

Measuring ISP Response to VeriSign Site Finder

Benjamin Edelman
Berkman Center for Internet & Society
Harvard Law School

Motivation

- Why measure generally
- Special implications of disabling Site Finder
 - If find Site Finder disabled in many places
 - If find Site Finder disabled in only a few places

Methodology I:

Surveys / Form Submissions

- Problems
 - Burden on respondents
 - Representativeness
 - Accuracy
- Results - Submissions to date report blocking by
 - AOL
 - Earthlink
 - San Bernardino County School System
 - Time Warner Cable (partial?)
 - ...

Methodology II: Site Finder Log Analysis

- Problem
 - Availability of log files for analysis
 - “[A]s a matter of privacy and operational security we do not provide our log files to any outside party”
 - Tom Galvin, VeriSign spokesman (email)
- Results
 - Unknown

Methodology III:

Inference from selected user clickstreams

- Idea: Get data about selected users / networks accesses to Site Finder content, look for trends.
- Problems
 - Potential privacy concerns
 - Obtaining necessary data
 - Users with non-default nameservers
 - Users who request Site Finder content manually
(other than by mistyping domain names)
 - Works best for large networks

Reference:

Inference from selected user clickstreams

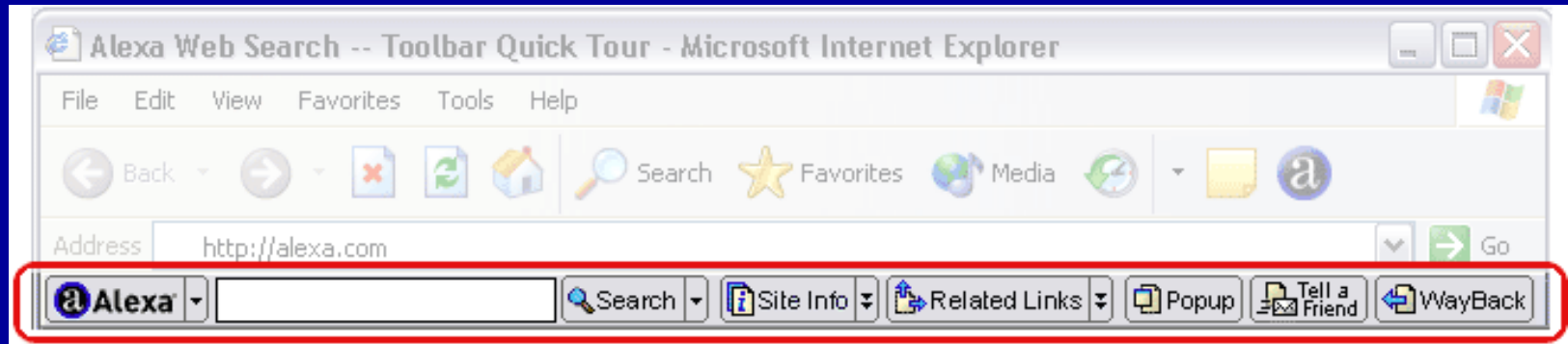
see also

Technical Responses to Unilateral Internet Authority:

*The Deployment of VeriSign
“Site Finder” and ISP Response*

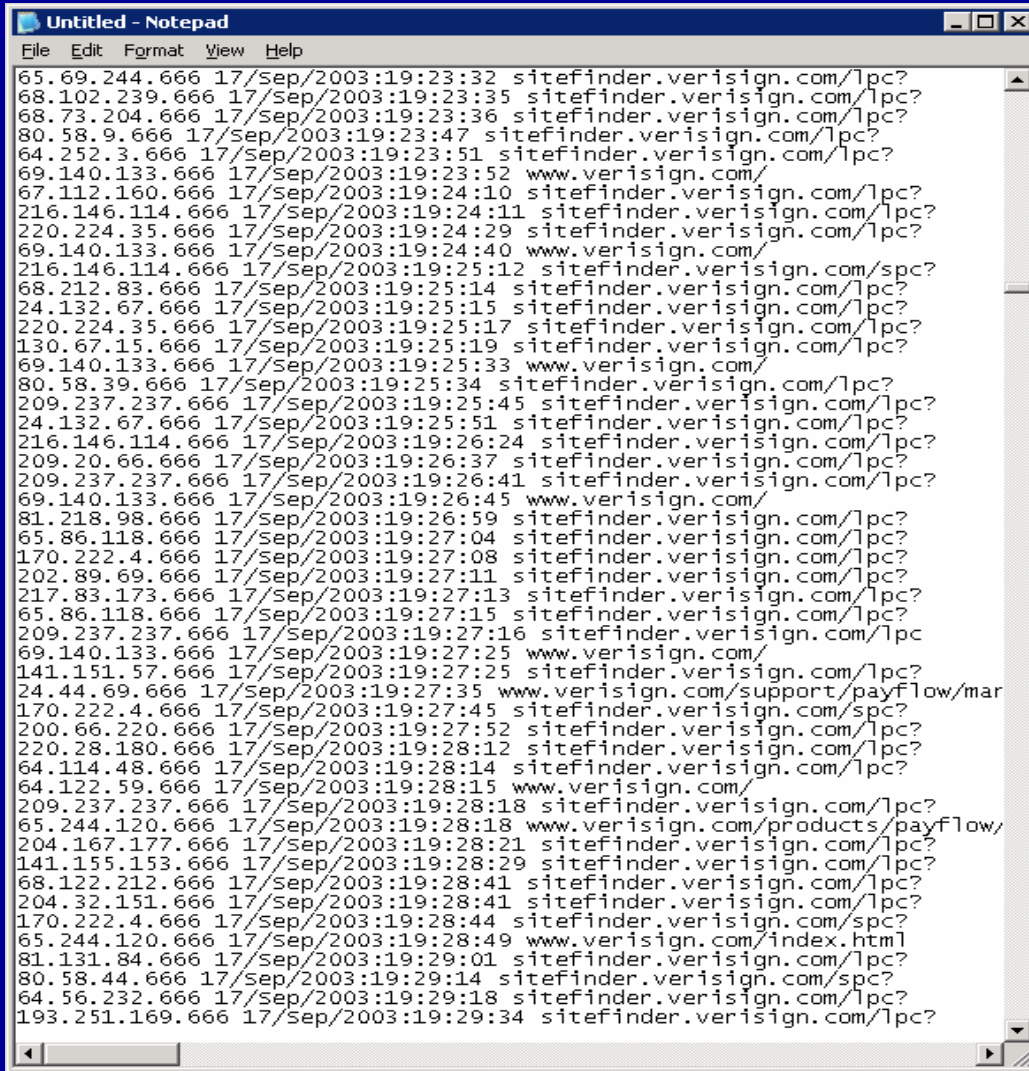
<http://cyber.law.harvard.edu/tlds/sitefinder>
with Jonathan Zittrain

Intro to Alexa Toolbar



- Provides search shortcut, related links, etc.
- 10+ million downloads. Active users unknown.
- Representativeness
 - Users generally representative of Internet community.
 - Possible under-emphasis on technical community.
 - Possible over-emphasis on Southeast Asia.

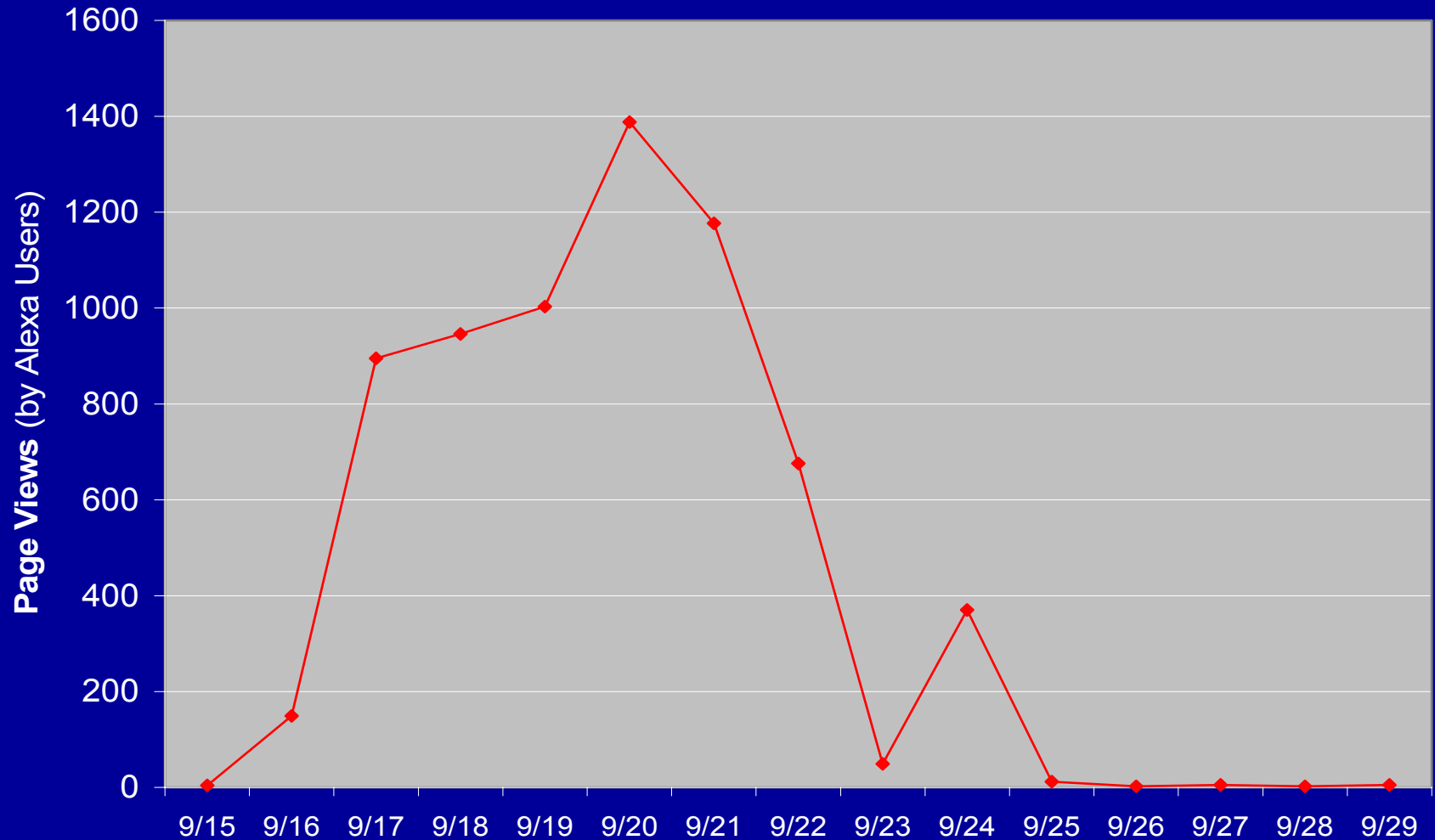
Alexa Data Set



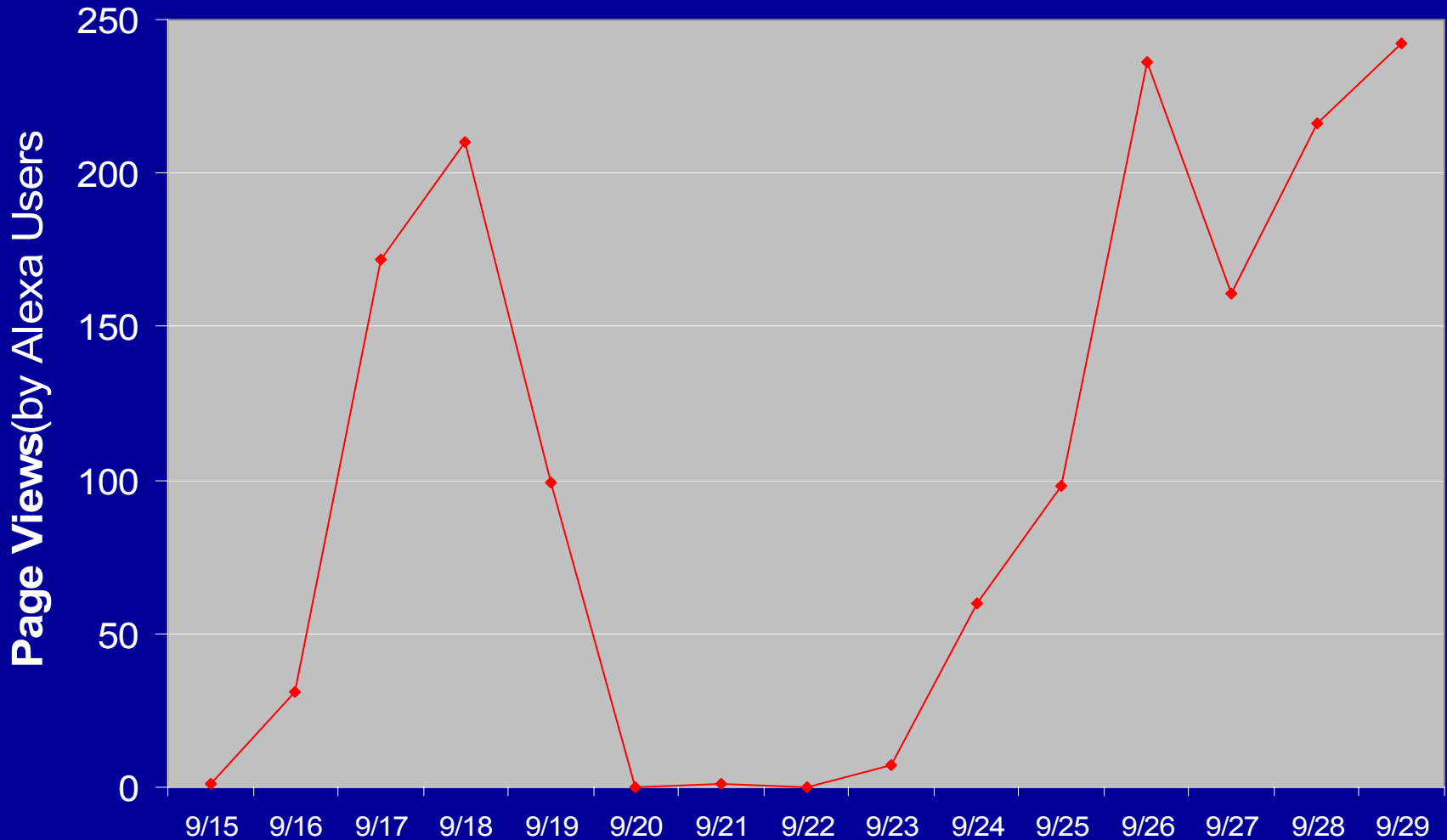
```
Untitled - Notepad
File Edit Format View Help
65.69.244.666 17/Sep/2003:19:23:32 sitefinder.verisign.com/lpc?
68.102.239.666 17/Sep/2003:19:23:35 sitefinder.verisign.com/lpc?
68.73.204.666 17/Sep/2003:19:23:36 sitefinder.verisign.com/lpc?
80.58.9.666 17/Sep/2003:19:23:47 sitefinder.verisign.com/lpc?
64.252.3.666 17/Sep/2003:19:23:51 sitefinder.verisign.com/lpc?
69.140.133.666 17/Sep/2003:19:23:52 www.verisign.com/
67.112.160.666 17/Sep/2003:19:24:10 sitefinder.verisign.com/lpc?
216.146.114.666 17/Sep/2003:19:24:11 sitefinder.verisign.com/lpc?
220.224.35.666 17/Sep/2003:19:24:29 sitefinder.verisign.com/lpc?
69.140.133.666 17/Sep/2003:19:24:40 www.verisign.com/
216.146.114.666 17/Sep/2003:19:25:12 sitefinder.verisign.com/spc?
68.212.83.666 17/Sep/2003:19:25:14 sitefinder.verisign.com/lpc?
24.132.67.666 17/Sep/2003:19:25:15 sitefinder.verisign.com/lpc?
220.224.35.666 17/Sep/2003:19:25:17 sitefinder.verisign.com/lpc?
130.67.15.666 17/Sep/2003:19:25:19 sitefinder.verisign.com/lpc?
69.140.133.666 17/Sep/2003:19:25:33 www.verisign.com/
80.58.39.666 17/Sep/2003:19:25:34 sitefinder.verisign.com/lpc?
209.237.237.666 17/Sep/2003:19:25:45 sitefinder.verisign.com/lpc?
24.132.67.666 17/Sep/2003:19:25:51 sitefinder.verisign.com/lpc?
216.146.114.666 17/Sep/2003:19:26:24 sitefinder.verisign.com/lpc?
209.20.66.666 17/Sep/2003:19:26:37 sitefinder.verisign.com/lpc?
209.237.237.666 17/Sep/2003:19:26:41 sitefinder.verisign.com/lpc?
69.140.133.666 17/Sep/2003:19:26:45 www.verisign.com/
81.218.98.666 17/Sep/2003:19:26:59 sitefinder.verisign.com/lpc?
65.86.118.666 17/Sep/2003:19:27:04 sitefinder.verisign.com/lpc?
170.222.4.666 17/Sep/2003:19:27:08 sitefinder.verisign.com/lpc?
202.89.69.666 17/Sep/2003:19:27:11 sitefinder.verisign.com/lpc?
217.83.173.666 17/Sep/2003:19:27:13 sitefinder.verisign.com/lpc?
65.86.118.666 17/Sep/2003:19:27:15 sitefinder.verisign.com/lpc?
209.237.237.666 17/Sep/2003:19:27:16 sitefinder.verisign.com/lpc
69.140.133.666 17/Sep/2003:19:27:25 www.verisign.com/
141.151.57.666 17/Sep/2003:19:27:25 sitefinder.verisign.com/lpc?
24.44.69.666 17/Sep/2003:19:27:35 www.verisign.com/support/payflow/mar
170.222.4.666 17/Sep/2003:19:27:45 sitefinder.verisign.com/spc?
200.66.220.666 17/Sep/2003:19:27:52 sitefinder.verisign.com/lpc?
220.28.180.666 17/Sep/2003:19:28:12 sitefinder.verisign.com/lpc?
64.114.48.666 17/Sep/2003:19:28:14 sitefinder.verisign.com/lpc?
64.122.59.666 17/Sep/2003:19:28:15 www.verisign.com/
209.237.237.666 17/Sep/2003:19:28:18 sitefinder.verisign.com/lpc?
65.244.120.666 17/Sep/2003:19:28:18 www.verisign.com/products/payflow/
204.167.177.666 17/Sep/2003:19:28:21 sitefinder.verisign.com/lpc?
141.155.153.666 17/Sep/2003:19:28:29 sitefinder.verisign.com/lpc?
68.122.212.666 17/Sep/2003:19:28:41 sitefinder.verisign.com/lpc?
204.32.151.666 17/Sep/2003:19:28:41 sitefinder.verisign.com/lpc?
170.222.4.666 17/Sep/2003:19:28:44 sitefinder.verisign.com/spc?
65.244.120.666 17/Sep/2003:19:28:49 www.verisign.com/index.html
81.131.84.666 17/Sep/2003:19:29:01 sitefinder.verisign.com/lpc?
80.58.44.666 17/Sep/2003:19:29:14 sitefinder.verisign.com/spc?
64.56.232.666 17/Sep/2003:19:29:18 sitefinder.verisign.com/lpc?
193.251.169.666 17/Sep/2003:19:29:34 sitefinder.verisign.com/lpc?
```

- Date through Sep. 29 as to requests for web pages on verisign.com by Alexa Toolbar users
- Data elements:
 - User IP address (*I*24)
 - Date & time
 - URL requested

Some Noteworthy Networks: China

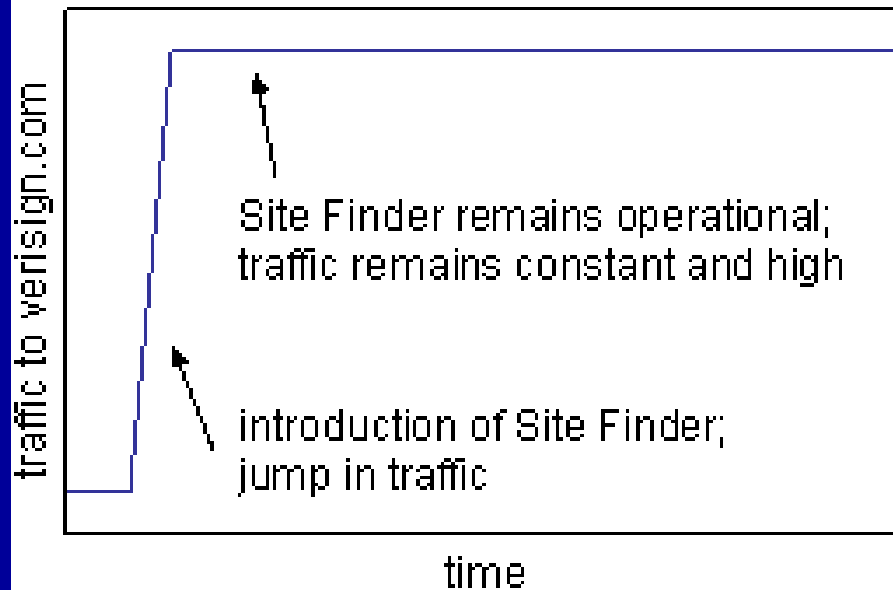


Some Noteworthy Networks: Adelphia

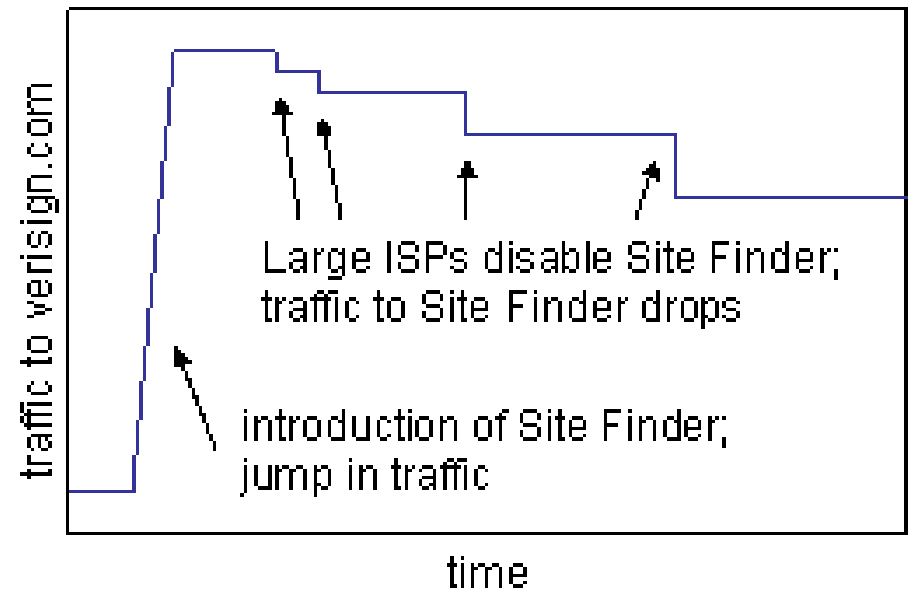


Aggregation: Idealizations

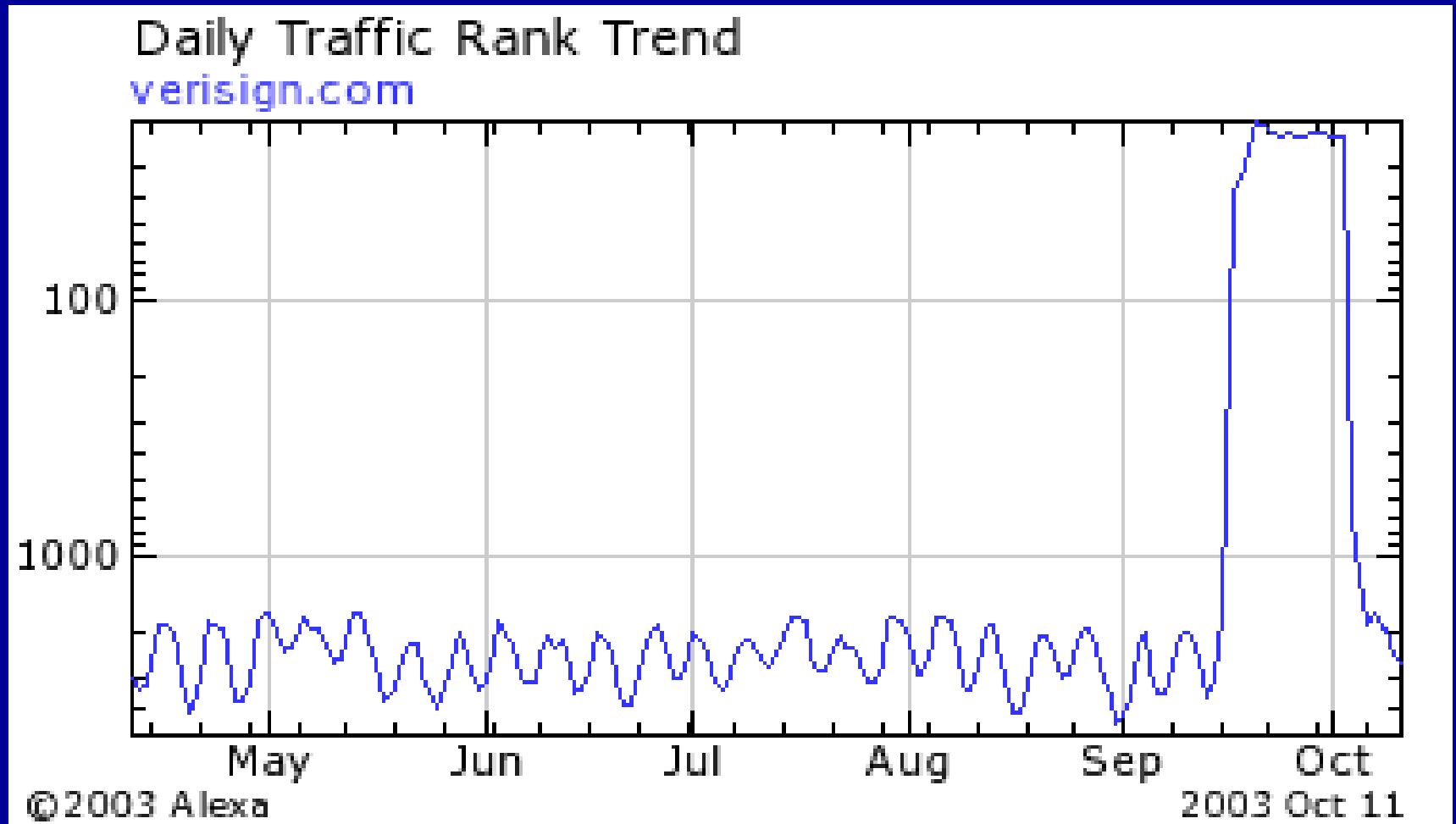
idealization of traffic to date



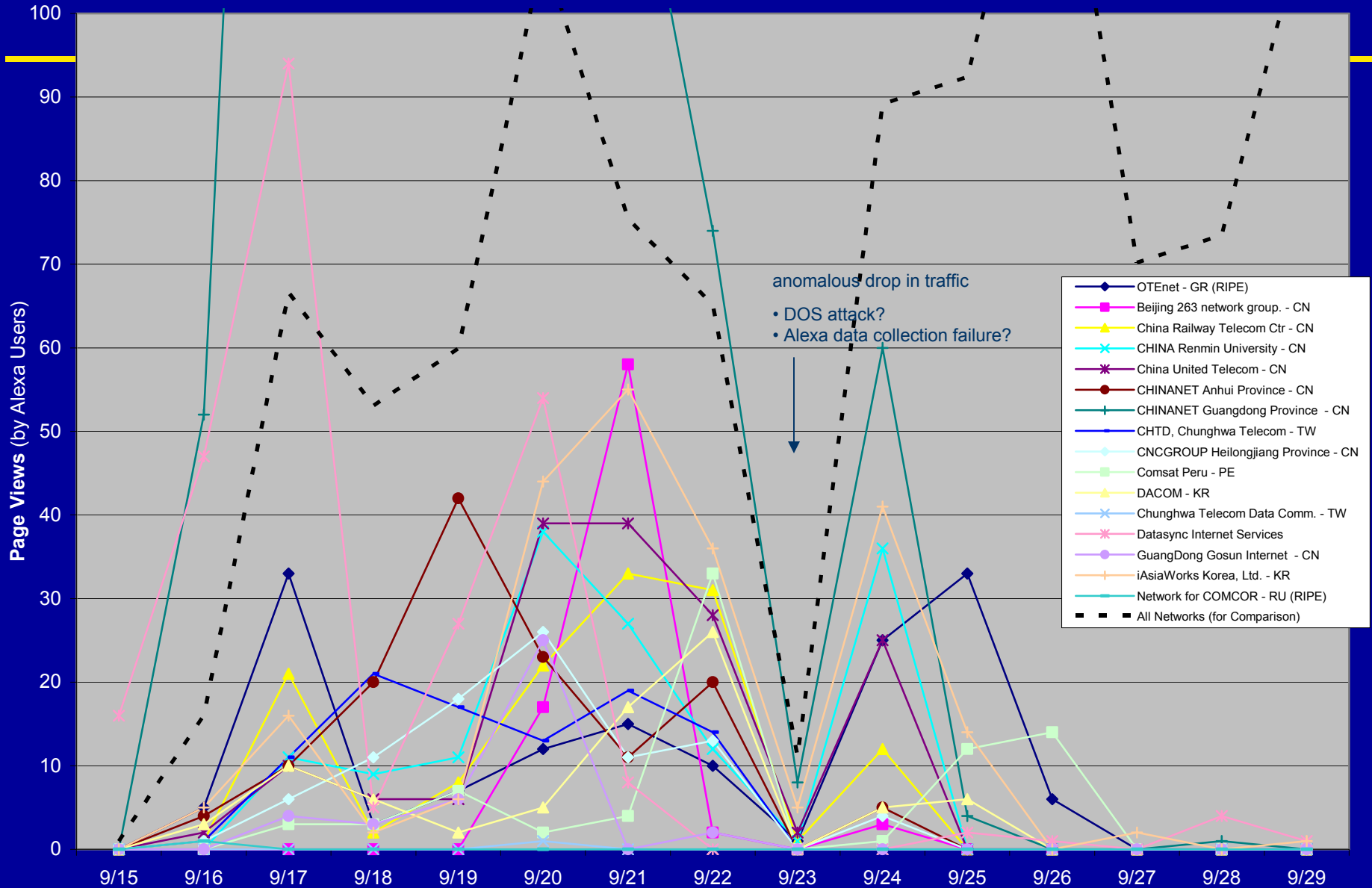
possible future trend in traffic



Aggregation: Data (ordinary, public Alexa rank chart)



Aggregation: Data (summed across all networks)



Additional Analysis with Add'l Data

- Possible data sources
 - Site Finder web server log files
 - Google Toolbar log files, other toolbars
- Could address Site Finder blocking
 - by smaller networks
 - by networks with fewer or no Alexa Toolbar users

Analyzing Page-Views on Site Finder

How many Site Finder users:

- Click on a “*Did you mean?*” link?
- Click on a “*Popular Category*”?
- Read the *Terms of Use*?
- Leave Site Finder without clicking on anything?

Search the Web:

Search

Did You Mean ?

We did find these similar Web addresses:

■ www.doesnotexist.com ■ www.doesnotexist.net

Search Popular Categories:

[Travel](#)

[Entertainment](#)

[Gambling](#)

[Shopping](#)

[Gifts](#)

[Computers](#)

[Autos](#)

[Insurance](#)

[Small Business](#)

[Investing](#)

[Health & Fitness](#)

[Home & Garden](#)

[Career](#)

[Education](#)

[Reference](#)

Copyright© 2003 **VeriSign, Inc.** All Rights Reserved
[Privacy Policy](#) | [Terms Of Use](#) | [Content Filtering Preferences](#) | [Help](#)

Analyzing Page-Views on Site Finder

URL	% of Accesses
sitefinder.verisign.com/lpc – <i>default Site Finder page</i>	81.58%
sitefinder.verisign.com/spc	16.19%
sitefinder.verisign.com	1.60%
sitefinder.verisign.com/help.jsp	0.18%
sitefinder.verisign.com/preferences.jsp	0.16%
sitefinder.verisign.com/terms.jsp	0.15%
sitefinder.verisign.com/privacy.jsp	0.11%
sitefinder.verisign.com/whatsthis.jsp	<0.01%
sitefinder.verisign.com/pdf/sitefinderdevguide.pdf	<0.01%

Note: No data available in Alexa logs as to clicks within “Did you mean?”

Session-Based Analysis

- “What proportion of users click on...” rather than “What proportion of page-views are for...”
- Difficult using Alexa Toolbar data because low order byte of IP address is unavailable.
 - 12% of user-sessions include at least one “*Popular Category*” view.
 - (“Session” = “request from same /24 within 5 minutes”)

Benjamin Edelman

<http://cyber.law.harvard.edu/edelman>
edelman@law.harvard.edu