

Documentation of Gator Advertisements and Targeting

[Benjamin Edelman](#) - [Berkman Center for Internet & Society](#) - [Harvard Law School](#)

[[Background](#) - [Motivation](#) - [Methodology](#) - [Research & Analysis](#) - [Testing](#) - [Suggestions](#)]

Abstract

The Gator Corporation designs software to display advertisements on users' computer screens, triggered in part by the specific web sites users visit. The author has developed an automated method of determining which specific advertisements Gator has associated with which web sites, data that may be helpful to web site operators, policy-makers, and others in assessing Gator's practices. This article offers listings of more than seven thousand specific sites targeted by Gator as well as analysis of the advertisements shown. An interface is also available to let interested Internet users to test Gator's advertisements on their own.

Related Projects

- [Washingtonpost.Newsweek et al. v. The Gator Corporation - Edelman Expert Declaration](#)
- ["Adware": Research, Testing, Suits](#)
- [Other work by the author](#)

Background

[The Gator Corporation](#), a California company founded in 1998, [calls itself](#) "the leader in online marketing" and [reports](#) on its web site that it has more than 30 million users. As Gator explains in its "[Quick Tour](#)" (page since replaced by Gator; [preserved by archive.org](#)) Gator allows an advertiser to show its ad while a customer is viewing any site on the web, and Gator [promises advertisers](#) (page since removed by Gator; [preserved by archive.org](#)) that it can even display their ads when users visit competitors' sites. Gator's [ad formats](#) include [pop-ups](#), [pop-unders](#), and [sliders](#).

[According to Gator](#), its method of displaying advertisements is more effective than others'. Gator says it offers with a clickthrough rate of more than 10%, purportedly more than 35 times higher than average web sites. Gator attributes this success to "deliver[ing] special offers precisely related to the users' interest at the time they are making purchasing decisions or shopping for relevant goods and services."

Gator's "[Rates](#)" page indicates that its campaigns begin at \$25,000 for a promotion of unspecified duration. A [Business 2.0](#) article suggests that Gator's rates are significantly higher than the ordinary cost of placing an advertisement on a web site by contractual agreement with the operator of that web site.

Gator's activities have prompted a number of legal challenges. Facing complaints from the Internet Advertising Bureau as to Gator's activities, then including banner ads that tended to cover web site operators' own banner ads, Gator in 2001 [sued](#) the IAB and subsequently [settled](#). In 2002, Gator was sued by media companies including the New York Times and Washington Post in a case in which the author worked as an expert. (See [declarations and other case documents](#).) Gator has also been sued by LL Bean, Lending Tree, Six Continents Hotels, and UPS, among others. Details are in the [pending suits](#) section of the author's related "[Adware": Research, Testing, and Suits](#) index.

Motivation

[Return to top](#)

Web site operators often wish to learn whether their sites are targeted by Gator advertisements. In the past, answering this question required installing the Gator client on a computer and checking whether Gator displayed ads subsequent to repeated requests for the specified web sites. But this approach suffers from at least three distinct deficiencies:

1. Manual testing of a site using the Gator client software requires obtaining and installing Gator software.

Users may consider it undesirable to install this software due to concerns about security, performance, and interruption during their web browsing. Some users may be unable to install this software on their computers due to organizational policies prohibiting the installation of unauthorized software generally or Gator specifically.

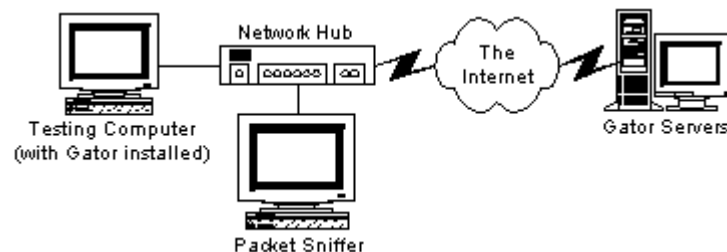
2. Experience shows that Gator imposes a delay between ads -- tending not to show one popup if another has been shown only a few minutes before. Manual testing therefore typically requires an extended period of browsing a web site -- at least a few minutes and potentially considerably longer.
3. Experience shows that Gator targets its ads at specific directories, files, keywords, or transactions. As described in #2, Gator might simultaneously have in its inventory numerous ads to be shown under various permutations of these circumstances and under various patterns of delay or other factors unknown to the web site administrator conducting tests. Gator's focused targeting of particular browsing circumstances further compounds the extended period of browsing required by #2.

With these deficiencies in mind, the author set out to devise a more efficient way of determining whether Gator had targeted its ads at a given web site and, if so, what ads were targeted at what portions of the site.

Methodology

[Return to top](#)

The author posited that the Gator client likely communicates frequently with Gator servers in order to report to the server which web sites a user is viewing and in order to learn from the server which advertisements to display in response. To determine the specific means of communication between the Gator client and Gator servers, the author installed a packet sniffer to monitor all communications between the Gator client and Gator servers, as shown in the diagram to the right. (According to [Telecom Glossary](#), a [packet sniffer](#) is software that enabled "monitoring network traffic in order to recognize and decode certain packets [data transfers] of interest.")



The author observed that the Gator client sends standard HTTP requests (both GETs and POSTs) to Gator servers and receives standard HTTP responses in return.

The HTTP responses received from Gator servers fall into at least three distinct categories:

1. Some responses from Gator servers contain plain text URL references to specific advertisements stored on Gator servers. The advertisements in turn fall into two distinct categories:
 - a. Some advertisements are graphical, taking form of one or more .GIF or .JPEG graphics, sometimes along with an HTML file that frames the graphics in a web page, all compressed in a ZIP file.
 - b. Other advertisements take the form of web pages hosted on external public URLs. These URLs are also expressed in plain text within the advertisement files referenced in #1.
2. Some responses from Gator servers contain lists of keywords that often bear clear and recognizable relationships with the content available on the requested web site. For example, in response to a request for microsoft.com, Gator servers send the client a file that includes the plain text "scid=KB;en-us;q49500" -- recognizable, to experienced users of Microsoft's web site, as a reference to a particular technical article on Microsoft's site. Casual experience indicates that when users visit web pages (on the requested web site) that contain these keyword strings, users are particularly likely to receive Gator advertisements. The keywords seem to include host names (e.g. *support.microsoft.com*), directory names (e.g. */netherlands*), filenames (e.g. *bp_faq.html*), and terms to be sought in URL parameters (such as search terms or other requests to a web server script).
3. Some responses from Gator servers contain references, again in plain text and via ordinary HTTP links, to .grp files that seem to refer to "groups" of sites. A referenced .grp file typically contains text that seems to be the title of the group (e.g. *Electronics Popunder Group*) as well as associated

commentary as to the purpose or significance of the group (e.g. *Group to add a 3 hour delay between all tag-alongs*) and/or special rules pertaining to it (e.g. *do not add new advertiser without having a popunder exclusion added for expedia*).

These methods of data collection do not speak to the frequency with which any given advertisement is shown on a targeted web site. However, empirical testing confirms that most ads found via these methods are in fact shown frequently -- more often than not, by and large -- on computers with Gator installed that visit the corresponding portion of the targeted sites and that have not recently visited related sites that trigger the same or similar advertisements. (The latter restriction is as described in the text of the first example rule in category 3, immediately above.) Experience therefore suggests that when a Gator response in category 1 (above) references an advertisement, that advertisement is likely to be shown to a Gator user who subsequently visits the targeted portion of the site at issue. Future research will seek to determine the method of prioritization among competing advertisements, the required times (if any) between advertisements, and the other specific conditions that trigger advertisements or that cause their suppression.

In the course of determining the specific methods of communication used by Gator, the author reviewed a large number of specific configuration files transferred via ordinary HTTP from Gator servers to the Gator client. However, all data interpreted from these files was found in plain text, without encryption of any kind.

This research is a work in progress. Gator configuration files seem to include extensive additional information beyond that reported here, and the author continues to work to better understand the full contents of these files. Send suggestions for specific research deliverables to [Ben Edelman](#).

Research & Analysis

[Return to top](#)

This section provides a sampling of web sites found to be targeted by Gator, as well as analysis of the specific portions of these sites targeted and the specific advertisements shown.

The following links detail 7,472 specific web sites targeted by Gator for advertisement display:

Alphabetical listing (by second-level domain):

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#) [numbers](#)

Sites by top-level domain:

[.ORG sites](#), [.EDU sites](#), [.GOV sites](#), [.MIL sites](#), [.INT sites](#)

Gator seems to particularly intensively target certain categories of web sites:

<i>Type of site / source</i>	<i>% of sites targeted by at least one Gator advertisement</i>	<i>Average number of Gator ads targeted at each targeted site in category</i>
Sample targeted sites		
Credit Cards Sites Source: Google search on 'credit cards'	65.5%	12.7
Targeted sites include americanexpress.com (43 ads), ftc.gov (2), bankrate.com (36), discovercard.com (2), hypermart.net (2)		
Movies Sites Source: Google search on 'movies'	64.5%	34.5
Targeted sites include imdb.com (14 ads), go.com (35), digitalcity.com (18), hollywood.com (17), yahoo.com (309)		
California Sites Source: Google search on 'California'	60.2%	8.8
Targeted sites include ca.gov (3 ads), berkeley.edu (5), usc.edu (2), caltech.edu (2), ucDavis.edu (4)		
Jobs Sites Source: Google search on 'jobs'	57.9%	12.2
Targeted sites include monster.com (19 ads), jobs.com (6), hotjobs.com (11), opm.gov (7), jobs.net (5)		
Refinancing Sites Source: Google search on 'refinancing'	56.5%	19.6
Targeted sites include bankrate.com (36 ads), cnn.com (27), householdauto.com (5), forbes.com (8), homestore.com (24)		

[Full table available, detailing 115 additional search terms](#)

Note that all analysis operates at the level of second-level domain name. URL listings were stripped to second-level domain names, then data was retrieved as to Gator targeting of each second-level domain name. In each category, the second column reports the proportion of sites found to be targeted by at least one Gator advertisement. The third column reports the average number of Gator advertisements targeted at those sites that are targeted by at least one Gator advertisement.

The following link provides a partial listing of advertisers whose products or services are promoted using Gator:

[Gator advertisers](#)

Testing indicates that, in addition to advertisements targeted at specific web sites, Gator also shows certain additional advertisements subsequent to user requests for numerous web sites. (Gator seems to prevent such advertisements from being displayed subsequent to user requests for certain sites, including Gator's own site and sites specified in court orders). As of May 2003, an additional 17 advertisements were placed in this category of general advertisements, including advertisements about replacement ink cartridges, cell phones, lotteries, business cards, and DVD rental. Using the testing system below, these ads can be viewed by specifying the string *notrgs* as the domain name to check.

Targeting to user profiles. The research described in this document entails replication of communication between Gator client and server using a fresh installation of Gator that has not been used to visit sites other than those necessary to conduct this testing. Whenever the Gator client checks with the Gator server as to advertisements to be shown (subsequent to user requests for a given site), the client sends the server a unique user ID and machine ID as well as the user's IP address, time zone, and zip code. Based on this data, and the associated user profiling that it makes possible, Gator servers could customize the advertisements to be shown to each user. Testing to date has shown no signs of profile-based targeting of this sort, and tests with different or invalid user IDs and machine IDs yield the same results as tests with the original IDs as set by Gator's installation software. Nonetheless, Gator's [marketing materials](#) prominently offer a targeting capability ("... based on past behavior"). This "past behavior" targeting may reflect profile-based targeting not yet detected by the author. Alternatively, it may refer only to display of an ad as a user leaves a targeted web site (or some number of minutes or clicks after a user leaves that site). Future research may attempt to more fully flesh out Gator's profiling systems by comparing the ads shown to users with divergent browsing histories.

Testing

[Return to top](#)

This section allows interested users to advertisement targeting by Gator. Simply enter the domain name to check into the specified field and press Submit. The subsequent page reports 1) the specific graphical advertisements and specific web page popups and popunders shown by Gator in response to requests for the specified domain name; 2) the specific host names, directory names, filenames, or URL keywords targeted by Gator for the display of particular targeted advertisements; 3) the specific category groupings in which Gator places that domain name.

Update: At 7:21pm on May 22, Gator began blocking tests made through this system. Traceroute indicates that the block takes place at the level of a router operated by [Level 3 Communications](#), Gator's ISP. Requests from other hosts (other than the host used for this testing system) are unaffected. For the moment, for lack of ability to obtain new results, the testing system reports cached results obtained in testing of April-May 2003. The author is currently evaluating additional possible modifications; send [email](#) with questions or suggestions.

Name:	<input type="text"/>
Email address:	<input type="text"/>
Special interests / comments:	<input type="text"/>
	<input checked="" type="checkbox"/> Receive updates about this project
	<input type="checkbox"/> Receive information about the author's other projects
Domain name to check: (e.g. <i>harvard.edu</i>)	<input type="text"/> (required)
	<input type="submit" value="Submit"/>

This testing system provides up to five graphical advertisements for each requested site. Additional data is available upon request from the author. Available data typically includes other advertisements, other targeted URL fragments, other group memberships, and more information about each advertisement -- including which keystings are associated with each advertisement, advertisement style (as among Gator's [ad formats](#) of [pop-ups](#), [pop-unders](#), and [sliders](#)), the date on which each advertisement began to appear, and specified advertisement size on screen.

This testing system retrieves data directly from Gator configuration files, as understood by the author according to the [research methodology](#) described above. The author has performed extensive tests to confirm that his understanding of configuration files correctly maps to outcomes as experienced by ordinary users with Gator software installed. However, before taking action in reliance on these results, the author recommends conducting additional hands-on tests as well as manually examining Gator's configuration files as to the sites of interest.

All uses of this system are logged for future study, analysis, and publication. The results of selected requests may be merged with the author's prior data and reported via the links above.

Review some recent tests below:

heise.de - 5 ads found - 4 graphical, 1 textual - 5/23/2003 3:44:29 PM
vistaprint.com - 2 ads found - 5/23/2003 3:44:28 PM
evite.com - 3 ads found - 1 graphical, 2 textual - 5/23/2003 3:44:20 PM
heise.de - 5 ads found - 4 graphical, 1 textual - 5/23/2003 3:44:08 PM
bluemountain.com - 3 ads found - 1 textual - 5/23/2003 3:44:05 PM
hallmark.com - 4 ads found - 2 graphical, 1 textual - 5/23/2003 3:43:44 PM
americangreetings.com - 17 ads found - 15 graphical - 5/23/2003 3:43:35 PM
heise.de - 5 ads found - 4 graphical, 1 textual - 5/23/2003 3:42:43 PM
juno.com - 14 ads found - 11 graphical, 1 textual - 5/23/2003 3:42:12 PM
heise.de - 5 ads found - 4 graphical, 1 textual - 5/23/2003 3:41:37 PM

Suggestions and Extensions

[Return to top](#)

This research is a work in progress. Gator configuration files seem to include extensive additional information beyond that reported here, and the author continues to work to better understand the full contents of these files. Send suggestions for specific research deliverables to [Ben Edelman](#).

This work is supported by the [Berkman Center for Internet & Society](#) at [Harvard Law School](#). Thumbnails by [ABCDrawHTML](#).

Last Updated: May 23, 2003 - [Sign up for notification of major updates and related work](#).