

A Multi-Stage Filtering Scheme for Internet Access Control

Ghassan F. Issa

Applied Science University
Amman, Jordan 11931
e-mail: issa@index.com.jo

ABSTRACT

تعتبر شبكة الإنترنت بخدماتها المتعددة مثل البريد و التحدث الإلكتروني و الشبكة العنكبوتية الطريقة الأكثر فعالية لنقل و تبادل المعلومات. إلا أن هذه الشبكة غدت مزدهمة بالمحتويات التي سببت تصادماً واضحاً مع عاداتنا و تقاليدنا بل تصادماً مع القانون. المواقع البذيئة و قليلة الاحتشام ذات المحتويات الإباحية و المواقع التي تشجع على الكراهية و العنصرية هي ليست إلا بعض الأمثلة من المواد المتوفرة لكافة مستخدمي شبكة الإنترنت و من ضمنهم الأطفال. و لحين توفر التشريعات و القوانين الصارمة لحماية المستخدمين فقد بدأ العديد بالبحث عن سبل لمراقبة و تنقية محتويات الإنترنت لجعلها أكثر أماناً للأجيال القادمة. ويستعرض هذا البحث بعض طرق الحماية من خلال دراسة و تقييم عدد من الأساليب المستخدمة لتنقية محتويات شبكة الإنترنت، كما يقدم البحث نموذجاً لبناء برمجيات التنقية متعددة المراحل و التي تستطيع منع الوصول الى المواقع الرديئة بطريقة فعالة.

The Information superhighway with its tools such as electronic mail, bulletin boards, chat rooms and the World Wide Web has become the ultimate source of communication and endless information exchange. This superhighway however, has been jammed with material causing serious clashes with our traditions, moral standards, and even the law. Pornography, indecency, crimes, stalking, and dangerous material are only few examples of some of the contents accessible freely to every user including our children. Until the existing legal laws sort out the confusion regarding the Internet and its contents, users are resorting to a variety of methods and tools to monitor and to filter the contents of the Internet hoping to make it a safer place for their future generations.

This paper presents a survey and an evaluation of some of the existing methods and tools used for Internet filtering. A proposal for a filtering program is being presented. This program overcomes the limitations of filtering techniques by its hierarchical organization which combines URL, word matching, and labeling approaches.

Keywords: Internet, World Wide Web, Content Filtering, URL Filtering, PICS Labels, Middle East.

* Jordan Journal of Applied Science, vol. 2 No. 1pp. 37- 49, 1999.

1. Introduction

Since its first introduction in the early nineties, the World Wide Web (WWW) has been considered as a new revolution in the information age. Unfortunately, with the lack of laws and regulations governing its contents, the WWW seemed like a fertile place attracting all forms and shapes of profit-seekers who are willing to market their products without any considerations to our moral standards . The Internet as a whole became a place which satisfies the taste of all kind of people including those with twisted minds and dangerous intentions. Thousands of sites were created which include pornographic material, sex newsgroups, and sites which teach children how to make explosives, how to burn down their schools, and even how to commit suicide[1].

A Recent study by a research group from the Carnegie Mellon University (USA) shows that there are about 450,620 pornographic images located in adult computer bulletin boards in the United States alone which have been downloaded by users 6,432,297 times [2].

Laws and regulations differ between various countries, they even differ from one state to another in one country such as the United States. Norms and traditions also are different [3]. For example what is considered acceptable in the United States may be totally rejected, and may be considered illegal in a more conservative country like Saudi Arabia. An article presented by Edwin Diamond et. Al. [2] looks at the current problems concerning indecency, pornography, hate speech and other related issues from a legal point of view. The paper shows that within the United States alone, there is a big confusion about the laws and first amendment, along with many problems that will probably take years to resolve. Local laws and regulations cannot alone prevent the existence of bad sites. A site can escape such laws by being easily moved to a different location, probably to a different county, with more lenient laws. Hyper links can be used to guide the user to the new location of the site with a simple mouse click.

In order to fight back, many schools, universities as well as public libraries are now using different methods to restrict and to control Internet usage in their facilities. They have done so despite of the many criticisms by freedom of speech activists and other organizations [4,5]. Countries such as Saudi Arabia^{*} and many other Middle Eastern countries would not take any chances and therefore have restricted the usage of the Internet to few selected government agencies and higher education institutions. This means that the public has been denied access to the Internet and will remain so until serious solutions come to existence. Some users in these countries have been willing to call long distance to get access from a nearby country that does not restrict Internet usage.

^{*}As of the writing of this paper, Saudi Arabia has launched a tightly controlled and centralized Internet services to the public. The service does not allow image viewing.

Such problems suggest that until the national and international laws are able to sort out the confusion regarding Internet contents, powerful and reliable Filtering Programs along with other monitoring techniques can be most effective.

Filtering can be accomplished in a variety of methods and can be applied effectively to eliminate pornographic material, eliminate advertisements, filtering the usenet, filtering email, ftps, chat rooms, textual materials, and other activities.

2. Approaches to filtering

Filtering the Internet can be achieved using several approaches. The most commonly used include URL Filtering, Word Matching Filtering, and Filtering using a rating system. Some filters use a combination of one or more of these approaches. Other Filtering methods include picture/image filtering and script filtering. This section presents a quick overview of the above mentioned filtering methods while highlighting some of their strengths and weaknesses.

2.1 URL Filtering

URL Filtering is an effective approach for blocking sites that are considered unsuitable for users [6,7,8]. Addresses (URL's: Uniform Resource Locators) of such sites are usually kept in a list which is searched before connecting to a site. A connection request will be denied if it's address matches any of the addresses in the list. The list of restricted addresses (Bad List) must be updated frequently to include all the new sites that come to existence. Usually companies producing filtering products must have special staff whose responsibilities are to keep searching and testing the Web for new sites to be added under the restricted list. The updated list must be downloaded periodically by customers.

Another form of URL filtering, which is considered a more restrictive approach, works by maintaining a list of allowed sites (Good List), and will limit access to only those sites included in the list. This approach is being preferred by some major companies who don't want their employees to waste time roaming around in cyberspace. Only the sites designated by the company are allowed to be visited.

URL filtering has been considered as one of the most effective tools for filtering despite some of its criticisms. The problem of maintaining and updating the list can be costly especially when it is estimated that a new site is being created over the net every 30 seconds. In many cases the users have no control over the list of restricted sites and they have to abide by the list provided by the filtering program. Many filtering programs have been attacked and even sued for restricting sites unjustly such as the case of restricting access to other competitive software products. In particular, one of the well-known filtering programs called Cybersitter from Solid Oaks was accused of denying users

from downloading their filtering program after searching the user's hard disks to find if they have visited any of the sites which criticize the Cybersitter program.

One of the best known filtering programs which works on the basis of URL filtering is CyberPatrol from the Learning Co. This program contains in its blocked list over 4800 web sites and about 250 newsgroups. Other packages include BESS from N2H2 Co., Agent Ware from Verity inc., and X-Shadow by ZDNet Products [8].

2.2 Word Matching Filtering

This approach relies on searching for and blocking inappropriate words on Web pages as they are downloaded [6,7,8]. Word matching can be applied to a variety of sources including chat rooms and e-mail. Word matching can be achieved by searching through the entire Web page or simply by searching only for keywords. In spite of the fact that Word matching approach can be very useful, it has many drawbacks. Searching through a document for specific words can be a slow process and can cause inconvenience for the user. Furthermore, word matching cannot prevent users from accessing obscene images unless the name of the image itself contains one of the prohibited words. Another major problem with word matching is the conflict on deciding what to consider as obscene and what is not, as the example of Medical terminology.

CyberSitter, from Solid Oaks Co., is a filtering program with an interesting database containing terms that can makeup bad sentences such as:
[you] [are] [a,an,too,to] [stupid,dumb,ugly,fat,idoit,dummy,...]

WizGuard is another software package which uses word matching to protect users against obscenity and sexually explicit materials. Well known Search Engines such as AltaVista and Lycos use Word Matching techniques to exclude particular words from their search results.

2.3 Using the rating system

The Massachusetts Institute of Technology has developed a set of technical standards called PICS (Platform for Internet Content Selection) [1,5]. Its concept is similar to the rating system used with movies. Sites are rated, for example, according to levels of violence and indecency, and PICS Labels are created and attached to the site. A filtering program which works on PICS labels is then used to filter sites which do not suit customers needs. Several organizations are now working very seriously with PICS standards and have furnished all sorts of free information for users to adopt PICS filtering. For example, the University of Michigan School of Information (USA) has developed a project called PICS Application Incubator which provides users with free advice and free source-code to start PICS labeling.

The problem with PICS however, is that it requires great joint efforts and cooperation of different organizations. Firstly, using a rating system such as

PICS is not mandatory and there are more existing sites which are not labeled than those with labels. Secondly, site owners have to be responsible and trustworthy enough to label themselves properly. Labels can be used to deceive users, and site contents can be changed after it has been labeled. Thirdly, the success of the labeling process requires the availability of software filtering programs which use them.

To overcome some of the problems associated with self labeling, independent labeling or rating agencies are being formed [9]. These agencies maintain independent lists on their servers containing site addresses and their labels. Filtering packages can access such servers before giving access to the user. If a site is already labeled by its owner, then its label is overridden by the label existing in the label agency. The problem with this approach, however, is that access to the label server will prolong response time unless the list of labels is stored locally on the users computer.

PICS labels presents an excellent and flexible approach to filtering, and has been recognized and used by many major software companies. This approach will become even more powerful given that the rating system is enforced on all web sites.

The Microsoft Explorer uses the PICS approach for its filtering and it allows the user to control the levels of material to be filtered. Other examples of filtering packages that use PICS labels include CyberPatrol, WizGuard 1.2, and Net Shepherd.

An example on how to self label a document is shown below. Notice that the label must be inserted in the HTML Header rather than in the body of the document. More information can be found in the W3C organization site <http://www.w3.org>.

```
<head>
  <META http-equiv="PICS-LABEL" content='
    {PICS-1.1 "http://www.gcf.org/v2.5"
      labels on "1994.11.05T08:15-0500"
        until "1995.12.31T23:59-0000"
          for "http://w3.org/PICS/OVERVIEW.html"
            ratings (suds 0.5 density 0 color/hue 1)}
  >
</head>
.....contents of document here.....
```

2.4 Other methods to filtering

2.4.1 Filtering images

There are several approaches for online filtering of pornographic images. A simple approach works by denying access to sites containing excessive number of images. The philosophy behind this filtering approach is that most bad sites contain large number of images as compared to textual material. A

simple filtering program, therefore, can easily determine if a web page contains large numbers of pictures, and if so, it can work on blocking it. Of course the apparent problem with this approach is that not all sites containing large numbers of images are bad sites, such as the case with medical, scientific, and engineering sites.

Other forms of image filtering do not block a site but work by filtering only the images having GIF or JPG extensions. Images with GIF extensions are usually used for advertisement by the web page owners. These images are too slow to load and can be extremely frustrating for users who have to wait for them. These images can also fill up the user's hard disk since they get copied to the cache area to speed up their future loading. WebFilter [10] is a program which can filter images by writing specific scripts for each web page to be filtered. It is useful for eliminating GIF images which usually contains advertisements.

A more sophisticated approach for filtering images works by analyzing the image before permitting the user to view it. The analysis is based on recognizing specific patterns in the image which resembles parts of human body, in addition to the percentage of skin color present in the image. A good example on this approach is a system called WIPE (Wavelet Image Pornography Elimination) developed at Stanford University [11]. It works by eliminating images using an algorithm which combines an icon filter, a graph photo detector, a color histogram filter, a texture filter, and a wavelet-based shape matching algorithm. A second program called IBCOW [12] was also developed at Stanford for classification of web sites based on the WIPE system. This program, however, takes about two minutes to classify if a given web site is objectionable or benign.

2.4.2 Script filtering

Script Filtering is yet another form of filtering which depends on maintaining a list of URLs for sites to be filtered. However, sites which are contained in the list do not get totally blocked but instead they get filtered. Each site in the list is attached to a script which describes its filtering scheme. Using customized scripts for each site permits the user to control exactly what to filter and what to allow from a given site. A script can be written for example to filter the annoying advertisements or offensive language from a given web page. The problem however is that a script must be written for each web page. Moreover, if a web page is changed, the script may have to be modified. To reduce the efforts of creating scripts for each web page, Script Templates can be created and stored in a library and used with minor modifications.

WebFilter [10] was one of the first packages based on Script filtering. Other packages include *Internet Fast Forward* by PrivNet Company, *Internet Junkbuster* for blocking advertisements, *ByProxy* which is written in Java, and many other packages.

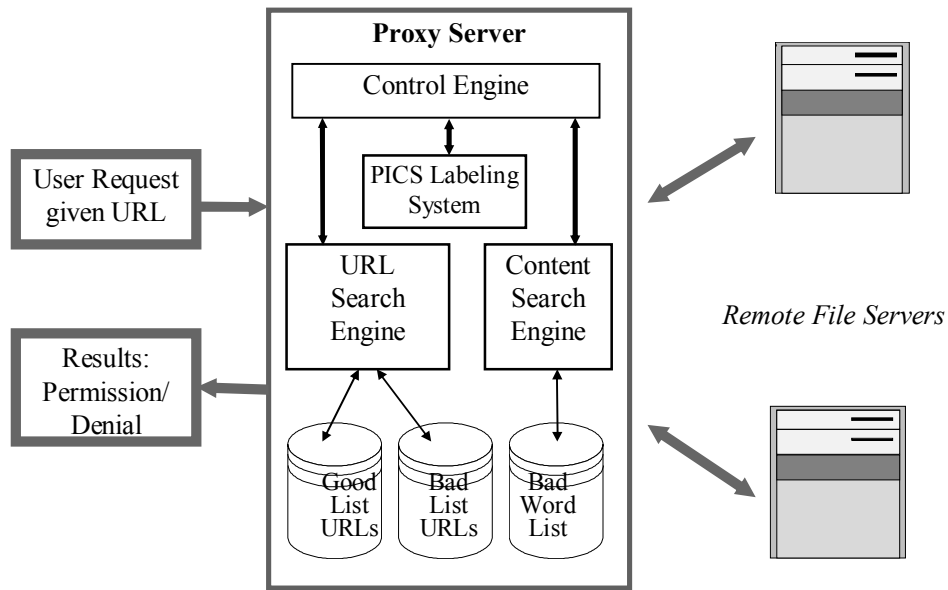


Figure 1. Filtering Program Acts as a Proxy Server

3. A model for Filtering

The filtering program presented in this paper (Fig. 1) takes advantage of PICS labeling, URL, and word matching techniques. The filtering components are organized into several stages beginning with searching for good URLs (sites known to be clean) and ending with the word matching component (Fig. 2). The program blocks access to sites which are considered unsuitable for users by acting as a proxy server which runs transparently on the user's computer. The user interacts with his/her browser by issuing a request to access a given site. However, this request does not go directly to the remote server, but instead gets received by the proxy server which performs the filtering procedure on it, and which decides whether to accept or reject that request. The filtering program is usually controlled by a super-user who can assign passwords to other users. The super-user has control over the filtering attributes and preferences, and has access to the various lists maintained by the system. These lists include the Bad-URLs list, the Good-URLs list, and the Bad-words list.

The main logic of the program is shown in Figure 3. The filtering stages are organized in such a way that rigorous filtering techniques are applied as needed with minimal amount of time possible, beginning with the least time consuming approach, to the most expensive one. Searching the Bad List (usually contains large number of site addresses) and searching for words (an inherently slow process) are performed as the last stages in the filtering process. Additional measures are taken to enhance the performance of the system with repeated use. Addresses of sites classified as Good sites by

either the PICS Parser or by the word matching component are stored back in the Good List. Similarly, addresses of sites found to be Bad by the word matching components are stored in the Bad List. This means that a future attempt to access a site which has been previously Parsed or searched would require only fetching its URL from the Good List or from the Bad List.

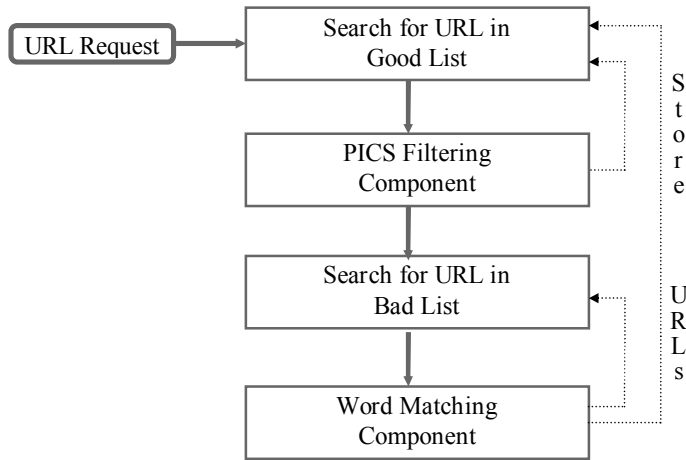


Figure 2. Multi-Stage Filter Components

3.1 Components of the Filtering Program

1. Parser for PICS Labels: a complete filtering program which works on PICS Labels.
2. Good Site List: contains a list of addresses (URLs) for sites which are considered useful, and which do not contain any controversial materials. This list can be customized by the super user (using a password). In order to keep the Good Site List as small as possible, multiple lists are created for each user of the system. The size of each list is also controlled by deleting the least frequently accessed sites.
3. Bad Site List: A large list containing hundreds of addresses (URLs) of sites which are considered not suitable for users. This is usually an encrypted list for safety, however new URLs can be added to it by the super user or by other components of the system.
4. Word list: contains a list of terms considered not suitable or proper. The list can be customized by a super user.
5. Word Matching Search Engine: Searches for the existence of a string in a site that matches any of the words in the word-list. To reduce the search time, the search begins with site title, key words, then searches the contents.

- Control Engine: Controls the sequence of operations in the system by switching between the various filtering components.

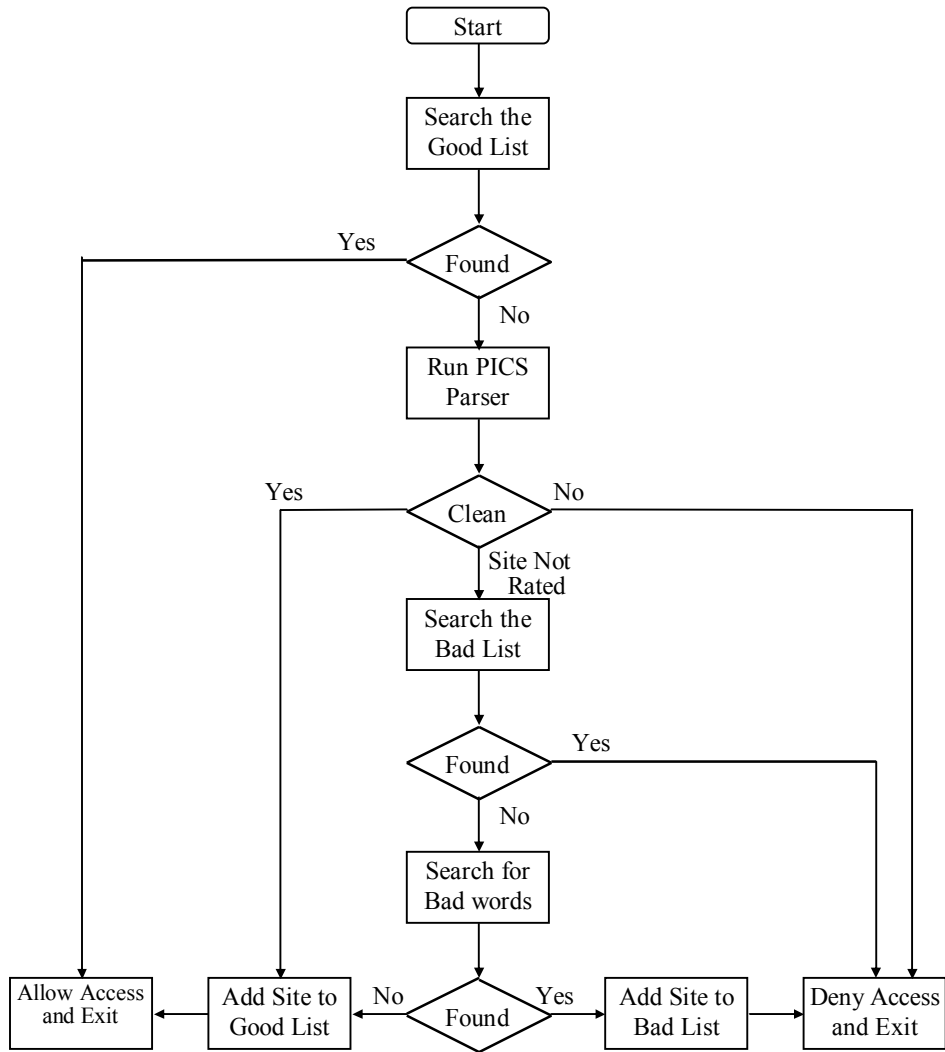


Figure 3. Main logic for Filtering Program

3.2 Operation of the Filtering program

- As a proxy server, the filtering program takes control over the local browser whenever there is a request to access a site given its URL. The program begins by searching the good list for the URL. If the URL is found, then immediate access is granted to the user. The advantage of searching the good list first, is that it is usually limited in size containing addresses of frequently visited sites which do not contain controversial material.
- If the requested URL is not found in the good list, then control is passed to the PICS Parser which checks labels of the site or the HTML document. If the site is not rated (i.e. Not labeled), then control is

passed immediately to the next stage. However, if the site was labeled, then there are three cases that can take place:

- the site is being rated as not suitable for the user, therefore access is denied.
 - the site is being rated as a suitable site for the user, therefore access is granted with no further filtering steps.
 - the site is found suitable, but user selects the option of further filtering stages to avoid non-trustworthy ratings.
3. The Bad URL List is searched only if the filtering procedure was not terminated by the PICS filter. If the requested URL is found in the list, then access is denied with proper warning to the user. However, if the URL is not found, it does not mean that the site is clean and further checking is required in the next step.
 4. If the requested URL is not found in either the good list nor in the bad list, then control is passed to the word matching procedure which searches the site for bad words using a list of bad words. The word matching procedure is a slower process comparing to URL matching since each word in the list must be searched for, until a match occurs. For such a reason this procedure is being used only if the URL search fails. Additionally, the word matching search attempts to save time by beginning the search starting first with the site title, keywords, and finally with the contents. If the requested site is found to be clean then its URL is added to the Good Site List, otherwise it is added to the Bad Site List.
 5. System maintenance and list updates: the Filtering program can be maintained by a super user whose duty would be to maintain the users lists. The program includes a single list for bad sites and another list of bad words. The list of bad sites must be updated frequently so that it will include new sites that come to existence. Such lists can be imported from several locations over the Internet and appended to the existing list. The list of bad words can also be modified by the super user who can add new words or delete existing ones. Each user on the other hand will have his/her own list of good sites to keep its size small and to reduce the search time. The super user can check the good site list and can modify it if necessary.

4. Conclusions and Recommendations

Providing Internet access to the public has become a priority task for many developed as well as developing countries. However, as soon as the WWW, an Internet service allowing multimedia transmission, was launched in the early nineties, thousands of sites containing controversial material were created, thus making the Internet a confusing and a dangerous place for many conservative countries, children education systems, public libraries as well as parents who fear the worst for their children. The lack of laws governing Internet contents, the differences in moral standards, clashes with freedom of expression and many other related issues have even made this problem a more complicated one. Fortunately, the Internet can be made much safer using different tools of monitoring, filtering, and other precautionary procedures.

This paper has tackled the issue of filtering as a means of controlling Internet access. However, each of the filtering methods discussed did have one or more limitations, indicating that no one filtering technique can guarantee perfect results alone. Limitations, however, can be overcome using well-designed software which must satisfy at least two objectives; effectiveness and performance. Effectiveness (how well a filtering program works on preventing access to bad sites) can be achieved by combining different filtering techniques such as PICS, URL matching, or word matching to produce what we refer to as Multi-Stage Filtering. Each stage in this case would play an important role in the overall filtering process.

In the case of performance of a filtering program, two issues are considered. First, the ability to perform the filtering tasks with minimal time and with little delay to the access time required to connect to a given site. Hierarchical organization of filtering modules can satisfy this objective. The filtering system begins with the least time-consuming filtering procedure which is to be located at the top of the hierarchy (searching the good list of URLs in our case), and ends with the most time-consuming procedure at the bottom (using the word matching procedure). Filters at the lower levels are accessed only if the top level procedures fail to perform the filtering task. The second issue is that a filtering program must be capable of enhancing its performance with repeated use. This can be accomplished through dynamic update of the lists containing both bad and good URLs. For example, if the filtering program activates the word matching module which goes through the slow process of searching for bad words in a given site then the URL of that site is to be stored back in either the good or the bad list of URLs depending on the outcome of the search. In this way, if the same document is to be accessed in the future, it would require searching the list of URLs instead of searching for words, thus saving a great deal of time. Similarly, if the PICS parser determines that a document is suitable for the user, the document URL is stored in the good list, thus requiring no future parsing for the same document.

It is important to realize, however, that Internet control can be achieved using several approaches which includes filtering amongst them. The best result can

be achieved by establishing a well-defined strategy which combines one or more of the following methods:

Summary of Methods to control Internet Access

1. Restrict Internet access to only selected people in government and higher education institutions.
2. Have users sign agreements (Acceptable Use Policy) which prohibits them from accessing unlawful material and apply strict penalties to those who violates such agreements.
3. Use monitoring programs to track down violators.
4. Use filtering programs and apply it to global proxy servers or FireWalls.
5. Use filtering programs and apply it locally to prevent children and irresponsible users from accessing unlawful materials.
6. Provide Internet access to children using Menu-Driven programs with pre selected topics.
7. Allow access to Internet only through public places which can be controlled using steps 2 to 6.
8. Adult supervision is necessary. Children are to use the Internet only during some specified hours in the day when an adult is available for supervision. The adult's password must be used to get Internet access.
9. Establish a rigorous educational system for Internet users including parents, children as well as public and private institutions.
10. Countries must work seriously in establishing sound communication laws to eliminate the confusion between legality and freedom. Internet users must be aware of such laws.

Bibliography

- 1] Paul Resnick, (March 1997) "Filtering Information on the Internet", **Scientific American**, pp. 106-108
- 2] Edwin Diamond and Stephen Bates, (Oct 1995), "Law and Order Comes to Cyberspace", **Technology Review**.
- 3] Weinberg, Jonathan, "Rating the Net," **Hasting Communications and Entertainment Law Journal**, vol. 19, No.2, pp. 453-482.
- [4] Cisetr Steve, (Oct. 1993) "Protection and the Internet," **Apple Library**, sac@apple.com
- [5] LaRue, James, (1997) "Internet Freedoms and Filters: Roles and Responsibilities of the Public Librarian on WWW," web site: www.sni/~jlarue/
- 6] Belkin, Nicholas . J; Croft, W. Bruce, (Dec1992), "Information Filtering and Information Retrieval: Two sides of the Same Coin," **Communication of the ACM**, vol 35, No. 12, pp. 29-38.
- 7] Stevens, Curt, (Dec 1992), "Automating the Creation of Information Filters," **Communication of The ACM**, vol 35, No. 12.
- 8] Railsback, Kevin, (1997), "Policing the Net," **ZD Internet magazine**.
- 9] Pual Resnick and Jim Miller, (1996), "PICS: Internet Access Controls Without Censorship," **Communication of the ACM vol. 39 No. 10**, pp.87-93.
- 10] Boldt, Axel, "Filtering the Web using WebFilter", Web site: <http://math-www.uni-paderborn.de/~axel>
- 11] James Z. Wang, Jia Li, Gio Widerhold, Oscar Firschein, (1998), "System for Screening Objectionable Images," **Computer Communications**, vol. 21, No. 15, pp. 1355-1360.
- 12] James Z. Wang, Jia Li, Gio Widerhold, Oscar Firschein, (1998), "Classifying Objectionable Websites Based on Image Content, " Sanford University, Stanford, CA 94305
Web site: <http://www-db.stanford.edu/~wangz/project/imscreen/JCC98/>