specification process or is it entirely determined by Google's engineers in an objective manner with a *single* purpose to protect the advertiser base. This is one of the important issues that I investigated as a part of my studies of how Google manages detection of invalid clicks.

As stated before, filters are specified by engineers usually using the feedback approach described in Section 9.1.1 (although there are exceptions to this approach, such as the specification of the doubleclick filter that is discussed below). These new filters are produced by engineers in response to some previously missed attacks and, therefore, are specified with a single purpose to protect advertisers. However, some of the filters have parameters associated with them. For example, consider the following filter stating that if signal X associated with a click is above the threshold level a then mark the click as invalid. The value of this threshold parameter a determines sensitivity of the filter and how many clicks are identified as invalid. If parameter a is set low, then the filter will mark more clicks as invalid, and Google will forgo some of the extra revenues by not charging advertisers for these additional clicks. If a is set high, then fewer clicks will be marked as invalid by the filter; but advertisers may be charged for some of the truly invalid clicks missed by the filters. Thus, it is crucial to set the threshold value a properly and fairly. As stated before, determining the threshold value a is both an engineering *and* a business decision because it determines both accuracy rates of filtering out invalid clicks and extra revenues for Google from charging for additional clicks.

I have spent a significant amount of time trying to understand who sets these threshold parameters, how, and what are the procedures and processes for setting them. In particular, I tried to understand if it is an *entirely* engineering decision that tries to protect the advertisers from invalid clicks or any of the business groups at Google are involved in this decision process with the purpose of influencing it towards generating extra revenues for Google.

As a result of these investigations, I realized that it constitutes *exclusively* an engineering decision with no inputs from the finance department or the business units, except the following two cases:

- The first one was a special case when one particular IP address was disabled because of inappropriate clicking activities, and a business unit requested the Click Quality team to conduct an additional investigation since it was an important customer associated with that IP address, and restore it if the investigation results were negative. When I was explained what had happened, I felt that Google's actions were reasonable in this particular situation.
- The change in the doubleclick policy that was considered in Winter 2005 and implemented in March 2005. It turned out that the change in the doubleclick policy (i.e., not to charge advertisers for the immediate second click in a doubleclick) had non-trivial financial implications for Google. Being a publicly traded company at that time, this change would have had a noticeable effect on Google's total revenues with corresponding implications for the financial performance of the company. Therefore, this policy change had legitimate concerns for Google's management, and these financial implications have been discussed in the company. Still, despite

its noticeable negative effects on its financial performance, Google decided to abandon the old doubleclick policy and not to charge advertisers for the second click, which was an appropriate action to take.

In conclusion, with the exception of the doubleclick, I found Google's processes for specifying filters and setting parameters in these filters driven exclusively by the consideration to protect the advertiser base, and, therefore, being reasonable.

Doubleclick constitutes a special case. For me, the second click in the doubleclick is invalid, as I argued in Section 8, and the advertisers should not be charged for it. It is not clear to me why it took Google so long to revise the policy of charging for doubleclicks. Nevertheless, this policy was revised in March 2005 despite the fact that the company lost "noticeable" revenues by taking this action.

9.1.5 History of Google Filters. Whatever I have described in this section so far, constitutes the current state of affairs for Google filters. In this subsection, I will describe the history of development of Google filters. First of all, I would like to point out that most of the descriptions in this subsection are not based on documents provided to me by Google but rather on the verbal descriptions by the members of the Click Quality team based on their recollections of the past events and on the "folklore" evidence since none of the team members I interviewed were even around or involved in the click fraud effort when the AdWords program was introduced in February 2002.

Google's invalid click detection efforts started when the PPC-based version of the AdWords program was launched in February 2002. These efforts can be divided into the following three major stages:

- *The Early Days* (February 2002 – Summer 2003). These were the early days of the PPC model and of the click fraud characterized by extensive learning about the problem and determining ways to deal with it.
- *The Formation Stage* (Summer 2003 – Fall 2005). This stage started with the introduction of the AdSense program in March 2003, formation of the Google Click Quality team in the Spring/Summer 2003, launch of new filters and the intent to take the invalid click detection efforts to the "next level." It ended with the development of the whole infrastructure for combating invalid clicks and the consolidation of Google's invalid click detection efforts. This stage was characterized by significant progress in combating invalid clicking activities and developing mature systems and processes for accomplishing this task. Although the Click Quality team's solutions were still not perfect, based on the information provided to me by Google, I reached the conclusion that the invalid clicking problem at Google was "under control" by the end of 2005.
- *The Consolidation Stage* (Fall 2005 – present). By this time, Google had enough filters and perfected them to the level when they would detect most of the invalid clicking activities in the Left Part of the Zipf distribution (see Figure 1) and some of the attacks in the Long Tail. They would still miss more sophisticated attacks

in the Long Tail, and the Click Quality team continued working on the neverending process of improving their filters to detect and prevent new attacks. The Click Quality team has also been working on enhancing their infrastructure and improving their processes and methods for doing offline analysis and handling customer inquiries.

In the rest of this subsection, I will describe each of these stages.

The Early Days (February 2002 – Summer 2003). When AdWords program was launched in February 2002, Google had three filters installed at that time. These filters detected and removed only the very basic invalid clicks. Looking back at these early days of invalid click detection, it is not clear to me why Google engineers could not conceive and introduce some of the subsequently developed filters which are pretty basic and obvious, having the hindsight that we have now. Also, their invalid click detection efforts were quite slow at that time: during these 1.5 years no new filters were introduced, and the whole invalid click detection effort was based only on the three filters introduced during the AdWords launch in February 2002.

There are several extenuating circumstances that might have caused such a slow start:

- Click fraud was a really new phenomenon at that time, much less understood than it is now; therefore Google engineers were on a learning curve trying to understand the problems associated with click fraud and the ways to combat it. Moreover, when Google launched the original version of the AdWords program in 2000, it was based on the CPM, and *not* the CPC advertising model. Click fraud is quite different for the CPM than for the CPC model, which means that Google engineers had to learn about new types of the CPC-related fraud at that time. This switch and the related uncertainties might have also slowed their efforts to develop new CPC-based filters.
- Google was a much smaller and different company than it is now. It had much fewer financial, human and other resources, and these limited resources were significantly stretched back in 2002 when Google tried to allocate them among so many initiatives and projects at that time.
- To take the invalid click detection effort to the next level, Google needed to build an appropriate infrastructure, which might have been difficult for them to accomplish at that time because of the lack of resources and of the click fraud experience.
- Click fraud was of a different type in 2002 than it is now and invalid clicking was on a different scale than it exists now. It is quite conceivable that the initial three filters operated better and caught a larger percentage of invalid clicks back in 2002 than they would do so now since fraud patterns changed significantly since that time (the shape of the Zipf's distribution in Figure 1 might have been significantly different in those days). However, I could not examine appropriate data that would either support or refute this hypothesis and, therefore, my statement is purely hypothetical.

Unfortunately, it is hard to gather evidence supporting or refuting these claims because these events took place long time ago (measured in "Google time"). In fact, not a single person on the Click Quality team was either around or involved in the click fraud detection back in 2002. The only person from this era who is still at Google is on an extended leave and was not available for comments during my visits to Google.

It is hard to judge reasonableness of Google's invalid click detection efforts between 2002 and summer 2003 because there is simply not enough information available for this time period for me to form an informed judgment about this matter. One exception is the doubleclick policy that I have described before. As I have already stated, the second click in the doubleclick is invalid in my opinion, and Google should have identified it as such well before March 2005 (however, the detection and filtering out the third, fourth and other subsequent clicks was there since the introduction of the PPC model, and advertisers were not charged for these extra clicks).

The Formation Stage (Summer 2003 – Fall 2005). This stage started with the introduction of the AdSense program in March 2003 and the formation of the Google Click Quality team in the Spring/Summer 2003 (the first person was hired in April 2003 with the mandate to form the Click Quality team; several people joined the team during the summer of 2003, and the initial "core" team consisting of Operations and Engineering groups was consolidated by Fall 2003).

During this time period, two new filters were introduced in Summer 2003 and one more in January 2004. These three new filters remedied several problems that existed since the launch of the first three filters and significantly advanced Google's invalid click detection efforts. Besides the development of new and better filters, there was a separate effort launched to develop the whole infrastructure for doing the offline analysis of invalid clicks and managing customer inquiries about invalid clicks and billing charges.

Despite all these efforts, the new filters and the offline analysis methods still failed to detect some of the more sophisticated attacks (presumably from the Long Tail of the Figure 1) launched against the Google Network in 2004 and the first half of 2005. In response to these activities and as a part of the overall invalid click detection effort, Google engineers introduced some additional filters around Winter and Spring 2005, including the filter identifying the second immediate click in a doubleclick as invalid.

As a result of all of these efforts by the Click Quality team, a significant progress has been made in combating invalid clicking activities and developing mature systems and processes to accomplish this task. Although the Click Quality team's solutions were still not perfect, based on the information provided to me by Google, I reached the conclusion that the invalid clicking problem at Google was "under control" by the end of 2005.

The Consolidation Stage (Fall 2005 – present). By the end of 2005, all the major components of the invalid click detection program were in place, and Google had revised its doubleclick policy. There was evidence (as documented in Section 9.1.2) that the invalid click detection efforts worked reasonably well by that time. Therefore, Google entered the stage when it needed to fine-tune its current methods and prepare for the next level of more sophisticated attacks by unethical users, most likely belonging to the Long Tail of Figure 1. Currently, the Engineering unit of the Click Quality team is developing the Next Generation of Google filters designed for that purpose.

9.1.6 What is Missing in Google Filters. Although Google filters work reasonably well now, I found the following functionality not currently supported by them:

1. *Deployment of Data Mining Methods.* Google filters are rule-based and also anomalybased, as discussed in Section 9.1.1 (see Section 8.2 for the explanation of the rule-based and the anomaly-based approaches). In addition to these two approaches, Google can also develop classifier-based filters according to the principles discussed in Section 8.2 that are based on well-known data mining methods. These data-mining-based filters would classify the incoming clicks as valid or invalid with some degree of certainty and would filter out those clicks about which the classifiers are fairly certain that they are invalid. There exists a whole range of techniques developed in the statistical, machine learning and data mining communities over the last few decades on how to do it. The most challenging and contentious issue in building such classifiers is a *balanced* collection of truly valid and invalid past clicks for "training" the classifier. If the sample of these truly valid and invalid clicks is not balanced, then the resulting classifier built using this sample will be skewed and will produce poor results filtering invalid clicks. I discussed this issue at length with some of the members of Google's Click Quality team, and we had different views on the feasibility of building such a classifier for detecting invalid clicks at Google. I fully understand and respect their arguments. Nevertheless, I differ with them in my opinions on this matter.

2. *Using the Conversion Data in Filters*. None of the filters uses the conversion information that Google collects (if a click is followed by a conversion event on the advertiser's web site). This is the case because (a) only a fraction of clicks has the conversion information associated with them; (b) the majority of conversions occur only after a significant time period after a click on an ad occurred. Since the filters have a limited time window to decide if a click is valid or not (as discussed in Section 9.1.1), this means that the filters simply don't know if the conversion will take place or not by the time they need to make the decision. There are other, more technical reasons, why the Google engineers decided not to use the conversion data in filters. Nevertheless, I still think that the conversion data should be used in filters, even if its usage is limited.

3. *Developing More Advanced Types of Filters*. As I stated in Section 9.1.3, Google filters are quite simple. Despite its simplicity, they work reasonably well and detect a significant amount of invalid clicks, presumably, mainly in the Left Part and also some in the Long Tail of the Zipf distribution in Figure 1. However, to prepare for the "next level" of more sophisticated attacks in the future, Google should develop the next generation of more

advanced filters to stay "ahead of the curve" on detecting invalid clicks. As I stated before, the Click Quality team is currently working on the development of such methods.

I discussed these issues with some of the members of the Click Quality team. We were in agreement with some of these points, while had differences in opinions on some other issues. However, none of the observations made in this section (9.1.6) and the fact that Google does not support any of the functionality described in this section (9.1.6) imply that Google's efforts to detect invalid clicks are unreasonable.

***Conclusions*.** Google put much effort in developing infrastructure, methods and processes for detecting invalid clicks since the Click Quality team was established in 2003. These efforts were not perfect since Google missed certain amounts of invalid clicks over these years and it adhered to the doubleclicking policy for too long in my opinion. However, click fraud is a very difficult problem to solve, Google put a significant effort to solve it, and I find their efforts to filter out invalid clicks as being *reasonable*, especially after the doubleclick policy was reversed in March 2005.

9.2 Offline Detection Methods

The online stage of the process of detection and removal of invalid clicks is followed by the offline stage. In this stage, there are no real-time constraints on how fast the deployed methods should be able to detect invalid clicks. Therefore, more extensive and more computationally involved detection methods can be deployed in the offline stage without any time limits imposed on the analysis process. In particular, the analysis of invalid clicks can be performed over a larger set of clicking data and over a longer time horizons than in the online filtering stage. Also, many more factors can be considered as a part of this analysis. This lack of computational constraints and the deployment of more extensive clicking data results in better analysis and better detection methods that could determine additional invalid clicks not detected by the online filters.

The offline detection methods can be characterized by the following two dimensions:

- *When the detection occurred*: before the customer complained or after. The two alternatives are:
    - *Proactively*: detection methods are applied before the customers complain about invalid clicks.
    - *Reactively*: investigation of invalid clicking activities occurs after a customer complains and as a response to this complaint. This is not truly an invalid click *detection* method, but is rather a post-factum analysis and investigation of inappropriate clicking activities.
- *Means of analysis*:
    - *Automated*: detection of invalid clicks is done by a software system.
    - *Manual*: detection is done by a human inspector who investigates a reported problem.

When studying interactions between these two dimensions, I would like to point out that all the reactive analysis is done manually, which also implies that the automated analysis can be done only proactively since there is no automated reactive analysis.

I next describe automated offline detection methods (which are proactive based on the previous comment) and then the manual inspections.

9.3 Automated Offline Detection Methods

Google deploys the following two types of offline detection systems:

- *Alerts*: are used for detecting more complex and more subtle patterns of invalid clicking activities that may or may not be valid (there is simply not enough evidence that these clicks are invalid). Since these clicks cannot be safely removed by filters, the filters pass them as valid, and it is the job of alerts to identify them in the offline analysis stage and pass these suspicious clicks to human experts for manual investigations.
- *Auto-termination system for publishers*: This automated system detects suspicious AdSense publishers who are either automatically terminated, are warned, or are subsequently investigated manually, depending on how serious their inappropriate activities are.

In the rest of this section, we describe these two automated systems.

9.3.1 Alerts

There are two types of alerts:

- Those that monitor various invalid clicking detection activities and warn the Click Quality team if some of these activities go wrong. For example, such an alert may warn the team if any of the database servers are down or some disks are full.
- Those that monitor Google's logs for abnormal querying or clicking activities.

Although both types of alerts are relevant, I will focus on the second type of an alert in this report because they contribute more to the invalid click detection efforts.

This second type of an alert checks for various complex conditions – more complex than the ones used in filters. The values of the threshold conditions in these alerts can be set more "aggressively" because the alerts do not actually filter out any clicks but rather alert human inspectors about abnormal activities so that they can study the causes of these alerts and decide on appropriate actions. Finally, these alerts take into the consideration a broader set of deciding factors and can monitor these factors over longer time periods. Therefore, these alerts provide the second “line of defense” against invalid clicks by doing additional type of analysis that is different from the type of monitoring that filters do. Thus, the alerts are able to catch some of the additional invalid clicks that filters missed.

Google engineers provided me with an example of a certain set of invalid activities against an advertiser that arrived from multiple IPs in a semi-coordinated manner. Google filters missed these invalid clicks, while the alerts caught them because they checked for a different set of conditions in a manner that filters could not do for various technical reasons. Therefore, the alerts could “connect the dots” better than filters in this particular case and could detect the aforementioned invalid clicking activities. This demonstrates that filters and alerts complement each other in the process of detecting invalid clicks and, therefore, both of them are needed in this process.

Alerts are issued in two ways:

- Placed in some log that Click Quality inspectors can examine using some browsing and querying tools
- Periodically delivered over email to particular Click Quality personnel for subsequent investigations.

Therefore, when alerts are issued, they are subsequently manually investigated by the Click Quality team, based on their priority, to determine what caused the alert and which corrective action (if any) should be taken.

The first alerts were introduced in the fourth quarter of 2005 and were subsequently improved and enhanced since that time. The type of the attack described above was detected only recently using a newly introduced type of an alert.

9.3.2 Auto-Termination System for AdSense Publishers

Initially, all the terminations of the AdSense publishers for inappropriate behavior were done manually. Currently, it is a mixture of manual and automated terminations, with the auto-termination rates growing steadily.

Auto-Termination System is an automated offline system for detecting the AdSense publishers who are engaged in inappropriate behavior violating the Terms and Conditions of the AdSense program. It examines online behavior of various publishers and either immediately terminates or warns the publishers who are engaged in the activities that the system finds to be inappropriate.

More specifically, the Click Quality team has developed a set of conditions indicative of a strong possibility of inappropriate behavior of the publishers. If certain combinations of these conditions hold, the Auto-termination system would take one of the following actions depending on the severity of these conditions:

- *Automatically terminate* the publisher if the violating conditions are really severe;

- *Automatically warn* the publisher if the violating conditions are indicative of inappropriate activities but are not as severe as in the previous case. This warning happens when certain "flags are raised," but not enough hard evidence is accumulated to be certain that the publisher is engaged in inappropriate activities. As a part of the warning, Google requests the publisher to disengage from these activities and gives a grace period to the publisher. If these inappropriate activities do not stop within a certain time period, the publisher is terminated by the autotermination system.
- *Request for a Manual Inspection:* Pass the publisher's case for a manual inspection by the team of Google's investigators in case the auto-termination system does not have strong evidence to terminate or even warn the publisher. This request is placed in the inspection queue and is subsequently retrieved and inspected by one of the Click Quality investigators using the inspection tools described in Section 9.4.

The decision to terminate, warn or manually inspect the publisher is based on a set of various conditions pertaining to publisher's behavior that were developed by Google's Click Quality team based on their extensive prior experiences in dealing with the AdSense publishers.

The first prototype of the auto-termination system was built in the early 2005 and the system was launched in the summer 2005. Recently, Google has developed major enhancements to the current version of the auto-termination system deploying an alternative set of technologies.

9.4 Manual Offline Detection Methods

Both the advertisers and the publishers can be investigated for the invalid clicking activities that either happened to or originated by them. Investigation requests are generated from various sources. In particular, investigations of *advertisers* come from the following sources:

- *Advertiser complaints*: an advertiser notices unusual clicking activities and requests Google to investigate those activities for the presence of invalid clicks.
- *Alerts*: alerts detect unusual patterns of behavior of advertisers and trigger manual investigations of these patterns.
- *Customer service representatives*: they may request to investigate an advertiser based either on the advertiser's request or based on their own initiatives.

Investigations of the *publishers* come from the following sources:

- *Publisher's complaint*: publisher notices some suspicious activities on his/her site and asks Google to investigate them.
- *Advertiser's complaint*: an advertiser notices some suspicious clicking activities on its ads coming from a certain publisher and requests Google to investigate that publisher.

- *Auto-termination system*: the auto-termination system requests a manual investigation of a publisher in those cases when it cannot automatically terminate a publisher, as described in Section 9.3.2.
- *Classifier*: Google has an automated system that examines publishers' behavior, as described in Section 9.3.2 and classifies publishers as possible spammers or "clean" publishers. If a publisher is classified as a spammer, that publisher is subsequently being investigated.
- *Detection of duplicate publishers*: Google has a system that detects multiple publishing accounts opened by the same person or an entity. Such cases are manually inspected after detection.
- *Second-review publishers*: some publishers, who had prior disputes with Google, request Google to be re-investigated.
- *Customer service representatives:* Google's CSRs may notice suspicious activities on the publishers' websites and issue requests to investigate these publishers.
- *Requests from the Click Quality team:* in some cases, members of the Click Quality team noticed some suspicious activities on the part of the publishers. An investigation request is generated for such publishers by Click Quality members in such cases.

These investigations can be proactive or reactive, i.e. in response to the advertiser's inquiry about suspicious activities or charges. Google's goal is to do as many of these investigations proactively as possible, which is indeed the case since many of the investigations listed above are indeed proactive. Another goal is to investigate the suspicious publishers in the early stages of their inappropriate activities *before* they are paid for these activities by Google.

Once a request to do an investigation is submitted to the Click Quality team, it is being prioritized and entered into a queue. The Click Quality team has developed a whole process of how these investigation requests propagate through the system and being eventually handled by various members of the Operations unit of the Click Quality team.

Also Google has developed several Inspection Systems that allow members of the Click Quality team to investigate different inspection requests. Depending on the nature of this request (see above), different Inspection Systems are used by Click Quality investigators since each inspection system deals with only specific types of investigations. Although Google has several types of inspection systems, the most important and the most frequently used ones are those that investigate:

- *Advertisers,* i.e., invalid clicking activities pertaining to particular advertisers.
- *Publishers,* i.e., invalid clicking activities associated with particular publishers.
- *Duplicate accounts,* i.e., whether a particular individual or an entity has duplicate publishing accounts or had a previously terminated publishing account(s) with Google.

In addition, the Engineering team has a general inspection system that allows them to investigate various types of abnormal activities detected and reported by automated invalid click detection systems.

All these inspection systems constitute some kind of browsing and reporting tools (reminiscent of various commercially available Business Intelligence products) that were developed in-house by the Click Quality team and that allow the Click Quality investigators quickly and visually examine various clicking, querying and browsing activities of different entities (publishers, advertisers, users, etc.) and try to discover unusual patterns of behavior indicative of inappropriate activities.

The basic idea behind most of these investigations is to discover *unexpected* behavior of the entities being investigated (such as publishers, users, etc.). Based on an extensive experience that the Click Quality team has developed investigating very large numbers of requests and based on certain good understanding of “normal” clicking, querying and browsing activities on the Google Network, the Click Quality investigators look for the deviations from these “normal” behaviors using the inspection tools described above. Once such deviations are discovered, the investigator “drills down” into the problem and uncovers the reasons causing these deviations and, most likely, the source and reasons for the inappropriate activity or a set of activities.

The outcomes of these investigations is the determination of whether

- The invalid clicks are present
- No invalid clicks are present
- It is unclear if invalid clicks are present

The first two cases lead to the obvious actions. The last case constitutes a special situation that is subsequently studied by several additional members of the Click Quality team. If the team still cannot reach a definitive conclusion, then a “benefit-of-a-doubt” action is taken. For example, in the case of an advertiser inquiry about invalid clicking activities, the advertiser is given credits for those clicking activities that the Click Quality team has not resolved as being valid. Similarly, if clearly documented inappropriate activities are detected for a publisher, the publisher’s account is terminated by the Click Quality team. If they cannot be clearly documented, then the publisher is issued a warning and being “watched” by the Click Quality team. If the publisher continues inappropriate activities over some time, he/she is being subsequently terminated. When a publisher is terminated, *all* the clicks (valid and invalid) from the terminated publisher within a certain time period are credited back to the affected advertisers.

These inspection systems have been developed by Google over an extensive period of time and are constantly improved to extend their functionality and make them better for the investigators to do their inspections more effectively.

I have personally observed several such inspections and can attest to how successfully they have been conducted by Google's investigators. This success can be attributed to (a) the quality of the inspection tools, (b) the extensive experience and high levels of professionalism of the Click Quality inspectors, and (c) the existence of certain investigation processes, guidelines and procedures assisting the investigators in the inspection process.

Some additional evidence that the offline inspection methods work reasonably well:

- Small reinstatement rates for previously terminated publishing accounts for the AdSense program. Previously terminated AdSense publishers can appeal to Google, and their requests are investigated together with reasons of why their accounts have been terminated. If the Click Quality team had terminated such an account for an invalid reason, such an account is reinstated. This actually happens periodically, but the reinstatement rates are quite low. I realize that this is not a highly reliable reason since it can be interpreted as Google being excessively defensive about reinstating previously terminated publishers. However, based on the evidence that I have seen, I think that the Click Quality inspectors try to be fair to both publishers and advertisers and approach this problem very professionally.
- The Click Quality team applies sampling methods to select random AdSense publishers and see how well the Click Quality investigators would detect spammers in this random sample. They compare spamming publishers' detection rates for these samples against their overall detection rates. The results are comparable.

My only concern with these manual inspections is about scalability of the inspection process. Since the number of inquiries grows rapidly, so does the number of inspections required to investigate these inquiries. As stated before, Google tries to automate this process by letting software systems do a sizable number of inspections. Still, the number of manual inspections keeps growing significantly over time, based on the numbers that I have seen. This means that Google has a challenging task of expanding and properly training its team of inspectors to assure rapid high-quality inspections of inquiries in the future.

One of the complaints about Google's investigation system that I keep hearing is that Google is quite secretive and does not provide meaningful explanations of the inspection results neither to the advertisers nor to the publishers. After examining how their inspection systems work, I can understand this secrecy. If Google provides such explanations, then the unethical users can gain additional insights into how Google invalid click detection methods work and would be able to "game" their detection methods much better, thus creating a possibility of massive click fraud. To avoid these problems, Google prefers to be secretive rather than to risk compromising their detection systems and the advertiser base.

Finally, I would like to point out that when Google terminates an AdSense publisher, *all* the clicks generated at that publisher's site over a certain time period (valid and invalid) are credited to the advertisers whose ads were clicked on that site.

9.5 Performance of Invalid Click Detection Methods

The performance of online filters was discussed in Section 9.1.2. For the reasons presented in Section 8, it is hard to come up with good direct and objective performance measures of these filters, such as accuracy and error rates. Therefore, Google engineers resort to the *indirect* performance measures of the filters, such as the following measures, that provide only *some* evidence that the filters perform reasonably well:

1. Newly introduced and revised filters detect only few additional invalid clicks. As explained in Section 9.1, a recently introduced filter managed to detect only 2%-3% of its invalid clicks not detected by other filters already. Similarly, some newly introduced filters were not even moved into production because they hardly caught any new clicks.

2. The offline invalid click detection methods, described in Section 9.2 detect relatively few invalid clicks; therefore, the online filters capture a very significant percentage of detected invalid clicks. This observation does not provide irrefutable evidence that the filters work well since it can simply be attributed to the poor performance of the offline methods. However, the Click Quality team put much thought into developing reasonable offline methods. Therefore, even if they did not perform that well, the low ratio of the offline to the online detections of invalid clicks would still provide *some* evidence that the online filters perform reasonably well.

In addition to these two arguments, the Click Quality team provided me with the following additional indicators supporting the claim that Google's whole invalid click detection system performs reasonably well:

3. The number of inquiries about invalid clicks for the Click Quality team increased drastically since late 2004. However, the number of refunds for invalid clicks provided by Google did not change significantly over the same time period. Therefore, the number of refunds per inquiry decreased drastically since late 2004. Since each inquiry about invalid clicks leads to an investigation, this means that significantly fewer investigations result in refunds. This statistic can be interpreted in several ways. First, it can be an indication that Google's invalid click detection methods have significantly improved over this time period and that reactive investigations do not find any problems when searching for invalid clicks. Second, this statistic can mean that Google tightened its refund policies and is less generous with its refunds than it used to be. Third, this statistic can mean that more advertisers are looking more carefully into their logs and are more suspicious about invalid clicks since this problem received wide attention in the media and the public discourse in general. Therefore, they may request Google to investigate suspicious clicking activities even if nothing really happened. I examined investigative activities of the Google Click Quality team and can attest that it consists of a group of highly professional employees who do their investigations carefully and professionally. Therefore, I do not believe in the second

reason stated above. The third reason is quite possible since advertisers are indeed concerned about invalid clicks and request Google to investigate suspicious clicking activities more frequently than before. However, the number of inquiries increased so significantly that I would expect that the number of refunds would also increase somewhat. Since this did not happen, I attribute this effect to the fact that Google's invalid click detection methods work reasonably well by now.

4. The total amount of reactive refunds that Google provides to advertisers as a result of their inquiries is miniscule in comparison to the potential revenues that Google foregoes due to the removal of invalid clicks (and not charging advertisers for them). The number of inquiries about invalid clicks increased drastically since late 2004, as I indicated in Point 3, showing that advertisers are paying more attention to invalid clicking activities (and also perhaps due to the growth of the advertiser base), especially since click fraud attracted much attention lately. Also, the Click Quality team does a careful and professional analysis of these inquiries based on my knowledge of their activities. These two observations put together imply that the total amount of refunds provided by Google can be used as an indirect proxy of how many invalid clicks Click Quality team fails to detect and remove proactively. I understand that this statistic is far from perfect as a proxy for many reasons. Nevertheless, it provides some indirect evidence that Google filters work reasonably well.

5. As explained in Section 9.4, the Click Quality team conducts Quality Assurance offline analysis of the clicking traffic by periodically sampling certain clicking activities, passing these cases to the Click Quality investigators who examine them for the presence of invalid clicks and thus estimate how many invalid clicks were missed by the offline filters. As explained in Section 9.4, the results of these tests demonstrate that the invalid click detection methods perform reasonably well.

6. Another indirect piece of evidence provided to me by Google is that Conversions-PerDollar (CPD) rates on various partner sites of Google Network are not significantly lower than on their "flagship" Google.com site. CPD is the statistic determining the number of conversions that occurred divided by the dollar amount spent on advertising. This statistic shows how effective advertising campaigns are for the advertisers. Since Google spent much effort over the past 4.5 years to make sure that Google's AdWords program works reasonably well, it now serves as the "golden standard" against which other programs are compared at Google. Since CPD numbers for other parts of the Google Network approach that of at Google.com, this is an indication that other advertising programs work as well as AdWords works on Google.com. Since other parts of the Google Network are affected by invalid clicking activities significantly more than Google.com, this is an indication to the Click Quality team that their efforts to combat fraud on other parts of the Google Network are as effective as on Google.com. This is another indirect piece of evidence that Google's efforts to detect invalid clicks on the rest of the Google Network are as effective as on Google.com.

Conclusions about the performance of invalid click detection methods. As a scientist, I am accustomed to seeing more direct, objective and conclusive evidence that certain methods and approaches "work." Having said this, I fully understand the difficulties of obtaining such measures for invalid clicks by Google, as previously discussed in this report. Moreover, one can challenge most of the reports pertaining to invalid clicking rates published in the business press by questioning their methodologies and assumptions used for calculating these rates. Most of these reports would not stand hard scientific scrutiny.

Still, as a scientist, it is hard for me to arrive at any definitive conclusions beyond any reasonable doubt based on Points (1) – (6) above that Google's invalid click detection methods "work well" and remove "most" of the invalid clicks – the provided evidence is simply not hard enough for me, and I am used to dealing with much more conclusive evidence in my scientific work.

Having said this, the indirect evidence (1) – (6) specified above, nevertheless, provides *a sufficient degree of comfort* for me to conclude that these filters work reasonably well. Finally, this statement should *not* be interpreted as if I find Google's effort to detect invalid clicks (a) unreasonable, or (b) not working reasonably well. It only states that Google did not provide a compelling amount of conclusive evidence demonstrating the effectiveness of their approach that would satisfy me as a scientist.

Finally, the measures (1) – (6) above are only *statistical* measures providing some evidence that Google's filters work reasonably well. This does not mean, however, that any particular advertiser cannot be hurt badly by fraudulent attacks, given the evidence that Google filters "work." Since Google has a very large number of advertisers, one particular bad incident will be lost in the overall statistics. Good performance measures indicative that filters work well only mean that there will be "relatively few" such bad cases. Therefore, any reports published in the business press about particular advertisers being hurt by particular fraudulent attacks do not mean that the phenomenon is widespread. One simply should not generalize such incidents to other cases and draw premature conclusions – we simply do not have evidence for or against this.

9.6 Economic Considerations Pertaining to Detection of Invalid Clicks

Since invalid click detection methods have a direct impact on Google's revenues, I also examined some of the economic consequences of detecting invalid clicks. I present some of my findings in this section based on the performance data over the past 12 – 18 months provided to me by the Click Quality team.

First of all, most of the revenue that Google foregoes due to discarding invalid clicks comes from the filters since they identify most of the invalid clicks. The second source of the forgone revenues comes from the terminated AdSense publishers (as stated before, *all* the clicks made on the terminated publisher's website generated over a certain time period are credited back to the advertisers *regardless* of whether they are valid or invalid). However, this second type of revenue is relatively small in comparison to the foregone revenues due to filters. The third source of the foregone revenues comes from the AdWords credits. However, these AdWord credits are miniscule in comparison to the other sources of foregone revenues. In summary, the most significant source of foregone revenues, by far, are Google filters. Hence their performance is the most crucial factor for the whole invalid click detection program (note that this observation does not mean that Google focuses mainly on this part of the invalid click detection program since other parts are also important).

Second, as I concluded in Section 9.1, the invalid click detection process is currently driven by the Click Quality team with the major objective to protect advertisers and other stakeholders against invalid clicks; it is not being influenced by Google's business units or the finance department, except the two cases reported in Section 9.1.4. The first one was a relatively minor case where Google's actions were understandable in my opinion.
The second one pertains to charging advertisers for doubleclicks and is more serious. As I stated in Section 9.1.4, it is unclear to me why it took Google so long to revise the policy of charging for doubleclicks.

Third, based on the numbers provided to me by Google for the last few quarters, I conclude that the amount of revenues that Google forgoes for crediting advertisers for invalid clicks is insignificant in comparison to the amount of revenues Google risks to lose if it loses trust of the advertisers. Therefore, it makes no business sense for Google to go after these extra revenues and that the best long-term business policy for Google is to protect advertisers against invalid clicks. Policy reversal on the doubleclick is a good example of this. By not charging advertisers for the doubleclick since March 2005, Google lost a "noticeable" amount of revenues. However, the revenues lost as a result of this action are insignificant in comparison to the revenues that Google risks to lose if it loses trust of the advertisers. Therefore, reversing the doubleclick policy makes sense not only from the legal, ethical and public relations point of view, but it is also a sound economic decision.

The economic consideration described above is aligned with the legal consideration of risking legal actions if Google does not do a reasonable effort to protect advertisers against invalid clicks. It is also aligned with the ethical, public relations and marketing considerations of serving and satisfying the needs of its advertising customers. Therefore, based on all these economic, legal, ethical and public relations considerations, the best long-term business strategy for Google is to protect its advertiser base against invalid clicks in the best possible manner.

9.7 History of Invalid Click Detection Efforts

In Section 9.1.5, I have already described the history of developing Google filters and identified three stages of this process. In this section, I will enhance this history to the entire invalid click detection effort and will follow the three-stage framework described in Section 9.1.5.

The Early Days (February 2002 – Summer 2003). These were the early days of the PPC model and of the click fraud that immediately followed the launch of the revamped AdWords program. The main invalid click detection activities focused on filters at that time. There was no significant infrastructure developed for dealing with invalid clicks, partially, because these invalid activities were so new and Google was still learning about them. In particular, the Click Quality team was not formed at that time, and customer inquiries were handled by the Customer Service Representatives during that period.

The Formation Stage (Summer 2003 – Fall 2005). This stage started with the introduction of the AdSense program in March 2003, formation of the Google Click Quality team in the Spring/Summer 2003, launch of new filters and the intension to take the invalid click detection efforts to the “next level.”

The Click Quality team consisted on the Engineering and Operation groups. While the Engineering group focused on the development of online filters and other invalid click detection software, the Operations group focused more on the offline detection methods and on the development and implementation of proper inspection methods and processes.

This stage ended with the development of the whole infrastructure for combating invalid clicks and the consolidation of Google’s invalid click detection efforts. This stage was characterized by significant progress in combating invalid clicking activities and developing mature systems and processes for accomplishing this task, including the development of the whole system of inspections of invalid clicking inquiries by the Operations group.

Although the Click Quality team’s solutions were still not perfect, based on the information provided to me by Google, I reached the conclusion that the invalid clicking problem at Google was “under control” by the end of 2005. In particular, several massive attacks were launched against the Google Network in 2005, and Google managed to detect and remove large volumes of invalid clicks at that time: one can clearly see major spikes on the charts plotting detected invalid clicks during this time period. This indicates that, although not perfect, Google detection software managed to remove massive amounts of invalid clicks during these attacks.

The Consolidation Stage (Fall 2005 – present). By this time, Google’s infrastructure for detecting invalid clicks has been established and needed to be consolidated at this point. Google had enough filters and perfected them to the level when they would detect most of the invalid clicking activities in, presumably, the Left Part of the Zipf distribution (see Figure 1) and some of the attacks in the Long Tail. These filters would, presumably, miss more sophisticated attacks in the Long Tail, but the Engineering unit of the Click Quality team continues working on the never-ending process of improving the filters to detect and prevent new attacks. Similarly, the Operations unit has been working on further improving the offline invalid click detection and inspection processes and on developing various enhancements to their infrastructure and to their customer inquiries management systems and processes.

10. Conclusions

As explained in Section 8, all the conceptual definitions of invalid clicks assume human intent. This means that none of these definitions can be operationalized in the sense that invalid click detection methods can be developed that would algorithmically identify invalid and only invalid clicks satisfying these definitions. This is the fundamental problem of invalid clicks that makes click fraud a difficult problem to solve.

In the absence of a conceptual operationalizable definition of invalid clicks, an alternative approach is to use operational definitions of invalid clicks that can be of the following form:

- *Anomaly-based* (or *Deviation-from-the-norm-based*). A click or a group of clicks is invalid if its behavior significantly deviates from the normal behavior, where normal behavior is established based on the average day-to-day activities.
- *Rule-based*. A click or a group of clicks is invalid if it satisfies certain conditions defined by human experts. These experts can be either local experts from Google or some global standardization committees that collectively develop rule-based standards of invalid clicks.
- *Classifier-based*. A click is invalid if a data mining classifier labels it as "invalid." This labeling is done based on the past data about valid and invalid clicking activities used for "training" the classifier to decide which clicks are (in)valid.

Google has built the following four "lines of defense" against invalid clicks: pre-filtering, online filtering, automated offline detection and manual offline detection, in that order. Google deploys different detection methods in each of these stages: the rule-based and anomaly-based approaches in the pre-filtering and the filtering stages, the combination of all the three approaches in the automated offline detection stage, and the anomaly-based approach in the offline manual inspection stage. This deployment of different methods in different stages gives Google an opportunity to detect invalid clicks using alternative techniques and thus increases their chances of detecting more invalid clicks in one of these stages, preferably proactively in the early stages.

Since its establishment in the Spring and Summer of 2003 the Click Quality team has been developing an infrastructure for detecting and removing invalid clicks and implementing various methods in the four detection stages described above. Currently, they reached a consolidation phase in their efforts, when their methods work reasonably well, the invalid click detection problem is "under control," and the Click Quality team is fine-tuning these methods. There is no hard data that can actually prove this statement. However, indirect evidence provided in this report supports this conclusion with a moderate degree of certainty. The Click Quality team also realizes that battling click fraud is an arms race, and it wants to stay "ahead of the curve" and get ready for more advanced forms of click fraud by developing the next generation of online filters.

In summary, I have been asked to evaluate Google's invalid click detection efforts and to conclude whether these efforts are reasonable or not. Based on my evaluation, I conclude that Google's efforts to combat click fraud are reasonable.