Association for Information Systems
AIS Electronic Library (AISeL)

ACIS 2001 Proceedings | Australasian (ACIS)

2001

Censorship rules: The Topology and Data Topography of Australian Adult Websites

David Harte
Edith Cowan University, dharte@student.cowan.edu.au

Mark Brogan
Edith Cowan University

Censorship rules: The Topology and Data Topography of Australian Adult Websites

David Harte, Mark Brogan

Edith Cowan University, Perth, Australia
dharte@student.cowan.edu.au

Abstract

This paper presents the results of preliminary research into the topology and data topography (amount of content per location) of Australian adult content websites. The research was conducted as part of a larger research project into the effectiveness and compliance implications of the Australian Government's Broadcasting Services Amendment (Online Services) Act 1999.

Keywords

Internet Content Control, Adult Website Topologies, Data Topography,
IT Public Policy Research

BACKGROUND

According to research conducted by the Inktomi Corporation, by 2000 the Web had come to comprise over 1 billion unique pages (Inktomi, 2000). The number of web sites was estimated in the same survey at 4,951,247. Pastore (2000) cites a Cyveillance study estimating the number of pages being added daily at around 7 million. In Australia, the focus of this study, the rate of growth is substantial, with an estimated more than 6 million adult Australians accessing the Internet as of November 1999, an estimated increase of 100% since February 1998 (NOIE, 2000). The content available on Internet spans the range of human interests, including such diverse topics as religion and pornography. Since the adoption of the Web in the early 1990s, adult web sites have mushroomed on the Internet.

In the era of the 'dot-bomb', the online adult industry stands out as a rare success story in Business-to-Consumer (B2C) online business. Martin (1999) reckons that the online adult industry earned $480 million in 1998. Koerner (2000) suggests that this is an under estimate and places earning for the same year at $1 billion. According to Buskin (2000) the top three adult sites earned over $100 million each in the same year, and the projected earnings for the industry in the year 2003 are estimated at $3 billion (Koerner, 2000). To put this figure into perspective, the E-commerce giant Amazon.com lost over $317 million in one quarter of 2000 (Hansell, 2000), and the total revenue for the adult publication giant Playboy was $138 million. Koerner (2000) cites a Neilsen-NetRating statistic that 17.5 million web users accessed adult websites in January 2000 alone.

The number of adult sites on the Internet is unknown but Mark Tiarra, president of the industry group, United Adult Sites (UAS) cited in Buskin (2000) suggests an educated guess at about 200,000. This number of sites helps to illustrate the scope and scale of the task of controlling access to adult content on the Web.

The extent and nature of adult materials has provoked a global public policy controversy over the need for Internet censorship. Australia became a player in the censorship game with enactment of the Broadcasting Services Amendment (Online Services) Act in 1999. The Act came into operation on 1 January 2000(Cwlth) (BSA). The Act was formulated against a background of protest from industry and cyber rights advocates.

THE ECU STUDY: RESEARCH QUESTIONS AND METHODS

The intractable public policy controversy that has emerged around online censorship is grounded in opposing positions about the social impacts of exposure to online pornography and the rights of the individual. The content control debate in the US and Australia has followed this pattern, with the rationale that content control protects children. The Communications Decency Act (1996) was the US Government's response to the calls for censorship and control. The US Supreme Court however, ruled (Reno Vs ACLU) that this law contravened the 1st Amendment to the US Constitution. Australian citizens do not currently enjoy comparable constitutional protection.

Public debate has mostly not been informed by grounded research into the measured impact of the regime on the online industry or investigation of popular assumptions about the operation of the adult web sites. We have sought to address this deficiency through an investigation of site topology and data topography of Australian Adult Websites. Data gathered on site topology and topography can be used to:

*assess the impact of the regime on the so-called data deficit (expected to grow as a consequence of the regime with significant cost burdens to the industry); and
*to assess the likely effectiveness of content blocking technologies and related tactics.

Economics: Shifting Content and the Data Deficit

During debate regarding the Bill in 1999, the Internet Industry and others argued that a likely consequence of this legislation would be that adult sites would be forced overseas resulting in losses to the local industry (Budde, 1999). If large volumes of such content were shifted, from local .au domain download to intercontinental backbone download, it would present a much higher cost to the service provider. The retrieval of content over the congested backbone is much more expensive than from a server within the Australian domain:

> "If a significant proportion of data in Australia is moved overseas, Australians must pay to import that data in the first place. Further, where before that data could be stored in Australia to satisfy later requests (that is, it could be cached or mirrored), now the data must be imported each time a content seeker requests it, multiplying the cost of providing the data. Billing mechanisms in the internet are relatively immature. Current technology is unable to premium bill content subsets – all content is charged at a flat rate. As such, there is no way to assign the increase in cost to those end users which are forcing the cost increase with the result that all end users must pay for use by the few. "(Scott, 1999a)

The issue of the so called 'data deficit' has been identified as contributing to the high cost of Internet access in Australia.

> "Forcing the content out of Australia also means that inbound traffic into Australia is increased. Australian carriers are currently forced to buy content from US carriers, but must give Australian content to the US carriers for free. One of the justifications for this is that traffic is 70:30 in the US carriers' favour." (Scott, 1999b)

This imbalance in carriage charges has the potential to force many Australian ISPs out of the market, especially when competing with overseas free ISPs and ICHs. The shift of significant content volumes from the .au domain to overseas servers and the subsequent worsening of the 'data deficit' is an important research question with serious implications for Australian ISPs and Internet services generally in Australia.

Measuring the deficit: Towards a research design

If the deficit has grown since 1999 as a consequence of online censorship, then a coarse measure of such growth might be expected in increased bandwidth consumption by ISPs. However this information is not readily accessible and its reliability is compromised by the widespread use of proxy servers. Alternatively, if the impact on the data deficit is real, then its causal explanation lies in an underlying trend in site hosting, involving the re-location of sites off-shore preceding or following promulgation of the Broadcasting Services Amendment (Online Services) Act.

The alternative topology/topography method was selected for deficit analysis. Trends in site hosting are easier and more reliably measured than changes in traffic volumes. Using packet tracking applications (such as Ping, Traceroute, NeoTrace), the geographic location and the IP addresses of web sites can be logged, providing the basis of tracking. The InterNic and AusNic registries can also be used to identify the owners of the virtual domain and subdomain sites. Sites can be examined in detail using sitemapping software (such as SiteMapper) to reveal both their topography in terms of the amount of content and also their topology. With sophisticated site tracking it is possible to determine if sites have changed their host country. This process can be sufficiently sophisticated to avoid errors of misattribution.

Geographical Domain Mapping

Data topography can be thought of as an exercise in geographical domain mapping. Geographical domain mapping describes the physical location of hosted content. The use of individual country domain designations such as .AU, .UK etc does not guarantee that the website is hosted within these countries (Australia and the United Kingdom in the example). The geographic location of the website is completely decoupled from the URL. However, it is apparent that the use of country domain designations helps to promote the popular misconception of Internet domains, which map directly to geographic domains. In the sense that Internet domains can be misleading in terms of geographic names, and these names cannot be used reliably to determine the place of content hosting.

THE TECHNICAL PROBLEM

Blocking content

In June 1998 the Commonwealth Scientific and Industrial Research Organisation (CSIRO) Division of Mathematical and Information Sciences released an investigation of content blocking in a report entitled Blocking Content on the Internet: a Technical Perspective (McCrea, Smart & Andrew, 1998). The Report was commissioned by the National Office for the Information Economy.

McCrea et al (1998) concluded that application level filtering using Universal Resource Locator (URL) blacklists, was feasible. The 1999 Report identified a number of products categorisable as end user, server or ISP and hybrid products all employing the application level blocking principle.

The second Report was both significant and influential, with application level blocking subsequently operationalised within the approved industry code. The use of application level filtering facilitates the complaint, review and takedown structure of the co-regulatory system that exists today. The URL of any website, which becomes the subject of a takedown notice, will be added to the blacklist of all the filtering packages included in Schedule 1 of the Code. The procedure for this is based on the ABA issuing a notice to filter manufacturers of prohibited content discovered as the result of complaint investigation and may also issue a notice to Internet Service Providers.

The Consequences of Search Engine Manipulation

Probably the major problem for commercial adult websites is advertising their presence on the Web. As indicated earlier, the Web is growing at a rapid pace and the problem of website promotion is exacerbating. Some studies have shown that most users find content on the Web using search engines, and a prominent listing on search engine is vital (Coopee, 2000). With this fact in mind, successful webmasters devote considerable time and effort to make sure their site features prominently in search engine query results (Glidewell, 2000). This is not a trivial task as the success or failure of a website may depend on a high rating from a search engine.

Adult content websites use many techniques to manipulate search engines to gain a prominent position in a query listing. The search algorithms of popular search engines are deduced by submitting multiple pages and studying the success rate for different techniques (Glidewell, 2000). Some websites use pages of keywords to attract a listing. In this study one website was found to have 201 pages of content comprising nothing but keywords included in the HTML metatags and the main body of the page. Many websites were found with multiple subdomains and subdirectories. This can result in many more search engine database entries, a problem for both client and server side content filtering.

Website Referral and Link-Back agreements

The marketing strategies employed by adult sites depend heavily on website referral. This is a major contributing factor to the prevalence of portal sites, with the large megasites paying the small websites for successful referrals. The marketing strategy is employed extensively in adult websites on the Internet. The scheme works as follows:

- A large adult site in the US offers to supply some content (usually about 200 images) to a smaller portal site, for marketing purposes.
- The portal site uses the supplied images to attract customers and maintains a hyperlink to the main megasite(s). The only content on the portal is the supplied images.
- The large site operator pays a royalty to the portal operator for every referral that results in a membership transaction.

The use of this strategy has resulted in a large percentage of adult websites being mere portals with hyperlinks to many of the large megasites, with very little content hosted by the portal sites. The use of what are known as Link-Back agreements (Glidewell, 2000) is also highly prevalent. The strategy works as follows:

- An agreement is made between two large content providers to hyperlink each other's site within client-push JavaScripts.
- When a customer accesses a site, each attempt to leave the site activates a JavaScript to open another of the site's webpages.
- When the website has exhausted all its possible content offerings, it activates the hyperlink to the other website in the link-back agreement.
- The new website follows the same protocol and may eventually link-back to the original referring site

The use of this marketing strategy has three effects:

- It maximises the possibility of making a sale and therefore a profit;
- It increases the topological complexity of adult websites; and
- It increases the Internet download volume generated by adult websites considerably.

RESEARCH FINDINGS: AUSTRALIAN ADULT WEB SITE TYPOLOGY AND TOPOGRAPHY

The following is a summary of principal research findings arising from the ECU Study:

The myth of Australian hosted objectionable content

Substantive research began in last quarter 1999, i.e. a full three months before implementation of the Broadcasting Services Act Online Services Amendment Act was promulgated on 1 January 2000. A variety of search engines (GoEureka, Anzwers, Yahoo, Altavista) were used in the study to identify Australian adult content. All searching was confined to the .au domain. Searching resulted in an average of about one thousand five hundred (1500) response hits. It must be pointed out that this figure is quite dynamic due to the nature of the Internet and the content provided. Repeated tests using different search engines revealed variability in this figure of about thirty per cent (30%).

These results required considerable refining and sorting to remove multiple page hits from the same website and discontinued links. This yielded a total of eight hundred and ten (810) discrete search engine hits. This figure was further processed to yield the URLs of discrete websites. The final total of discrete website URLs was two hundred and fifty five (255). These URLs were collected in a database for later SQL query analysis.

In this study of over fifteen hundred (1500) search engine response hits, not one commercial website hosting significant levels of adult content was located within Australia. Most of the content found was usually that provided by the large US hosted commercial sites in return for referrals.

Based on the .au domain, as of last quarter 1999:

1. The Australian online adult content industry was highly concentrated -only twenty three (23) operators were identified in this study and anecdotal evidence from within the industry would seem to confirm this finding;
2. The majority of Australian sites identified (98.4%) are portal sites to offshore hosts principally in the US;
3. Only twenty three (23) (or 9.0%) of the commercial adult websites were virtual domain sites;
4. One hundred and seventy two (172) (or 67.4%) of sites used subdomain or subdirectory domain names;
5. Two hundred and nineteen (219) (or 85.9%) of adult sites identified are hosted in the US (despite the .AU designation);
6. All of the commercial sites identified use third-party credit card validation services, again hosted in the US;
7. There is virtually no content apart from images used for marketing, website referral, and site promotion, hosted on Australian servers. The actual content for sale is hosted on, mainly US servers.

INTERPRETATION OF RESEARCH FINDINGS

Offshore hosting and the data deficit

It was widely expected that introduction of online censorship would lead to an *exodus* of adult web sites from Australian ICHs' and a contraction in ISP revenues. The research however would seem to contradict this position. By last quarter 1999 and months before the introduction of the new censorship regime, in excess of eighty five per cent (85.9%) of .au adult sites were already being hosted off-shore. Similarly, claims that the new regime would substantially add to the data deficit are unsustainable given that very little adult content appears to have existed on Australian servers by late 1999.

The real implication for the Internet Industry, given the current content control protocol, is in terms of loss of revenue from the removal of portal sites and not increased overhead in terms of bandwidth.

It must be stressed that this situation is dependent on the current arrangements remaining in force. Any attempt at increased ISP based filtering and control could change this state.

A comparison of hosting charges in the US and Australia

Comparison of hosting costs between the US and Australia reveals a massive differential between the two domains. Some offshore free website ICHs are now offering free hosting to adult sites with 25 Mb of storage

and the use of CGI scripting . This storage capacity would not be sufficient to run a major adult site, but would be suitable for a sophisticated portal site capable of generating good earnings, if constructed properly. The large American ISP CaveCreek with over 200 servers is one of the largest in the world. CaveCreek is reputed to be host to over 12,000 adult sites, one of the largest collections of adult websites in the world (Glidewell, 2000). A comparison of CaveCreek's hosting charges (http://www.cavecreek.com/) and some of the large Australian ISPs (Ozemail, iiNet, etc) reveal a major disparity in charges. As indicated earlier, adult content sites are high storage, high volume websites. For a high volume adult site hosting 500Mb of content and 300 Gb of volume per month, the cost of hosting in Australia is 20 to 30 times more expensive.

This level of hosting overhead was identified by some Australian adult site operators as the primary reason to host offshore.

They argue it is simply too expensive to operate a high volume adult site from Australia. They identified this factor as the defining one for the Australian adult content industry. However, it is obvious that this may also have implications for the general E-commerce industry in the future. With adult content on the Internet put at 70 to 80% of total traffic (Koerner, 2000), it is this factor that perhaps contributes most to the so called 'data deficit'.

Research findings: Effectiveness of content controls

As mentioned in the Technical Background section, the CSIRO concluded that URL match filtering was technically feasible. The conclusion was that the use of blacklists of banned URLs would not seriously degrade system performance. It may be argued however that this prognosis may only be maintained if the size of the URL blacklist is reasonable. The maintenance of a blacklist with potentially billions of URLs would not be a trivial problem. Even partial URL filtering using wildcard values would be problematic.

Effectiveness of content controls was not a research question in the ECU study. But data typography information gathered suggests that the topology of the Internet adult content industry is extreme. Case study showed that sights may have potentially thousands of webpages resolving to just one IP address. Multiply this number by the amount of referral links to and from other sites plus the very high number of adult sites globally, and the true nature of the scale of the problem is revealed. URL filtering is easily circumvented leaving the content unaffected and untouchable.

CONCLUSION

This research has shown that most Australian online adult content is in fact hosted in the US. The reasons for offshore hosting is almost totally financial and pre dates introduction of the Broadcasting Services Act (Online Services) Amendment Act. The removal of adult sites from Australian servers should not contribute much to the so called 'data deficit'. The researcher conclude that of grounded research on data topology and typography had been conducted by the Australian Government in 1999, it would have found that legislation aimed at suppression of objectionable content hosted in Australia would have little to do.

The topology of sites suggests that it may be technically impossible and financially very expense to implement any effective content control, including the URL filtering endorsed by the CSIRO. As suggested, the present content control arrangements may be totally ineffective and very expensive. The extreme topology of the online adult content industry is a virtual latter day 'Gordian Knot' intricately woven into the fabric of the World Wide Web. It may be that the only way to control adult content on the Web would be to 'cut the knot', but this would surely result in severe damage to the Web itself. As is all too often the case, the remedy may prove more fatal than the disease.

REFERENCES

Broadcasting Services Amendment (Online Services) Act (Cwlth) (1999), [online]. Available WWW: http://scaleplus.law.gov.au/html/comact/10/6034/rtf/122of99.rtf [2000, April 20]

Budde, P. (1999, August 24) *Regional ISPs Under Threat from $150m Censorship Compliance*. The Australian.

Buskin, J. (2000, April 17) *E-Commerce (A Special Report): Industry by Industry --- The Web's Dirty Little Secret: What porn sites lack in respect, they make up for in profits.* Wall Street Journal. [online] Available WWW: http://www.proquest.com/ [2000, August 8]

Communications Decency Act (U.S. Congress) 1996. [online]. Available WWW: http://www.epic.org/free_speech/censorship/copa.html [2000, February 10]

Coopee, T. (2000, June 12*) How to climb the search engine rankings*. InfoWorld. [online]. Available WWW: http://www.proquest.com/ [2000, August 23]

EFA (2000). *NCD:FAQ - Australian Internet Censorship Laws.* Electronic Frontiers Australia. [online] Available WWW: http://rene.efa.org.au/liberty/faq_auscens2.html

Glidewell, R. (2000, April) *Business lessons from online porn.* Upside 12 (4). pp. 194-208. [online] Available WWW: http:// www.proquest.com/ [2000, August 8]

Hansell, S. (2000, July 27*) Amazon Reports Losses of $317 Million in the Second Quarter.* New York Times. [online] Available WWW: http://www.proquest.com. [2000, August 8]

Internet Industry Association. (1999) *Internet Industry Code of Practice* IIA [online]. Available WWW: http://www.iia.net.au/Internet Industry Association.htm [2000, January 20]

Internet Industry Association. (2000) *IIA Guide to ISPs: Content Regulation Checklist* Available WWW http://www.iia.net.au/index2.html [2000, 28 September]

Kahaner, L (2000, June 12*) Content matters most in search-engine placement.* Informationweek. (790) pp. 172-178. [online] Available WWW: http://www.proquest.com/ [2000, August 23]

Koener, B (2000, March 27) *A Lust for Profits Pornography is a huge a growing cyberspace draw. But will salacious Web sites be a hit on Wall Street?* U.S. News & World Report 128 (12) pp. 36-44. [online] Available WWW: http://www.proquest.com/ [2000, August 8]

Martin, L. (1999). *Cyberporn: Slipping Through the Net.* [online] Sydney Morning Herald Online. Available WWW: http://www.smh.com.au/news/9903/20/features/features3.html [2000, April 12]

McCrea, P & Smart, B & Andrew, M. (1998*) Blocking Content on the Internet: a Technical Perspective.* CSIRO.

The National Office for the Information Economy (NOIE) (2000*)The Current State of Play: Australia and the Information Economy* [online], NOIE, Available WWW:

http://www.noie.gov.au/projects/information_economy/ecommerce_analysis/ie_stats/state_of_play.htm [2000, May 5]

Pastore, M. (2000) *The Web: More Than 2 Billion Pages Strong* [on-line]. Available WWW: http://cyberatlas.internet.com/big_picture/traffic_patterns/article/0,,5931_413691,00.html [2001, July 10]

Reno v ACLU (US Supreme Court No. 96-511, 26 June 1997) [online]. Available WWW: http://www.epic.org/cda/cda_decision.html [2000, April 15]

Scott, B. (1999a) *A Layman's Guide to Internet Censorship in Australia* [online]. Available WWW: http://www.gtlaw.com.au/pubs/sobershock.html [1999, September 12]

Scott, B. (1999b). *The Dawn Of A New Dark Age - Censorship and Amendments to the Broadcasting Services Act* [online]. Available WWW: http://www.gtlaw.com.au/pubs/newdarkage.html [2000, April 12]

COPYRIGHT