Internet Filtering and Censorship Proposals

Summary notes from the talk
by Siva Sivasubramanian at UNSW

Introduction:

I am not going to discuss censorship, its merits or demerits. That is outside the scope of my talks and I am leaving that to my fellow speakers in the ensuing sessions.

I am going to focus on filtering, discuss methods of filtering; merits and demerits of different filtering techniques, and how filtering could lead to Clean Internet Initiative.

Current Objections to Filtering:

Objections stem from Ethical, Commercial and Technical angles.

Ethical objections are based on who decides what to be filtered and if that 'who' should really have that authority.

Let me leave this objection to be discussed by others to follow.

Commercial Objections are based on the legitimate premise that it would cost more to provide Internet Service

Technical Objections are based on the premise that filtering will slow the Internet and filtering is an imperfect science at best.

The technical and commercial objections are very legitimate and let us discuss them in detail.

What are the requirements for filtering?

Once the filtering moves from individual level to that of an ISP level, complexities creep in.

What is legitimate content for some could be offensive content for others.

Take the example of SPAM, it is a legitimate business operation for some, they call themselves as bulk mailing business.

In an office, you can filter spam relatively easily, what is not business related is SPAM, add a caveat; excepting some reasonable personal mails – Now the nightmare for the system staff begin.

Extend this reach to the next level do this Spam filter at an ISP level. How do you classify SPAM across all your users? What is Spam for some could be legitimate mail for others.

What we call junk mail in our mail boxes are treasured by several folks – they meticulously go through them. My neighbour usually gives me an extra bag to store all her junk mails. She is a junk mail junkie! How do you handle such cases during a content filtering?

With difficulty offcourse!

At Office, when I tell my boss that I filtered off a million spam last week, he asks me for proof that I did not accidentally filter a legitimate business mail.

What will you do when your customers ask? How do you prove them that your filtering did not affect them?

With difficulty offcourse!

These situations could lead to filtering being customer or customer segment sensitive.

Here I have made a leap of faith, I could be wrong, the filtering may be imposed as universal leveller in which case it would reduce the technical impediments but create other more vexing issues. – Again discussion issue for my learned colleagues.

So let us again assume that filtering is going to be selective. It could be customer segment or context sensitive.

Let me make it clear, we are trying to identify the filtering requirements and are using Spam as a datum to derive our requirements. I hasten to add that Spam filtering and Internet filtering are different but the techniques and principles from one domain can be borrowed to another.

Filtering Techniques

Index or Reputation Based Filtering

Analysis or Content Based Filtering

Indirect Methods or Proxy Methods of Filtering

Index based filtering:

These techniques based on source criteria and Content filtering techniques are based on analysis of content.

We could borrow a lot from the SPAM filtering techniques and technologies. They follow either method admirably well.

Spam filtering uses both brute force of computing power to sift out trash from treasure as well uses nimble techniques that segregate trash even before they hit the threshold of their mail systems.

We need to borrow heavily from those techniques.

Both Index and Analysis techniques are individually inadequate but in combination, they provide a decent solution.

So could they be tried in the content filtering context?

Index Based

Reputation Services (used by AV and Anti Spam Vendors)

Global Indexing of sites and URLs for potential offensive content

This is inadequate,

(a) Non English Content

This is traffic driven, what about those ethnic sites that have lesser traffic than English sites?

So you need local Indexing of sites or self learning Indexing

This is an extension one needs to build for each situation.

(b) Hidden content (IP address only access for classified access)

(c) Is easily bypassed by reasonably savvy persons.

Analysis based Filtering

Done through content analysis

It is the technique whereby content is blocked or allowed based on analysis of its content, rather than its source or other criteria. It is most widely used on the internet to filter email and web access.

Do you evaluate picture or text or both?

How do you evaluate in multiple languages?

How do you circumvent the tricks the content makers play?

Is this a limitation or challenge? It is like that thin but subtle difference between Protection and Capture.

Indirect Methods or Proxy Methods of Filtering

Let us borrow techniques from pathology, for example Heart Attack is determined through blood test where presence of certain enzymes confirm heart attack.

Similarly,

Offensive content could be identified by the presence of infection it carries. Often sites hosting such contents tend to be laced with malware. Such sites tend to scam for the victims do not complain.

Therefore virus / malware filter software could be used to block such contents.

Contents that are clean but offensive are the ones more difficult to catch. We need precise definitions for such contents but the problem is such definitions at best are evolving and will continue to evolve.

Let me make it clear once again, that I am suggesting that we borrow the techniques, principles and building blocks from Spam domain in our pursuit of building Internet filtering capabilities.

Limitations of filtering

Filtering keeps those voyeurs in spirit from becoming voyeurs in practice.

It is like our windows and sliding doors in our houses – Keeps honest ones out, criminals can still get in.

Present Scenario - is it a Challenge or Limitation?

This weakness is a great opportunity

Not for criminals

For technologists, academics – to devise newer and better technology

All problems can be solved

With time and resources

So are we seeing the dawn of next super fast but clean Internet?

I don't have a solution now and here but have the confidence that we will get there sooner that we think.

In conclusion

Ethical objections – Let us leave it to my learned colleagues

Commercial Objections, it will cost more in the short term yes, but cost will rapidly fall. Example bandwidth costs; Moore's law has never failed

Technical Objections– The filtering is at best an imperfect science now. Yes, but perfection is there for grabbing, no new technology is needed but innovative adaptation of available technology is what is needed now.
This is the challenge I expect is going to reward the technology the most

Key fact to remember: Every source or every type of content have their own unique signature, they may mutate but sill is very unique, the trick is to recognize it and keep one step ahead – We did it in Spam control. Can we do it with filtering?

Like the space ideals leading to material science, atomic exploits – leading to nuclear medicine, number theory advances leading to cryptology; - this challenge will enrich the technology to the next level.