Address-based Content Blocking: An Introduction

Brussels 2nd Februrary 2007

Malcolm Hutty
malcolm@linx.net

Contents

- Context for blocking
- Methodologies for blocking
- How easy is it to avoid?
- Other considerations

Context for Blocking

Purposes of Content Blocking 1

- Protection
  - Help the users to avoid material that they do not wish to encounter
- Compliance
  - Prevent users from accessing material that they are actively seeking

Purposes of Content Blocking 2

- Protection
  - User does not want to access blocked material
  - User will not deliberately subvert blocking system
  - User's normal usage will usually not strain the blocking system by introducing difficult cases
- Compliance
  - User wishes to access blocked material
  - User may deliberately subvert blocking system

Examples of Protection

- “Phishing”
- Viruses and other malware
- Protecting ordinary users from viewing child pornography
- Helping children not to mistake “gambling” for “computer games”

Examples of Compliance

- Preventing terrorists accessing “bomb making” instructions
- Preventing paedophiles accessing child pornography
- Preventing gamblers accessing gambling sites

Examples of mixed cases

In these cases, some users may wish to be blocked, some may not

- Preventing teenagers accessing pornography
- Preventing Muslims accessing extremist ideologies
- Preventing the “curious” accessing banned material

Methodologies of Blocking

Methodologies of Blocking

- End-user filtering
- Web Proxy filtering
- DNS poisoning
- IP blocking
- Hybrid IP blocking/proxy filter

- (Alternative to blocking: removal)

End User Filtering

- Methodology
  – Software installed on each PC prevents access to certain materials
- Financial Costs
  – From around €50 per PC
  – Falls on customer
- Non-financial costs
  – Choice of sites to block can be questionable
  – Classification of sites can be questionable

End User Filtering 2

- Features
  - Commonly targets web, e-mail
  - Rarely targets Games, IM, Peer-to-Peer etc
  - Vibrant commercial market means state of the art is continually advancing
  - Customer has choice of a wide range of reasons for sites to be blocked (e.g. pornography, violent imagery, gambling, racism, even “lack of educational value”)

Web Proxy Filtering

- Methodology
  - All web traffic passed through a proxy cache, which selectively refuses access to particular web pages
- Financial Costs
  - Very high (€100,000s for an ISP with 50,000 customers)
- Non-financial costs
  - Can slow down network traffic
  - Can reduce network reliability
  - But no overblocking

Web Proxy Filtering 2

- Features
  - Centralised mandatory blocking of all web traffic
    - Generally, limited block-list from a qualified source e.g. court, IWF
  - Does not block non-web traffic

DNS Poisoning 1

- DNS is the system that translate human-readable addresses into machine-readable Internet protocol addresses
  - Example DNS address: www.google.com
  - Corresponding IP address: 216.239.59.147
- Every ISP provides a “DNS resolver” to look up these translations for its customers.
  - Each customer configures their PC to use their ISP’s DNS resolver as part of the process of connecting to that ISP
  - Whenever they visit a new website (or use any other Internet resource), their PC contacts the DNS resolver to discover the IP address to contact
  - Customer could instead configure their PC with any other DNS resolver, e.g. from an American ISP or one they run themselves

DNS Poisoning 2

- Methodology
  - ISP configures DNS resolver to lie about existence of sites to be blocked
- Financial costs
  - Low (Can be less than €5000 per ISP)
- Non-financial costs
  - Massive over-blocking, as a whole domain is blocked (e.g. all of MySpace, Geocities, terra.es etc)
  - Surprisingly difficult to implement without errors

DNS Poisoning 3

- Features
  - Blocks more than just web;
  - But non-use of DNS by site operators can limit effectiveness; and
  - Over-blocking is a serious problem, and can cause user rejection

IP Address Blackholing 1

- Methodology
  - ISP prevents all traffic from routing to specified IP addresses
- Financial costs
  - Depends on length of block list
- Non-financial costs
  - High level of overblocking due to shared web space (e.g. all of MySpace, Geocities, terra.es etc)

IP Address Blackholing 2

- Features
  - Blocks access for all protocols
  - Over-blocking is again a serious problem

IP Blackhole/Proxy Hybrid (Cleanfeed)

- Methodology
  - Use the same technology for IP-based blocking to route only selected traffic to a proxy; the proxy decides what to block
- Financial Cost
  - Less than full proxy, but still substantial
- Non-financial costs
  - Over-blocking greatly reduced / eliminated

But does it work?

How hard is it to avoid so-called "mandatory" blocking?

Proficiency levels required for avoidance

VERY HIGH	Advanced network software research
HIGH	Good understanding of networking principles. Basic software development skills.
MODERATE	Can search for and find obscure or complex software. Can follow complex instructions. Capable of imagining secondary uses of “dual-purpose” software.
LOW	Aware of common applications e.g. peer-to-peer. Capable of following written instructions to download, install and use such software.
VERY LOW	Can use web browser, e-mail. Cannot set up own computer to use Internet

Avoiding Blocking Systems 1

- End User Filters
  - Removal by PC owner (**LOW** expertise)
  - Surreptitious by-pass by PC user (**MODERATE** to **VERY HIGH** expertise)
- DNS poisoning
  - Use different ISP's DNS resolver (**LOW** expertise)
  - Run your own DNS resolver (**MODERATE** expertise)
  - Avoid or confuse DNS (**MODERATE** expertise)
  - DNS-SEC will make this obsolete

Avoiding Blocking Systems 2

- All methods except End-User Filters
  - Use Peer-to-Peer (**LOW** expertise); only provides access to content, not applications such as gambling sites
  - "Anonymizer.com" style tunnel (**VERY LOW** expertise)
  - Create your own encrypted tunnel (**MODERATE** expertise)
  - Confuse the blocking system with technical attacks[1] (**MODERATE** to **VERY HIGH** expertise, variable effectiveness)

[1]Simple examples include URL Character encoding, web file-path traversal with ".." etc

Other considerations

What gets blocked

- A blocking system is only as good as the list of items to block
  - Recognised authority or private company as source?
  - Only “known” material will be blocked
    - Not password-protected or otherwise hidden material
- Address-based blocking can only block material that has an address
  - Excludes Peer-to-Peer, e-mail, Instant Messages etc

Hybrid approach and the Oracle attack[1]

- Nature of Attack
  - With a **MODERATE** to **HIGH** technical expertise, a user can reverse engineer a hybrid system and discover the list of blocked sites.
  - This may have occurred already, although this is not known
  - Over time the difficulty is likely to be reduced to a **LOW** expertise requirement.
- Implication
  - Hybrid blocking systems could be inadvertently “advertising” a verified guide to child pornography for paedophiles

[1] The “Oracle Attack” was discovered and named by Dr Richard Clayton of the University of Cambridge Computer Laboratory

Geopolitical issues

- Many undemocratic non-EU countries engage in censorship for domestic purposes
  - Blocking in the EU is cited as legitimising their censorship (e.g. China)
- Blocking material hosted in another county could be viewed as an “attack” on that country’s Internet access
  - Analogous to radio jamming
  - Especially credible if the effect of blocking “spills over” across jurisdictions, because EU networks serve non-EU countries too

Undermining the end-to-end principle

- The end-to-end principle is a basic organising principle of the Internet
- It says that intelligence occurs at the network edges, not in the core routers
- It permits technological development, including invention of web, VoIP, etc
- Requiring blocking at the network level undermines the end-to-end principle and the capacity for invention
- Arguably, it invites network operators to subvert the end-to-end principle further