Understanding the Network and User-Targeting Properties of Web Advertising Networks

Yong Wang* †, Daniel Burgener†, Aleksandar Kuzmanovic† and Gabriel Maciá-Fernández ‡
*University of Electronic Science and Technology of China, Chengdu, China
†Northwestern University, Evanston, IL, USA
‡University of Granada, Granada, Spain

***Abstract*—Advertising has become an integral and inseparable part of the World Wide Web. However, neither public auditing nor monitoring mechanisms still exist in this emerging area. In this paper, we present our initial efforts on building a network- and content-level auditing service for Web-based ad networks. Our network-level measurements – charting the network infrastructure and quantifying the ad platforms' delay performance – can help commissioners to evaluate their networks from end users' perspective, and let advertisers choose commissioners that better fit their needs. Our content-level measurements – understanding the ad distribution mechanisms and evaluating location-based and behavioral targeting approaches – bring useful auditing information to all entities involved in the online advertising business. We extensively evaluate Google's, AOL's, and Adblade's ad networks and demonstrate how their different design philosophies dominantly affect their performance at both network and content levels.**

I. Introduction

Advertising is thriving on the Web. Indeed, it has become hard to find a popular Web site that does not show banner ads, while rich media advertising is becoming commonplace for many sites. More importantly, advertising has enabled the de-facto business model for today's Web: it provides the main revenue source for many Web sites and services (*i.e.*, *publishers*); it gives an unprecedented opportunity to companies to advertise their products on the Web (*i.e.*, *advertisers*); finally, it brings huge profits (*e.g.*, [1], [2]) to *commissioners* who effectively tie together advertisers and publishers and provide the necessary infrastructure that hosts and serves Web ads.

While Web advertising has been explored from different aspects, (*e.g.*, monetary [3], [4], privacy [5], and E-Commerce [6]), to the best of our knowledge no effort has been invested in understanding how Web ad networks operate in the *network* level (how many data centers or servers are in such networks, what is their geographic distribution, or how effective are they in serving ads?) and *content* domain (which ads are served, where, and when?). Understanding these properties is essential for a number of reasons, some of which we outline below.

At the network level, knowledge about the performance of existing ad networks can help a new commissioner entrant in the market to choose whether to deploy its own network, or use Content Distribution Networks (CDN) services. Such analysis can further help in discovering regions that might show high discrepancies in response time, between the original site's content, served by publishers or associated CDNs, and ads, served by commissioner networks or CDNs, since high discrepancies may be annoying for Web users [7], [8]. Finally, knowledge about commissioners' ad networks is useful for commissioners themselves. For example, to understand how individual (and often *administratively independent*) ad networks within their own domain operate and whether unifying them would improve the overall performance or not.

At the content level, understanding which ads are served at given publisher sites, and how well they match user's interest profile or geographical location, is important for evaluating a commissioner's effectiveness in bringing the right content to the right audience. In addition, revealing the ties that exist between publishers and commissioners is important for establishing the necessary public auditing mechanisms in this domain. Moreover, gaining insights about the prevalence and effectiveness of location-based and behavioral ad targeting applied by various commissioners can provide useful auditing information to end-users and advertisers. Finally, such insights can provide invaluable information to advertisers and publishers when choosing which commissioners to work with.

In this paper, we perform a large-scale measurement and analysis of Web ad targeting networks on the Internet. Our key contribution lies in developing an ad monitoring procedure including a series of measurement metrics and in demonstrating that it is capable of effectively screening ad networks at scale. The methodologies we present in this paper are the rudiments of the network- and content-level auditing service that we intend to design for the Internet.

In our work here we provide the first insights from this effort; hence, we necessarily limit our study to a subset of Web ad networks and explore the questions outlined above. In particular, we originally aimed to evaluate five largest ad networks in the world. However, after determining the network infrastructure they use to distribute ads, we decided to focus on Google, AOL, and Adblade, as representatives of different design philosophies. Google uses its own large-scale distributed private network, and it is in this respect unique among all commissioners. We select AOL as the largest among the commissioners that use CDN services. Finally, Adblade is a commissioner that has a single point of presence on the Internet and hence differs dramatically from the other two networks.

At the network level, we first explore the generic delay performance. We then explore the delay *differences* between commissioners' ad networks and associated publishers' networks, and find that the discrepancy can be quite significant in certain regions. Our network-level experiments are not without surprises. For example, we find that Palo Alto, California, the unofficial capital of the Web advertising industry, has the

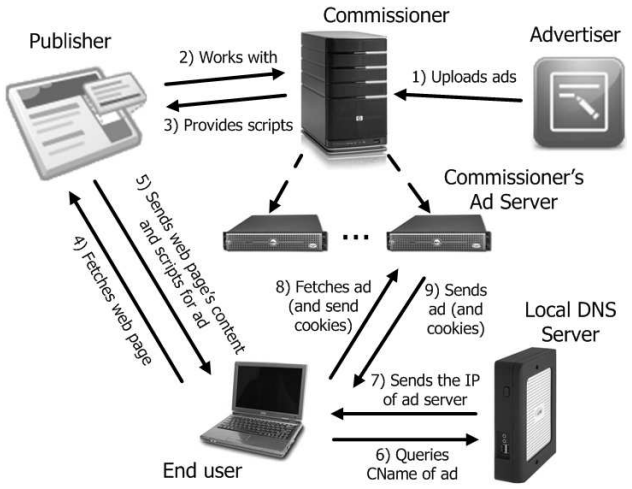

Fig. 1. How Web advertising networks operate

largest delay discrepancy *worldwide* between contents and ads for one of the explored commissioners. A more fundamental issue that we find is that even if Web content and ads are served by the same CDN, there exists no internal mechanism within a CDN to recognize and correct such anomalies.

At the content level, we explore the ad distribution mechanisms applied by different commissioners. Our key finding is that the network-level properties, *i.e.*, the use of CDNs, does influence the content-level ad distribution strategies. In particular, we find that CDNs distribute similar ads in given regions, yet CDN-based commissioners are somewhat behind others in achieving finer-grained location-based advertising. Finally, we explore the extent to which behavioral targeting is used and find that two of the three evaluated commissioners apply such ad delivering techniques.

This paper is structured as follows. Section III summarizes our measurement methodology. Section IV focuses on comprehensively charting the evaluated networks. Sections V and VI explore network- and content-level properties, respectively. We present related work in Section VII and conclude in Section VIII.

II. Background

In this section, we provide an overview of how Web advertising networks operate. Figure 1 shows the details. *Advertisers* in Figure 1 refer to any company or organization who want to advertise their products on the Web. To display advertisements (ads) to a large number of *publishers* in Figure 1 (who have their own Web sites) at scale, an advertiser can send its ads to a *commissioner* (Step 1 in Figure 1). The commissioner in Figure 1 provides the necessary infrastructure, *e.g.*, *commissioner's ad servers* shown in Figure 1, that host and serve Web based ads. Similarly, to attract a large number of advertisers to display their ads on its Web site at scale, a publisher works with a commissioner (Step 2 in Figure 1). Next, the commissioner provides appropriate scripts for this publisher who then vacates a space and embeds the scripts that generate ads for that space in its Web pages (Step 3 in Figure 1).

When an *end user* fetches the publisher's Web page (Step 4 in Figure 1), the publisher sends its Web page content as well as the scripts to the end user (Step 5 in Figure 1). Then, the end user executes the scripts and obtains the specific URLs of ads to be downloaded. Next, the end user queries its *local DNS server*, shown in Figure 1, to resolve the canonical Name (CName) of the URLs of ads into the IP address of an ad server hosting them (Step 6 in Figure 1). The local DNS server may consult ad server's authoritative DNS server to obtain the IP address of a nearby ad server (depending on if it caches such records) and sends the IP address to the end user (Step 7 in Figure 1). Finally, the end user further fetches the ads from the ad server (Step 8 in Figure 1), which in turn sends ads back to the requester (Step 9 in Figure 1).

In Step 9, a significant download delay would dramatically affect user's online experience. We measure such delay performance in Section V. Also, in Step 9, the commissioner could translate the end user's IP address into the user's geographical address. This helps provide location-specific ads that are more relevant to the user. We will analyze such location-based advertising technique in Section VI-B. In step 8, if it is the first time that the end user works with this commissioner, the commissioner usually sends cookies along with the ads. Cookies can help commissioners to label and distinguish Web users. Later, when the end user accesses other ads from the same commissioner, she would send the cookies together with the request for ads to the commissioner. Then the commissioner uses the cookies to identify the end user and provides ads that are relevant to the user's preferences, as determined by the user's browsing pattern. We will analyze such behavioral targeting technique in Section VI-C.

III. Methodology

Here we propose a methodology that allows us to obtain insight into an ad network's infrastructure and to evaluate its performance. As pointed out above, the utility of this methodology is twofold. Firstly, it helps potential new advertisers/publishers in the decision of choosing those commissioners which better meet their requirements. Secondly, it allows commissioners to evaluate their own networks, with the aim of detecting potential design flaws and points of failure with reduced quality of service.

Table I summarizes some of the usual questions that commissioners, advertisers, and publishers have when participating in the online advertising business. Certainly, they would value the ability to perform auditing *tests* to ad networks before making any business decisions. We propose a number of such tests, which can be grouped in several sets, to evaluate issues such as (i) measure network delay, (ii) analyze the ad distribution mechanisms, (iii) detect the use and effectiveness of location-based and (iv) behavioral targeting strategies. In this paper, we design and present such ad auditing tests.

The proposed methodology follows a process which is carried out in three phases. During these phases, several different features of the network operation are measured and subsequently evaluated to extract desired information. We perform the measurements of the considered features using an evaluation platform, which we describe in more detail in

	Delay	Distribution	Location	Behavioral
Commissioners' interests				
What is the appropriate design philosophy for my network, if I am a new entrant in advertising industry?	X	X	X	X
What is the average delay of my network when serving ads? What is the performance of other ad networks?	X	X		
Is the currently used (when applicable) CDN network degrading or improving my ad network performance?	X	X		
Advertisers' interests				
How fast does a commissioner deliver ads to end users?	X			
How well is a commissioner able to distribute ads considering users' location information?			X	
How well is a commissioner able to distribute ads considering users' profiles?	X			X
Publishers' interests				
How does the publication of ads coming from a given commissioner affect user perceived performance when visiting my Web site?	X			
How relevant are the ads shown in my Web sites to my users?			X	X

TABLE I
RELATION OF SOME USUAL ADVERTISERS', PUBLISHERS' AND COMMISSIONERS' QUESTIONS, AND THE AUDITING TESTS SUGGESTED TO OBTAIN USEFUL INFORMATION TO SOLVE THEM. (DELAY: SERVERS' DELAY PERFORMANCE, DISTRIBUTION: AD DISTRIBUTION MECHANISM, LOCATION: LOCATION-BASED ADVERTISING, BEHAVIORAL: BEHAVIORAL TARGETING)

Section III-A. Below, we list and describe the phases of our methodology.

Phase 1. Charting ad networks. In this preliminary phase, our goal is to collect data about the physical and the logical structure of the ad networks that we evaluate. (We provide details about specific networks in Section IV below.) In particular, we are interested in discovering, for a given network, how many content servers there are, how many DNS servers there are, what the geographical locations of the nodes are, what their naming conventions are (CNames of servers) and what their IP addresses are. Moreover, in this phase we also aim to discern whether an ad network is using CDN services to distribute its contents or not. Obviously, the use of CDN networks affects an ad network's performance, and how much is this done is one of the items we explore in the next phase. In Section IV we show how to obtain all this information for three different ad networks.

Phase 2. Network-level performance evaluation. During this phase, we perform an evaluation of ad networks from a networking perspective. The question we attempt to answer is that of how efficiently an ad network is able to serve ads to users. Specifically, we analyze the network delay feature, *i.e.*, locations that experience high discrepancy between ad and original site's content delays. These delay features are useful for commissioners to gain knowledge about poorly-served network regions, *i.e.*, zones in which the geographical distribution of the network nodes is not optimized. Our methodology exclusively focuses on delay measurements. This is not only because it is the most convenient measure that we can get, but also because ads are in general very small in size. Hence, network delay dominantly affects the ad download times [9], [10]. We give details for the execution of this phase in Section V.

Phase 3. Content-level performance evaluation. Ad networks should be designed to maximize the likelihood that users will click on the shown ads. Hence, such networks aim to bring ad content that is 'in harmony' with users' interests. This is usually achieved by selecting ads either by considering the contents of Web pages where they are embedded, *i.e.*, contextual advertising, or by employing other types of information such as users' profiles, navigation patterns, location, etc. In this phase, we perform an evaluation of the strategy that ad networks apply to choose appropriate ads among a set of possible candidates. Firstly, our goal is to discover how an ad network serves ads to users who are geographically dispersed and are accessing similar Web pages. This is useful to understand the policies applied by the ad service. Our second goal is to evaluate if and how an ad network applies location-based advertising. This gives an idea about how the network adapts to geographical dispersion of its users. Finally, we also evaluate the utilization of users' profiles in order to apply a behavioral targeting strategy by the ad network. Both the methodology for this phase and the results obtained when applied to three different ad networks are explained in Section VI.

A. Evaluation Platform

To achieve the above goals, we must select an appropriate measurement platform. First, we strive for a platform that

TABLE II
OPEN RECURSIVE DNS SERVERS

Region	# countries	# servers	% of total
N. America	25	33,645	43.73
Europe	50	26,294	34.18
Asia	40	14,019	18.22
S. America	12	1,405	1.83
Oceania	8	1,111	1.44
Africa	24	456	0.60
Total	159	76,930	100.00

TABLE III
PLANETLAB SERVERS

Region	# countries	# servers	% of total
N. America	3	139	49.29
Europe	22	103	36.53
Asia	7	29	10.28
S. America	3	8	2.84
Oceania	1	3	1.06
Africa	0	0	0.00
Total	36	282	100.00

would have a large geographical distribution to successfully emulate users' locations diversity and effectively chart the given ad networks. Second, the platform should be capable of going beyond delay measurements, *i.e.*, we want to be able to fetch contents from ad networks in order to evaluate their content-level properties. To achieve both of the above goals, we apply two platforms: an open recursive DNS platform (used in Section IV) and PlanetLab (used in Section V and Section VI).

Open recursive DNS. We use the open recursive DNS platform in phase 1 of the methodology (charting ad networks). Open recursive DNS servers are public DNS servers in the Internet which provide DNS resolution service to any requester, without any source-based filtering. Because we are capable of obtaining a large number of such vantage points (details below), we can effectively and accurately chart the ad networks. We achieve this by performing the translation between the CNames of the ad network servers into IP addresses which have values that depend on the requesters' location.

We use two approaches to locate open recursive DNS servers. First, we start from a large list of Azureus clients' IP addresses. We then look up the DNS servers' IPs of these clients and check if they are open recursive DNS servers. A second approach consists of retrieving the one million most popular Web sites from Alexa [11], and finding their authoritative DNS servers. Again, we check if they are open recursive DNS servers. By using these approaches, we successfully locate 76,930 open recursive DNS servers over six continents and 159 countries (Table II).

PlanetLab. The large number of open recursive DNS servers is indeed useful when they are used for mapping ad servers' CNames to IP addresses. However, such a platform is incapable of accurately measuring network delays. Moreover, it is also technically impossible to fetch the contents of advertisements using our open recursive DNS platform. Hence, these reasons drive us to a PlanetLab platform capable of achieving all these goals. Specifically, we recruit 282 servers from PlanetLab, which are geographically distributed in 36 different countries as shown in Table III. Despite the obvious difference in the size between the two platforms, we demonstrate below in Section IV-C that the PlanetLab platform still achieves fairly wide network coverage relative to the open recursive DNS one.

IV. CHARTING AD NETWORKS

In this section, we introduce the methodologies for charting ad content servers and ad DNS servers of each of the commissioners that we evaluate. We aim to address the following questions: *(i)* What is the philosophy that commissioners use to set up their advertising services, *(ii)* how many ad servers exist in each of the networks, and *(iii)* where are they geographically located?

A. Candidates Selection

We originally worked on 5 commissioners: Google, AOL, Yahoo, Microsoft, and Adblade. The first four are among the largest Web advertising companies [12], while we select the last one (Adblade) for the reasons explained below. After determining the network infrastructure used for distributing ads, we decided to study Google, AOL, and Adblade for these reasons.

Google uses its own large-scale distributed private network, and it is in this respect unique among all the commissioners. Moreover, it applies the so-called data-center approach, *i.e.*, distributed data centers exist all around the world, yet in a fairly moderate number of locations. Moreover, the Google infrastructure hosts two independent ad networks, Google's and Doubleclick's (recently acquired by Google). To distinguish between the two, we refer to them as Google-Google and Google-Doubleclick.

AOL, Yahoo, and Microsoft use the Akamai CDN network to distribute ads; hence, they share high similarity at the network level. Moreover, in comparison with Yahoo, which uses one CName with 2,278 different IP addresses, and Microsoft, which uses three CNames with 3,704 different IP addresses, AOL is the largest among them. In particular, AOL has four subsidiaries, Adtech, Adsonar, Advertising, and Tacoda; it has in total 12 CNames with 11,132 different IP addresses. Thus, we choose AOL as a representative of CDN-based advertisers.

Adblade is a commissioner that has a single point of presence, *i.e.*, a single server (or a cluster) located in Jersey City, New Jersey. Hence, its infrastructure fundamentally differs from the above two. We select it to understand what it looks like to serve ads without a distributed infrastructure. Moreover, it is also interesting to know its performance at the content level. Thus, we select it as a representative of other companies that apply the same philosophy.

TABLE IV
NUMBER OF IPS FOR EACH COMMISSIONER

Commissioners	# of IP	
	Ad content servers	Ad DNS servers
Google	306	6
AOL/Akamai	11,132	8,381
Adblade	1	2

TABLE V
GEOGRAPHIC DISTRIBUTION OF EACH COMMISSIONER

Continent	# of IP					
	Google		AOL/Akamai		Adblade	
	Ad	DNS	Ad	DNS	Ad	DNS
N. America	154	3	6,761	5,426	1	2
Europe	70	2	3,017	1,824	0	0
Asia	24	1	994	883	0	0
S. America	14	0	144	91	0	0
Oceania	0	0	178	124	0	0
Africa	0	0	38	33	0	0
Unknown	24	0	0	0	0	0
Total	286	6	11,132	8,381	1	2

B. Finding Canonical Names

In order to study the ad networks, we must first discover the CNames of their ad servers. Each ad company may use many different CNames, and they may vary over Web sites and geographical locations. In order to get a representative picture, we crawl the top 28,268 sites as listed by Alexa in the PlanetLab infrastructure with Firefox enabled, and record the DNS traces using Wireshark [13]. This gives us a list of URLs that were accessed by Firefox, and we search through them for the names of the ad companies that we are interested in. Then we use the `dig` tool to convert this list into CNames. This gives us 7 CNames for Google-Google, 34 for Google-Doubleclick, 49 for AOL/Akamai and its four subsidiaries, and 3 for Adblade.

We count the number of times that aliases for each CName are accessed in our DNS traces. Then, we select the most used CNames as candidate CNames to represent each company. This eliminates any potential error in missing a CName, as we are certain to have found the most popular CNames and we are limiting ourselves to studying the most used ones. We manually visit Web sites that show each of the CNames we select and confirm that an ad is delivered by a server at that name. It has the further benefit of eliminating servers that are not used for serving ads, *e.g.*, recording users' browsing pattern. Finally, for the ad content servers of three commissioners, we choose pagead.l.google.com for Google-Google, pagead.l.doubleclick.net for Google-Doubleclick, a950.g.akamai.net for AOL-Adsonar, 3 CNames for AOL-Adtech, 3 CNames for AOL-Tacoda, 5 CNames for AOL-Advertiser, and web.adblade.com for Adblade. [1] We do not include the details about the CNames of ad DNS servers due to space constraints and the fact that they can be easily obtained from the CNames of ad servers. We observe (see Section IV-C) that multiple CNames owned by the same company frequently map to the same IPs, so selecting the most used CName will give us a very representative sample of the company's ad network.

C. Mapping CNames to IP Addresses

To completely chart the ad networks for each commissioner, we query each of the selected ad servers' CNames from all the open recursive DNS servers and Planetlab servers. In many cases, the same CName is mapped into different IP addresses when queried over all vantage points. Table IV summarizes the number of IPs of both ad content servers and ad authoritative DNS servers for ad content servers of each commissioner in open recursive DNS platform, which covers all the discovered servers in the Planetlab platform. Comparing with Google which has 306 [2] ad content servers and 6 ad DNS servers, and AOL/Akamai which has 11,132 ad servers and 8,381 ad DNS servers, Adblade only has 1 ad content server and 2 ad DNS servers. Driven by the significant difference among commissioners, we explore the delay performance of each commissioner in Section V-A.

The difference of the discovery capacity between two platforms. As we mentioned above, the difference of the number of discovered ad servers between two platforms is not dramatic. For example, the DNS platform is able to discover 306 Google servers (153 servers for each of Google-Google and Google-Doubleclick), while the PlanetLab platform finds 286 servers (143 servers for each of Google-Google and Google-Doubleclick), which cover 93.5% of servers the open recursive DNS platform discovers.

Examining the IP Overlap among CNames. The CNames hosted by the same commissioner are usually mapped to the same IP addresses. Taking Google-Google as an example, we discover 153 ad servers for the CNames of pagead.l.google.com (set A), afd.l.google.com (set B), and partnerad.l.google.com (set C). We also examine the coverage relations of IP addresses among all sets, and find that sets A, B, and C are exactly identical. This is why it is sufficient to choose a representative CName for each commissioner, *e.g.*, pagead.l.google.com for Google-Google, to conduct our further experiments, as we introduced at Section IV-B.

D. Mapping IP Addresses to Locations

Using three different geolocation databases [14], [15], [16], we map the total of 11,132 ad content servers and 8,381 ad DNS servers into 77 different countries at 6 continents in the AOL/Akamai case. More than 60% of the ad content servers and the ad DNS servers are located in North America and more than 21% are located in Europe (Table V). In Adblade's case, we map a single ad content server and two ad DNS servers

[1] a627.g.akamai.net, a973.g.akamai.net, e1611.c.akamaiedge.net for AOL-Adtech, a1131.g.akamai.net, a1406.g.akamai.net, e922.p.akamaiedge.net for AOL-Tacoda, a949.g.akamai.net, a957.g.akamai, a1539.g.akamai.net, a1626.g.akamai.net, e1066.c.akamaiedge.net for AOL-Advertiser.

[2] Note that there may be a cluster of hundreds of machines behind each IP.

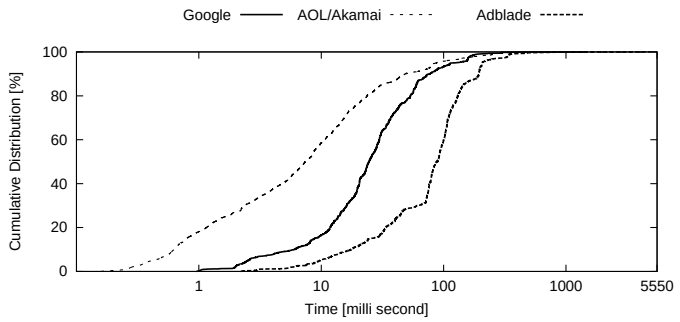

Fig. 2. Cumulative distribution of the delay for ad content servers.

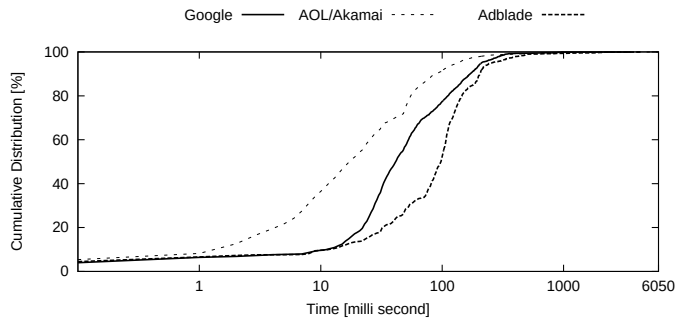

Fig. 3. Cumulative distribution of the delay for ad DNS servers.

to Jersey City, New Jersey. We fail to resolve the correct locations for Google's IP addresses that are all mapped to California, U.S., by means of all three geolocation databases. This is because Google registers all its IP addresses to its head company's address. Thus, we utilize a constraint-based geolocation approach [17] to discover the location of Google's IPs at the continent level.

V. Network-level performance

Here, we first show a comparative study on the delay performance for the selected commissioners. This information gives us the idea about their overall delay performance when serving ads. Next, we compare the delay performance between each commissioner's network and some of its corresponding publishers' networks. This result enables us to focus on performance discrepancies between these two networks that might be annoying for Web users. Moreover, such information might help commissioners to re-evaluate their server-placement strategies.

A. Delay Performance

Here, we evaluate and compare the delay performance for each of the three selected commissioner networks. For each one, we measure the delay between its ad content servers and all 282 Planetlab servers by directly sending `PING` probes for the CNames from all Planetlab servers twice a day, for a six-day long period. The results for the delay distributions of content servers are shown in Figure 2. Using common industry standards [18], we focus on the 95th percentile of the delay distribution (Table VI). Considering this metric, AOL/Akamai (87 ms) outperforms Google (122 ms), which is in turn better than Adblade (207 ms). Such trends also hold in Figure 2. This data bear out the expected result that AOL/Akamai's ad content servers are closer to end users than Google's and Adblade's. Given that AOL relies on the Akamai CDN network, which distributes a large number of servers around the world, while Google deploys its clusters of servers only at some big cities, it makes sense that AOL/Akamai performs better than Google. Finally, as Adblade puts its cluster of servers at a single location, it presents the worst performance.

We also measure the delay between ad DNS servers and all vantage points by using the KING approach [19]. In particular, for every vantage point, we first send a probe query, *e.g.*, a request for pagead.l.google.com, to seed the vantage point's local DNS server so that it do not have to contact the root level DNS servers in the near future. We then send another query to resolve an artificial domain that has a random number and the same subdomain name as the previous query, e.g., a request for random.l.google.com, to measure the delay between the vantage point and the ad DNS servers.

We present the cumulative distribution of the delay for ad DNS servers in Figure 3. Results considering the 95th percentile of the delay distribution (Table VI) show that the ad DNS servers are not as close to end-users as ad content servers are, which are in-line with the trends shown in Figure 3. This is a natural consequence of the discrepancy in the number of DNS and content servers in each network, *i.e.*, the number of DNS servers is smaller, except for Adblade. In this latter case, although the geographical location for Adblade's DNS and content servers is the same, there is a difference in the delay, due to routing effects ***etc.***

B. Ad vs. Publisher Networks

Here, we measure the discrepancy between publishers' and commissioners' response delays. Given that both publishers and commissioners may or may not work with CDNs, this discrepancy can be quite substantial. Certainly, such a substantial discrepancy can be annoying for Web users. For example, the appearance of a Web page might need the complete download of all elements in that Web page, depending on the browser and Web site setup. Late ad appearance might degrade the entire Web page rendering process [7], [8]. For these reasons, a publisher is necessarily interested in assessing how much the publication of ads coming from a given commissioner affects its Web site performance.

Since the aggregated delay information given in previous section may not be useful in this experiment, we choose

TABLE VI
95th percentile results for the delay distributions of content and DNS servers.

Commissioners	Ad Delay (ms)	
	Content servers	DNS servers
Google	122	210
AOL/Akamai	87	136
Adblade	207	255

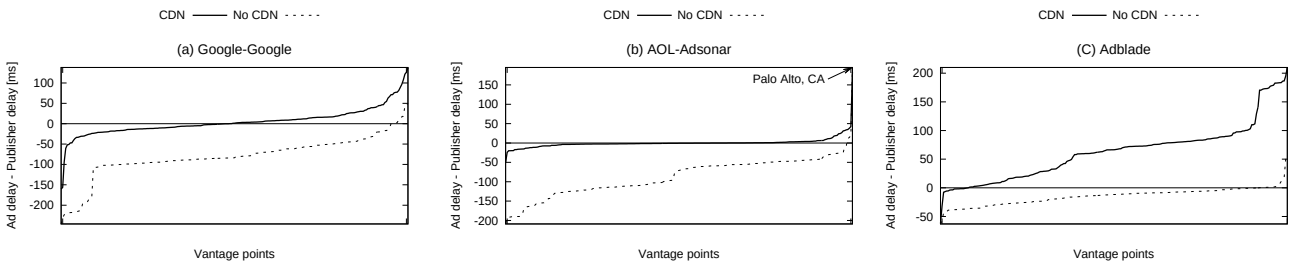

Fig. 4. Ad content network delay V.S. publisher content network delay

the Google-Google network as the representative for Google and AOL-Adsonar network for AOL/Akamai, and deploy additional experiments (details below). We select these two ad networks for two reasons: (i) Since each subsidiary company within a same commissioner uses the same network, the difference among them are not significant. (ii) Google-Google and AOL-Adsonar are also the two networks that support text-based ads. Such ads enable us to evaluate these commissioners at the content-level as well (Section VI). Hence, this approach enables us to comprehensively and comparatively understand these networks' network- and content-levels properties in experiments described in next sections.

To conduct this experiment, we proceed as follows. For each of 282 vantage points, we *simultaneously* send `PING` probes to the commissioner network and to the publishers' network containing over 600 Web sites that this commissioner works with, roughly half working with CDN, and half not working with CDN. We then measure the delay *difference* between the two networks and report averages. Since publishers may use a CDN network to distribute their contents, we further divide publishers into two categories: (i) those that work with CDNs (marked as 'CDN' in Figure 4) and (ii) those that do not (marked as 'no CDN' in Figure 4).

Figure 4 shows the results for the three selected commissioners. Whenever a value is below the y=0 line, the ad network outperforms the publishers' network at that vantage point. When a value is above the y=0 line, ads are delivered later than the Web content. The figure clearly shows that the largest discrepancy happens when one of the networks relies on a CDN, while the other does not. For example, when publishers are not using CDNs, Google-Google and AOL-Adsonar networks significantly outperform the publishers' network performance. On the contrary, in Adblade's case, when publishers are using CDNs, while the ad network is not, the ad delay is much higher.

More specifically, results in Figure 4 demonstrate that if publishers do not utilize a CDN network, commissioners' ad networks are typically better than publishers'. In particular, in only 1.76% of the vantage points for AOL-Adsonar case, 5.3% for Google-Google case, and 8.4% for Adblade case, do commissioners' ad networks perform worse than publishers'. However, when publishers are served by a CDN network, in 38.33% of the vantage points for the AOL-Adsonar network, 51.1% for the Google-Google network, and 91.6% for the Adblade network, commissioner's networks are worse than publishers'.

Next, we try to understand which regions commissioners underperform in. As pointed out above, answering this question is important since a huge discrepancy is typically annoying for Web users [7], [8]. Here, we evaluate differences higher than 100 ms. Surprisingly, we find that AOL-Adsonar present differences of 182.91 ms (publishers working with CDN case) and 194.14 ms (publishers not working with CDN case) in Palo Alto, California. We find it somewhat ironic that Palo Alto, the unofficial capital of the web advertising industry, is at the bottom of the worldwide list. For some reason, Akamai hosts ads for the Palo Alto area from Elmwood Park, New Jersey and Englewood, Colorado. A more fundamental issue here is that even if Web content and ads are served by the same CDN (Akamai in this case), there exists no internal mechanism within a CDN to recognize and correct such anomalies.

In the Google-Google case, there are almost no locations where publishers which do not work with a CDN exceed Google-Google over our threshold of 100 ms. Yet three countries, *i.e.*, South Korea, Japan, and Brazil, have discrepancies of 137.21 ms, 121.68 ms, and 107.54 ms. Although Google has deployed clusters of servers in Hong Kong, Taiwan, and the south portion of Brazil (found at Section IV-D), it would be also necessary for Google to rethink its network deployment in Korea and Japan, and other regions in Brazil to better accommodate its network.

VI. Content-level performance

In this section, we focus on the measurement and analysis of the three commissioners at the content level. Understanding the content-level performance is vital for all entities of the online advertising business for several reasons. As we discussed above, advertisers may not only care about network-level delay performance when choosing commissioners. They may also think about the way that their ads are distributed. We will comprehensively explore this issue in Section VI-A.

In addition to the ad distribution method, the ability to target local customers around their business is an important factor for advertisers. Therefore, questions such as, whether a commissioner utilizes location-tracking technology [20] and how much a commissioner uses it, are also on advertisers' minds when they choose partners, because such technology could significantly raise the click through rate (CTR) [21]. We will quantify such behavior in Section VI-B.

Adopting behavioral targeting [4] is a relatively novel technology that increases the effectiveness of targeting potential customers. Advertisers certainly prefer cooperating with commissioners that can support such technology, which could

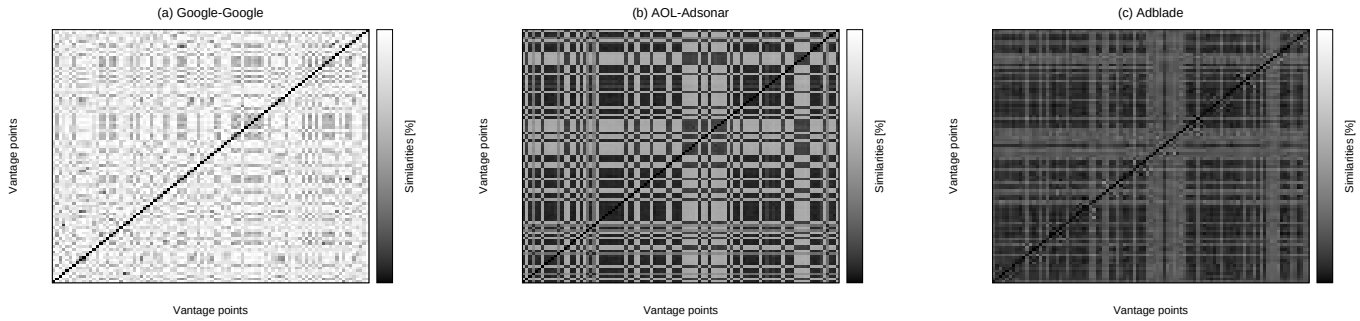

Fig. 5. Local similarities among vantage points (Dark=high similarity, light=low similarity)

also dramatically increase the click through rate. We analyze this issue in detail in Section VI-C below.

We choose Google-Google as the representative for Google, AOL-Adsonar for AOL, and Adblade here, since they support *text-based* ads that can be feasibly retrieved from the Web. More importantly, it is technically easier and more accurate to extract the contents *e.g.*, ad location information, from text-based ads than rich-media-based ones, as the parsing of the latter might leads to errors and, thus, to inaccurate results.

A. Distribution Mechanisms

To analyze ads distribution mechanisms, we now fetch the ads from the publishers' networks containing over 600 Web sites that we used in Section V-B. We accomplish this task from all Planetlab servers once a day in a consecutive period of five days. We first disable cookies at our experimental Planetlab machines in order to avoid behavioral targeting issues that we explore later. Then, for each vantage point, we calculate a *local similarity* metric between itself and any other vantage point in terms of the percentage of *identical* ads observed in both vantage points. In addition, we also compute a *global similarity* metric for each commissioner by averaging the 'local' similarities for all pairs of vantage points when retrieving this commissioner's ads.

Table VII shows that Adblade's global similarity is higher than that of AOL-Adsonar, which is in turn higher than Google-Google. A cause for such a sequence is that Adblade uses a single machine (or a cluster) to serve the requests; as this machine serves the whole pool of ads, it is expected that, after enough requests to publisher sites, all destinations (vantage points) receive all ads independently of their location (high similarity). AOL-Adsonar uses the Akamai CDN network to distribute ads. Although this would allow them to use different pools of ads depending on the location of the servers, as the essence of a CDN network is to share the same content over CDN servers, all vantage points still share a relatively high similarity. (We later prove that, in the advertising case, the servers in the same region indeed serve the same content). In the Google-Google case, as we discussed above, Google-Google has its own private network and, hence, the full control on the methods for distributing ads. Consequently, different vantage points that fetch ads from different ad servers experience a low similarity.

TABLE VII
GLOBAL SIMILARITY FOR EACH SELECTED COMMISSIONER.

Commissioner	Global similarity (%)
Google-Google	13.16
AOL-Adsonar	59.31
Adblade	72.62

Figure 5 shows the 'local' similarity among all pairs of vantage points. For each vantage pointed on the x-axis, a vertical 'stripe' corresponding to a value on the x-axis shows the similarities between this vantage point and others. The darker the color in a given (x,y) box is, the larger the similarity between x and y is. Coherently with the results shown in Table VII, the figure shows that the 'local' similarities in the Google-Google case are relatively low. This implies that Google-Google has a large pool of ads and distributes different ads into different servers. On the other extreme, the similarities in the Adblade case are quite high. This is because Adblade has a smaller pool of ads and puts all of them in the same machine (or a cluster of machines). Some relatively light stripes in the figure come from the fact that Adblade uses location-based advertising in the U.S., as we explain in Section VI-B below.

We explore the highly structured nature of Figure 5(b) (the AOL-Adsonar case), where a vantage point has either high or low similarities with others. To explore this issue, we cluster our vantage points based on their similarities in the following four regions: U.S., Canada, U.K., and 'Others'. We then explore intra- and inter-cluster similarities in Figure 6. The results show us that the intra-cluster similarities are much larger than the inter-cluster ones. For example, the similarity percentages among vantage points within U.S. range from

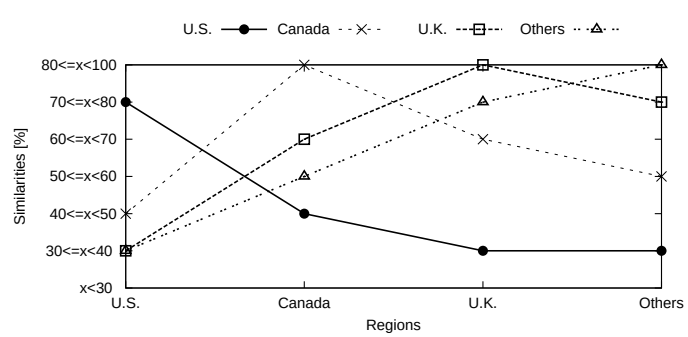

Fig. 6. Regional similarities in AOL-Adsonar

TABLE VIII
PERCENTAGE OF VANTAGE POINTS OBSERVING LOCATION-BASED ADS CONTAINING CITY OR STATE RELATED INFORMATION, OR WITHOUT LOCATION INFORMATION.

Commissioners	City	State	No info
Google-Google	31.58%	21.93%	46.49%
AOL-Adsonar	8.00%	12.00%	80.00%
Adblade	37.31%	0.00%	62.69%

TABLE IX
PERCENTAGE INCREASE OF OBSERVED 'SPORT' RELATED ADS WHEN BEHAVIORAL TARGETING IS ENABLED ('*local/uniform cookie*') COMPARED WITH DISABLED ('*no cookie*')

	Local cookie	Uniform cookie
Google-Google	25%	3%
AOL-Adsonar	13%	5%
Adblade	0%	0%

70% to 80%, while the inter-cluster similarity percentages between U.S. and other regions are about 30% to 50%. This phenomenon suggests that AOL/Akamai deploys location-based services, *i.e.*, it is more likely to put similar ads on the servers in a given country. Certainly, such a distribution is useful for targeting users at different markets. Moreover, except for the U.S., all other regions reach their peak of over 80% in their own regions. The reason why U.S. does not peak at over 80% is that AOL-Adsonar uses finer-grained location-based advertising, *e.g.*, city-level advertising, in the U.S., as we explain in Section VI-B below.

B. Location-Based Advertising

Commissioners could dynamically generate and send ads containing information about the location at which end users are. These 'localized' ads are more likely to target potential customers. Here, we quantify the percentage of vantage points in which the use of location-based advertising is observed (Table VIII). In particular, we parse the texts of received ads of each vantage point, and add one to the 'city' column if we could observe a match at a *vantage point* at city-level, or the 'state' column if the match is at the precision of state-level, or the 'no info' column if no association exists between the location of a vantage point and the texts of its received ads. Then, we compute the percentage of vantage points at which location-based advertising is observed.

Table VIII shows that the coverage of location-based advertising in the Google-Google case (31.38% + 21.93% = 53.31%) is wider than in the Adblade case (37.31%), which is in turn larger than in the AOL-Adsonar case (8.00% + 12.00% = 20.00%). This data makes sense because Adblade and AOL-Adsonar only apply location-based advertising within U.S. This is also the reason why the local similarities obtained between the vantage points in U.S. and the ones outside U.S. are to some extent lower than the similarities between the vantage points within U.S. (Figure 5). Also, this explains the fact that the local similarities of vantage points within U.S. do not achieve a threshold of 80% in Figure 6. In Google-Google case, since Google-Google deploys their advertising business all over the world, exploiting the same location-based technology is quite feasible.

C. Prevalence of Behavioral Targeting

Many commissioners claim to be able to more effectively reach users with behaviorally targeted ads. We want to examine the extent to which commissioners participate in behavioral targeting. Generally, commissioners track users by storing a cookie on their computer containing a unique identifier number, and then associating that number with various interest categories. We decide to use the interest category 'sports' in our tests of behavioral targeting because many Web sites in this category work for each commissioner. We first disable cookies in our PlanetLab servers in order to prevent behavioral targeting, and then retrieve the text-based ads from the previous list of Web sites (Section V-B), which may or may not be related to sports. These ads are processed and sports-related keywords are searched. We use about 30 keywords here to classify the 'sports' category , *e.g.*, sport, cycling *etc*.

After establishing this baseline ("no cookie"), we then visit Web sites known to work with the commissioner that fit in the category 'sports' with cookie enabled. After using these Web sites to establish a browsing pattern, we then repeat the above experiment, *e.g.*, retrieve the text-based ads from the previous list of Web sites, in order to observe the difference when behavioral targeting is used. If a commissioner uses behavioral targeting, then we should obtain a higher occurrence of sports related ads after establishing our interest in sports. This is referred to in Table IX as "Local cookie", as the cookie is established locally for each computer.

We finally repeat this experiment by browsing the same path from a local computer, copying the cookies from that computer to all PlanetLab nodes, and then retrieving the ads again. The purpose of this experiment is to give us an understanding of whether user profile information is geographically distributed or merely stored on the closest ad server to a user. This is referred to in Table IX as "Uniform cookie", as the same cookie is distributed to all computers.

Table IX shows the percent increase of 'sports' related ads over the experiment without cookies. Our results show that both Google-Google and AOL-Adsonar use behavioral targeting for the 'sports' interest category, whereas Adblade does not. Google-Google shows a 25% increase when cookies are enabled, and AOL-Adsonar shows a 13% increase.

The increases when a uniform cookie was distributed are fairly negligible (3% and 5% for Google-Google and AOL-Adsonar respectively). Apparently both Google-Google and AOL-Adsonar associate a user profile with interest categories only in a local machine, as the uniform cookie case shows very little increase over the situation without any behavioral targeting at all. This is not a problem for targeting the large number of users who do most of their browsing from a single location. However, users that browse from the same computer while travelling will not be given targeted advertisements while away from home.

VII. Related work

Real-world distributed service platforms have been evaluated previously *e.g.*, [22], [23], [18], [24]. In this context, our approach, *i.e.*, using vantage points to test the performance of the network access, has been applied before. To cite some of the examples, Gummadi *et al.* evaluated the performance of the Napster and Gnutella P2P networks [23]; Mao *et al.* conducted a measurement study to quantify the proximity of web clients to local DNS servers [24]. In a recent study, Huang *et al.* [18] performed a large-scale measurement study of popular CDNs. Our methodology, while similar in spirit to theirs, is different. In particular, at the network level, we avoid DNS-based delay measurements since they can incur a non-negligible error. Moreover, we go beyond only networking measurements and characteristics and evaluate service features as well. In particular, we explore ads similarity, the distribution mechanisms, as well as the prevalence of location-based and behavioral ad targeting policies. Such insights allow us to understand the inter-play between content-level properties and underlying networking design issues.

Another thread of research focuses on the evaluation of the advertising service, without considering the performance of the underlying network. In this context, Xu *et al.* are interested in quantifying the perception of users when location based advertising is used [20]. Others are interested in the evaluation of behavioral targeting strategies and their impact on the end users [4], [25].

For the privacy protection in the ad domain, Guha *et al.* [5] introduced Privad, a practical private online advertising system. In this system, the user profile information and ads are stored locally at user side. When serving an ad, the system selects one from the local pool of ads, rather than a distant ad server, based on user profile. In this way, the system protects user's privacy while enabling behavioral targeting, thus increasing profits, and reducing network delay overhead. As we discussed above, our goal differs from this work in that we build a network- and content-level auditing service for Web based ad networks. To the best of our knowledge, no previous work exists in this area.

VIII. Conclusions

In this paper, we deployed a Web ad auditing methodology that can be universally applied to arbitrary commissioners' networks to effectively monitor and help regulate the Web-based ad industry. Using this methodology, we performed an extensive network- and content-level analysis of three representative commissioner networks with divergent design philosophies. These range from distributing a large number of data centers (Google), to using CDN services (AOL), to standing up servers at a single location (Adblade).

Our findings are the following: (i) Both distributed architectures, namely Google and AOL/Akamai, manage to effectively bring ad content closer to the end users than Adblade. (ii) In three commissioners' ad networks, DNS servers are not as close to end users as ad content servers are. (iii) The discrepancy between publishers' and commissioners' ad networks can be quite high. Such scenarios do not happen only due to divergent ad networks architectures and the use of CDNs; we found that such problems can arise even when both content and ads are served by the same CDN due to the lack of coordination between publishers, commissioners, and CDNs. (iv) At the content level, we found that the explored commissioners deploy location-based and behavioral ad targeting at various levels of granularity. (v) Our results imply that CDN-based commissioners manage to effectively replicate ad content at regional levels, yet lag behind others in achieving finer-grained location-based advertising. (vi) On the contrary, data-center oriented commissioners are capable of collecting user profiles and applying behavioral targeting more effectively.

References

[1] "Google attracting 25% of online ad revenue," http://www.searchenginejournal.com/google-attracting-25-of-online-ad-revenue/3910/.

[2] "Washingtonpost.com: Every click you make," http://www.washingtonpost.com/wp-dyn/content/article/2008/04/03/AR2008040304052.html.

[3] G. Aggarwal, S. Muthukrishnan, D. Pal, and M. Pal, "General auction mechanism for search advertising," in *WWW '09*.

[4] J. Yan, N. Liu, G. Wang, W. Zhang, Y. Jiang, and Z. Chen, "How much can behavioral targeting help online advertising?" in *WWW '09*.

[5] S. Guha, A. Reznichenko, K. Tang, H. Haddadi, and P. Francis, "Serving ads from localhost for performance, privacy, and profit," in *ACM HotNets '09*.

[6] A. Metwally, D. Agrawal, and A. E. Abbadi, "Detectives: Detecting coalition hit inflation attacks in advertising networks streams," in *WWW '07*.

[7] "Pages with ads slow to display because of delay accessing googleads.g.doubleclick.net," http://www.google.com/support/forum/p/AdSense/thread?tid=1866d80fb41a6809&hl=en.

[8] "Slow loading web pages," http://discuss.pcmag.com/forums/thread/1004413251.aspx.

[9] T. Anderson, A. Collins, A. Krishnamurthy, and J. Zahorjan, "PCP: Efficient endpoint congestion control," in *NSDI '06*.

[10] N. Cardwell, S. Savage, and T. Anderson, "Modeling TCP latency," in *IEEE INFOCOM '00*.

[11] "Alexa," http://www.alexa.com/.

[12] "The online ad recession continues. is this what a reset looks like?" http://www.techcrunch.com/2009/07/31/the-online-ad-recession-continues-is-this-what-a-reset-looks-like/.

[13] "Wireshark," http://www.wireshark.org/.

[14] "Geo ip tool," http://www.geoiptool.com/.

[15] "Ip2location," http://www.ip2location.com/.

[16] "Maxmind," http://www.maxmind.com/.

[17] B. Gueye, A. Ziviani, M. Crovella, and S. Fdida, "Constraint-based geolocation of internet hosts," *Transactions on Networking*, 2006.

[18] C. Huang, A. Wang, J. Li, and K. W. Ross, "Measuring and evaluating large-scale CDNs," in *Microsoft Technical Report*.

[19] K. Gummadi, S. Saroiu, and S. Gribble, "King: Estimating latency between arbitrary Internet end hosts," in *SIGCOMM IMW '02*.

[20] H. Xu, L. Oh, and H. Teo, "Perceived effectiveness of text vs. multimedia location-based advertising messaging," *Int. J. Mob. Commun.*, vol. 7, no. 2, 2009.

[21] "So many ads, so few clicks," http://www.businessweek.com/magazine/content/07_46/b4058053.htm.

[22] B. Krishnamurthy, E. C. Wills, and Z. Yin, "On the use and performance of content distribution networks," in *IMW '01*.

[23] P. K. Gummadi, S. Saroiu, and S. D. Gribble, "A measurement study of napster and gnutella as examples of peer-to-peer file sharing systems," *SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 1, 2002.

[24] Z. M. Mao, C. D. Cranor, F. Douglis, M. Rabinovich, O. Spatscheck, and J. Wang, "A precise and efficient evaluation of the proximity between web clients and their local DNS servers," in *ATEC '02*.

[25] J. Jaworska and M. Sydow, "Behavioural targeting in on-line advertising: An empirical study," in *WISE '08*.