

Originally prepared for Monash Uni audience

Website Metrics

Dr Lisa Wise

Overview

- how do we measure the effectiveness of a website?
- how do we determine cost-effectiveness of providing online services?

**“Not everything that can be counted counts,
and not everything that counts can be counted”**

Albert Einstein

Metrics

- not just about webserver log analysis!!
 - logs have huge amount of data
 - not enough *relevant* data
- what do you want to know?
 - generate appropriate questions
 - what type of data answers those questions?
 - analyse *relevant* data
- metrics are interpreted measurements

Originally prepared for Monash Uni audience

Log files

```
138.251.242.60 - - [17/Oct/2002:06:42:41 -0400] "GET /img/contours2.gif HTTP/1.1" 200 263  
"http://wisebytes.net/illusions/mullerlyer.php" "Mozilla/4.0 (compatible; MSIE 5.5; Windows  
98; Win 9x 4.90; UTVI)"
```

```
138.251.242.60 - - [17/Oct/2002:06:42:41 -0400] "GET /img/machbands2.gif HTTP/1.1" 200  
293 "http://wisebytes.net/illusions/mullerlyer.php" "Mozilla/4.0 (compatible; MSIE 5.5;  
Windows 98; Win 9x 4.90; UTVI)"
```

```
138.251.242.60 - - [17/Oct/2002:06:42:41 -0400] "GET /img/herman2.gif HTTP/1.1" 200 613  
"http://wisebytes.net/illusions/mullerlyer.php" "Mozilla/4.0 (compatible; MSIE 5.5; Windows  
98; Win 9x 4.90; UTVI)"
```

- **who:** ip address, user,
- **when:** timestamp relative to GMT,
- **what:** request, protocol, response, kB,
- **how:** referrer, browser/OS

Originally prepared for Monash Uni audience

Logfile Analysis

- log files can be very large
- analyse using standard packages
 - Analog, Webalyzer
- <http://www.its.monash.edu.au/web/stats/>
- for your own site, grep out relevant lines (eg /web/slideshows), and then use an analysis package to get your information

What does it all mean?

- hits versus page view
 - a hit is when a server responds to a request so if a page comprises html and 5 images, the server will score 6 hits to return 1 page
- unique visitors
 - can be by user id (rare) or by ip address (same computer ... maybe, but not necessarily same person at computer)

What does it all mean?

- Visit / Path / Time spent at site
 - a visit is a sequence of requests from same host until there is a break of 30 mins (or whatever timeframe you set)
 - you can see what was requested over how long, but you don't know what the user was doing (reading vs coffee ...)
 - don't know what was in cache, when they used the back button, when they left your site

Issues

- how do you identify visitors / visits?
- what do you do about cached pages?
- what are people doing while a page is in their browser?
- what are they looking for?
- do they leave because they found it, or they're lost, or they're bored, or they hate your site?

Nettracker

- can do sophisticated custom reports
- uses log files and makes many inferences
- takes a lot of time to import data and generate custom reports
- doesn't provide real-time analysis (need to import log files)
- will be able to request reports in the future (will notify when service is available)

What to ask?

- metrics measure specific things
- need to decide what needs to be measured
- requires identifying goals of your website
- requires identifying goals of your users

... oh no !!! It's beginning to sound like user-centred analysis again !!!

Web server performance

- to look at web server performance we need to measure
 - latency, throughput, utilisation
 - performance under load (webserver / network)
- to look at web service performance
 - need to measure from the user's perspective
 - include connection speed, network congestion, server processing, transmission

Website metrics

- Portal: role is to provide a gateway to other services - metrics should reflect this
 - not total number of visits, but rather the relative use of links to access other services
 - need to analyse what referral rates other services get from the portal
 - is portal the preferred entry point?
 - popular links should be prominent, unpopular should be removed

Website metrics

- teaching site (eg WebCT)
 - lecture notes are to read or to print?
 - see print queues? ask students?
 - which revision tools are most popular?
 - useful or are students confused?
 - when are discussion groups used?
- need to formulate questions for analysis
- some info in Track Students tool

Business Unit Website

- How do users arrive at content pages?
(are any compulsory? eg enrolment)
- Do you expect them to come via the home page? (difficult to measure if cached ...)
- Look at search terms used - is this good metadata or poor navigation?
- Users inside or outside Monash?
- Are they target users?

Cost effectiveness

- look at how often your site is accessed and by whom
 - how difficult was it to prepare content?
 - is the “costly content” the most used and is it used by your target audience?
- how many people are needed for web site versus non-web solution?
 - is properly-costed web site more efficient?

Some References

- Jim Sterne, 2002, *Web metrics: proven methods for measuring website success*, Wiley
- Patrick Killelea, 2002, *Web performance tuning*, 2nd Ed, O'Reilly
- <http://www.analog.cx/> (especially the link on what the results mean)