PERGAMON

Telematics and Informatics 20 (2003) 71–95

TELEMATICS AND INFORMATICS

www.elsevier.com/locate/tele

Policies for content filtering in educational networks: the case of Greece

M. Avgoulea, C. Bouras *, M. Paraskevas, G. Stathakopoulos

Department of Computer Engineering and Informatics, Computer Technology Institute, University of Patras, Riga Feraiou 61, 26221 Patras, Greece

Abstract

The Internet removes all barriers to sending and receiving information. An increasing number of nations connect their schools on the Internet as an acknowledgment of its importance in education. Despite its undeniable usefulness, it also has certain perils. This study specifies and evaluates these, finding the technologies that are currently available to address filtering issues and comparing them. We present our proposed solution for the Greek School Network, and illustrate to what extent our solution successfully addresses the issues discussed.

Keywords: Content filtering; Educational networks; Internet; Policies for safe Internet

1. Introduction

The Internet is a powerful tool that can be used for training, research, business or mere fun. The well known advantages of the Internet distinguish this mode of communication from most others: the worldwide network of computer networks, has opened up vast possibilities for sending and receiving information in an efficient and effective manner, allowing rapid search and retrieval, and also enabling users to reach an unlimited audience.

However, questions arise when schools or other public networks offer connectivity to students or users of public libraries regarding the accessible content on the Internet. Without meaning that we should implement extreme measures of censorship or limitations on the content we allow our users to access, we should be alert and monitor what happens while we are online.

* Corresponding author. Tel.: +30-61-960375; fax: +30-61-996314.
E-mail address: bouras@cti.gr (C. Bouras).

0736-5853/02/$ - see front matter
PII: S0736-5853(02)00004-7

The rationale behind this is not a veiled prudishness or oversensitivity: schools' trustworthiness is endangered if the school network is used (systematically or occasionally) for purposes other than those for which it was designed and implemented. In addition, parents and the public believe their children should be protected from illegal, offensive and inappropriate content. The issue becomes increasingly important as the problem with Internet content seems to have social, cultural, pedagogic, scientific and other viewpoints. Society should be able to trust people who work for their children's education, and their national government to provide the best possible education to the students. While some may argue that it is a means for censorship and nothing more, others believe it is the only way in which societies can inform of and protect children from the potential danger the Internet poses. If the fundamental institution responsible for the children's education cannot provide means of protection against material that can be described as harmful or improper, then we enter dangerous territory.

It is obvious that we cannot ignore the problem. Every country should initiate programs and establish its own policies to deal with it. In fact many countries have already done that: Canada and Australia initiated public discussion on the content issue in conferences and in public to find the most effective ways to tackle the illegal and offensive content on Internet without raising public concern regarding human rights and an individual's freedoms.

The basic idea of the proposed solution for the Greek School Network is the provision of a safe environment in which kids can surf, to find useful educational links with information that will improve their knowledge, cultural and educational base without the anxiety of improper, sexually implicit content. This document examines all the potential dangers the Internet can pose and the possible solutions to them.

From our point of view, it is an unarguable fact that the Internet offers a vast mass of information, some of which is unsuitable for schools. One point that should not be overlooked is that education can help students determine which of that knowledge can be useful or interesting and which is illegal or improper for them. Schools should accept their responsibility and start certain initiatives to achieve these goals. The best solution is one that combines the right guidance of the students from the educational authority, the informing of parents and the training of educators, together with a technical solution. No technical solution alone can resolve such a problem, and whichever policies are established should take into account the above parameters.

The rest of the paper describes the problem in detail, looks at how other countries deal with the content issue and the technologies available to solve the problem, examines the solution we propose for the Greek School Network, and demonstrates how that solution will be implemented. Conclusions are then presented, and future work discussed.

2. Problem description

In the early 1990s, as the number of subscribers to proprietary online systems grew, transforming those communities from the moral equivalent of small towns into

large cities or even states, children began to meet people in online chat rooms who would engage in inappropriate conversations or encourage them to divulge information about themselves.

As the number of computers in schools and the number of children accessing the Internet from the classroom has grown exponentially over the past few years (there is some indication that Internet growth today has gone from exponential to linear), so too have the challenges facing educators trying to ensure that children have a positive experience when they go online. Today it is an indubitable fact that every child can easily gain access to objectionable material, as the Internet is not content zoned and minors are not prohibited from accessing adult material.

If the educational community ignores this problem, it will be like admitting its inability to deal with it or its indifference for the overspending of governmental money and the use of its network resources (the school network itself, but also its hardware and services) for purposes irrelevant to every possible educational goal. What is more important is that it renounces its role as a pedagogic institution at that critical issue which challenges the most technologically advanced countries and societies.

The realization of the importance of the problem and its settlement on the school environment, will increase the citizens' awareness and will be a good stimulus for Internet Service Providers worldwide to develop such technologies.

A related debate rages over what percentage of Web sites would truly be considered objectionable. Some advocates argue that sites that would be considered harmful to minors represent only a very small proportion of the Web. What is of greater concern, they say, is that perfectly benign and possibly very useful information could be blocked when software is used to screen inappropriate material. The actual extent to which adult-oriented material is available on the Internet is irrelevant, according to those who support government-mandated content controls. They believe that any amount of inappropriate content is too much when children are concerned. That point gets more significant regarding the nature of the school environment and the role it has played over the past decades, and more importantly the role it should play in the years to come, with the rapid increase of technology penetration in the classroom.

Pornography is not the only issue involved. Many adults are concerned about Web sites created by hate groups or devoted to topics such as bomb making and weaponry, gambling, alcohol or smoking. Although pornography on the Internet has captured the greatest attention on the part of policy-makers, it is not the only area of potential concern.

Some experts argue that the Web is expanding so quickly that it is virtually impossible to track every site that could be objectionable. The flip side of that argument is that it is better to minimize access to objectionable content as best we can, even if the occasional site slips through the cracks.

The Internet truly is like a vast library including millions of readily available and indexed publications, containing content as diverse as human thought. But in the Internet community, just as in any city frequented by millions of people, there are neighborhoods that are inappropriate for children to visit alone and strangers they

would be better off not meeting. Throughout the past decade, policy-makers, industry advocates, parents and teachers have tried to address these concerns, especially in the more technologically advanced areas of the earth. This decade may well be the decade of decisions in a much broader spectrum than ever before.

No matter which approach a school or school district decides to pursue, it should adopt an Acceptable Use Policy to govern the use of its network and computers. School officials should also take concrete steps to teach their students the 'rules of the road' when they go online, and how to evaluate the quality of online information.

The World Wide Web, however, is not the only source of concern. Children can receive email messages with pornographic file attachments or e-mails from Usenet groups, which communicate through an older Internet protocol, can contain postings from users that would be considered inappropriate. Of special concern, too, are Internet chat rooms and so-called Instant Messaging, where children can communicate online in real-time with adult strangers who may not have their best interests at heart.

A wealth of information is available on the Internet: groups' or individuals' opinions, governmental decisions, companies' advertisements on their software or hardware solution, researchers studies, etc. A number of white papers [1] are opposed to the use of censor ware programs—as they call it—in American libraries and schools, and describe their objections to the content filtering solution and their perception of the problem. Explanatory websites [2,3] provide links to additional information, and try to explain what filtering is. Their pages almost always contain the advantages and disadvantages of filtering, while companies try to explain why it is 'imperative' for parents or schools to buy their product. Some interesting links that are indicative of the ongoing debate are: the Joint Statement for the Record on "Kids and the Internet: The Promise and the Perils", [4] stating why libraries should not implement filtering policies, "Is Cyberspace Burning?" [5] from the American Civil Liberties Union, Communications White Paper [6] and Summary of the Internet Family Empowerment White Paper. [7]

[1] EFF NRC Censorware White Paper #1: http://www.eff.org/Censorship/Censorware/20010306_eff_nrc_paper1.html, http://www.eff.org/Censorship/Censorware/20010306_eff_nrc_paper2. html

[2] EFF Topics: Internet Blocking (Censorware)—http://www.eff.org/Censorship/Censorware

[3] Computer Professionals for Social Responsibility (Filtering FAQ)—http://www.eff.org/Censorship/Censorware/filtering_faq.html

[4] Joint Statement for the Record on "Kids and the Internet: The Promise and the Perils"—http://www.eff.org/Censorship/Censorware/19981214_ifea_nclis.statement

[5] Is Cyberspace Burning? How Rating and Blocking Proposals May Torch Free Speech on the Internet—http://www.aclu.org/issues/cyber/burning.html

[6] Communications White Paper—http://www.communicationswhitepaper.gov.uk/by_chapter/ch6/6_10.htm

[7] Summary of the Internet Family Empowerment White Paper—How Filtering Tools Enable Responsible Parents to Protect Their Children Online—http://www.cdt.org/speech/970716summary.html

3. Status in the other countries

Many countries seem to realize the potential problem in the use of school networks for Internet access. Each of the countries presented here establishes it's own policies and methods. However, their perception of the problem presents some surprisingly common characteristics. The countries surveyed are Australia, Canada, United States of America and the European Union. In our survey we browsed some interesting and concise pages regarding the pronounced decisions on the Internet content issue in Australia, [8] Canada, [9] the United States of America [10] and the European Union. [11]

3.1. Australia

The Internet is increasingly being used in Australian schools as a learning driving force in education. Forty-three percent of the children asked in Australia (from age 9–14) said that the Internet improved their perception about school.

The federal government of Australia responded to the issues posed by the entrance of the Internet into day-to-day life with the Broadcasting Services Amendment Act (1999), designed to protect citizens from illegal and offensive content. For the same purpose, a hot line was created to allow people to report what they think is illegal and at the same time express their worries about the content that can be accessed on the Internet. The industry plays an important role by establishing policies for the improvement of information exchange regarding web page content.

Some states have established policies which are mandatory for schools to implement. Acceptable Use Policies comply with certain standards for the publication of web pages, whilst in other districts general instructions and technical alternatives are given to help schools establish their own policy without any kind of enforcement.

Education authorities (governmental or non-governmental) stress the importance of accessing pages of a high quality. Educators at a national level do the evaluation and the pages are selected to meet certain specifications. In some cases, schools store pages certified as 'safe quality pages' in their cache, ensuring that their students have quick access to them. A large selection of web pages examined for their suitability in education with many searchable options on the selection criteria is offered. While these pages are available for the benefit of students reducing the bandwidth consumption, the cost for schools remains low because of the quicker access time to

[8] Using the Internet: a Positive Learning Experience (A Position Paper on Safe and Appropriate Internet Usage in Australia. EdNA Schools Advisory Group)—http://www.edna.edu.au/publications/use_internet/report.html

[9] Illegal and Offensive Content on the Internet: The Canadian Strategy to Promote Safe, Wise and Responsible Internet Use—http://www.connect.gc.ca/cyberwise/

[10] Congress Passes Filtering Mandates For Schools And Libraries—http://www.cdt.org/publications/pp_6.22.shtml

[11] European Union's Action Plan on promoting safer use of the Internet—http://europa.eu.int/ISPO/iap/decision/en.html

stored sites. In Australia decisions are made at the school level, despite the fact that in some states there is a trend for more obligatory policies decided at a higher level of authority. Recent studies point out that almost 98% of public schools connected to the Internet have established Acceptable Use Policies. Seventy-four percent of schools with AUP use software to block or filter pages with certain content, 64% uses rules of conduct and 28% uses an intranet in their effort to control access.

3.2. Canada

To-date, Canadians have established a wide range of partnerships to deal with issues of Internet content, bringing together various levels of government, law enforcement agencies, the private sector, not-for-profit organizations and the community at large.

The strategy of Canada's Government is to make Canada the most connected country in the world. The six pillars of the Connecting Canadians initiative are: Government On-Line, Canada On-Line, Canadian Content On-Line, Smart Communities, Electronic Commerce, Connecting Canada to the World.

Illegal content—content that violates Canada's laws—is of key concern. Child pornography and hate propaganda are particularly troubling, because they pose the greatest and most immediate risk to the safety and well being of Canadians. Enforcing the law in cyberspace, however, presents significant challenges, particularly in view of rapid technological change.

For Canadians it is important to distinguish between Internet content that is illegal, and content that is offensive to some, but is not illegal. Canadians think that control of illegal content is fundamentally an issue of law enforcement. The control and management of offensive content, however, calls for different approaches, such as empowering users, educating consumers to make informed choices, and establishing responsible industry practices.

The *Canadian Charter of Rights and Freedoms* guarantees all persons in Canada "freedom of thought, belief, opinion and expression, including freedom of the press and other media of communication". The federal government's approach is to involve a broad spectrum of Canadians in addressing the issues. Its priorities include supporting initiatives that educate and empower users, promoting effective industry self-regulation, strengthening the enforcement of laws in cyberspace, implementing hotlines and complaint reporting systems and fostering consultation between the public and private sectors, and their counterparts in other countries.

Some of the most important initiatives, studies and discussion in progress in Canada are presented below.

In 1994, the Government of Canada established the private sector Information Highway Advisory Council (IHAC) to provide advice on the best way to develop Canada's Information Highway. In its 1995 and 1997 reports, the Council focussed on the need to build awareness and educate Canadians about offensive content on the Internet, and also on the importance of working with industry, both to promote voluntary industry guidelines and to encourage research into filtering software. According to its recommendations, the federal government should: (a) fine-tune

existing laws to make them more applicable and enforceable in the changing world of global networks, and (b) encourage research and the development of technical tools that can protect users against offensive content on the Internet, and assist in law enforcement.

The Canadian Radio-television and Telecommunications Commission (CRTC) held an extensive public hearing on issues related to new media and the Internet, and came to essentially the same conclusions as IHAC in its May 1999 *Report on New Media*.

MNet's national, bilingual, Internet education program, *Web Awareness: Knowing the Issues*, is designed for parents, teachers, public librarians and community leaders. The Canadian Home and School Federation, the Canadian Teachers' Federation and the Canadian Association of Principals have endorsed MNet's programs.

A growing number of Canada's leading broadcast, cable, telecommunications and new media companies provide MNet with financial support, and several federal and provincial government departments are providing further funding.

The SchoolNet National Advisory Board, established by Industry Canada, produced a brochure which includes options available to address social issues, such as the appropriateness of online content, and responsible Internet use to assist educators who are introducing the Internet in the classroom.

Missing, an educational kit that teaches children how to surf the Internet safely and warns about predators who use the Internet to lure children into sexual encounters, has been distributed free-of-charge to 10,000 schools and libraries across Canada. The kit includes a computer game, a video documentary, a Web site and a guide for parents and teachers, and was sponsored by the federal and provincial governments, the private sector and high technology companies.

The Industry Canada study, *Content Filtering Technologies and Internet Service Providers: Enabling User Choice,* focuses on technologies that Internet service providers (ISPs) can put in place and Internet users can use. These technologies include child-friendly search engines and Web sites, as well as ISP-based filtering services and Web content labelling systems.

3.3. United States of America

In December 2000, the United States Congress passed legislation requiring Internet blocking technology to block pornographic materials in all public schools and libraries funded through certain federal programs.

To date the issue of content control in the United States of America is at a critical point: while the congress decided the mandatory implementation of filtering in schools and libraries, parents, organizations and politicians are steadily opposed to the legislation passed for the importation and use of that technology in public places and public services. Internet content filtering in schools and libraries is currently a highly controversial issue. Influential organizations like the American Civil Liberties Union (ACLU) and Computer Professionals for Social Responsibility (CPSR) are hostile to it, as is the American Library Association (ALA). Government bills

mandating filtering in schools and libraries face legal challenges on constitutional issues, and some libraries have already been sued for installing filtering software onto their computers. Parents, schools and libraries face difficulties to decide whether, and how, to filter Internet content (according to the report *Internet Content Filtering—Issues Facing Parents, Schools and Libraries* [12]).

Still, no one denies the fact that the Web is not content zoned, which means that children can access anything on it very quickly and easily. The main issue that parents are worried about is access to hard-core and child pornography, but there are other controversial issues, such as access to images of extreme violence.

The US Congress approved the mandatory import and use of filtering software against illegal and offensive content in all schools and libraries that receive federal financing. The *Children's Internet Protection Act—CIPA* passed the Senate and the Congress in December 2000 as part of a big government budget for 2001. Not later than 18 months after the date of the enactment of this Act, the National Telecommunications and Information Administration will initiate a notice and comment proceeding for purposes of:

- Evaluating whether or not currently available technology protection measures, including commercial Internet blocking and filtering software, adequately addresses the needs of educational institutions;
- Making recommendations on how to foster the development of measures that meet such needs; and
- Evaluating the development and effectiveness of local Internet safety policies that are currently in operation after community input. No funds made available under this title to a local educational agency for an elementary or secondary school that does not receive services at discount rates may be used to purchase computers used to access the Internet, or to pay for direct costs associated with accessing the Internet for such school, unless the school, school board, local educational agency, or other authority with responsibility for administration of such school both:
- Has in place a policy of Internet safety for minors that includes the operation of a technology protection measure with respect to any of its computers with Internet access that protects against access through such computers to visual depictions that are obscene, child pornography or harmful to minors; and
- Has in place a policy of Internet safety that includes the operation of a technology protection measure with respect to any of its computers with Internet access.

In the USA three factors justify the distinction of the rights of minors from those of adults: (1) the vulnerability of children, (2) their presumed inability to make critical decisions in an informed, mature manner; and (3) the significance of the parental role in child rearing. As for libraries, while school libraries have special

[12] Internet Content Filtering—Issues Facing Parents, Schools and Libraries http://www.safetyed.org/help/filtering.html

characteristics, public libraries are intended for free willing inquiry and access control is a more complicated issue.

3.4. European union

The European Union has published an Action Plan on promoting safer use of the Internet. While it recognises the positive benefits of the Internet (particularly in education), it states that the amount of harmful and illegal content carried over the Internet, while limited, could adversely affect the establishment of the necessary favourable environment for initiatives and undertakings to flourish.

A safer environment should be provided by combating illegal use of the technical possibilities of the Internet, in particular for offences against children and trafficking in human beings or for the dissemination of racist and xenophobic ideas, ensuring that consumers make full use of the Internet.

Europe should promote industry self-regulation and content-monitoring schemes, the development of filtering tools and rating systems by the industry and of course international cooperation. Cooperation from the industry in setting up voluntary systems of self-regulation can efficiently help to limit the flow of illegal content on the Internet. European coordination of representative and self-regulating bodies is essential for the Europe-wide effectiveness of such systems. While any hot-line reporting mechanisms should support and promote measures taken by the Member States, duplication of work should be avoided; where this is possible, hot-line reporting mechanisms could be established in cooperation with the law-enforcement authorities of the Member States, whereas the responsibility for prosecuting and punishing those responsible for illegal content should remain with the national law-enforcement authorities. This Action Plan should be of four years duration in order to allow sufficient time for actions to be implemented to achieve the objectives set.

For the members of the European union a quick search on the Internet produced the following results regarding the public awareness on the dangers that exist on Internet (indicative information produced in 1999): [13] Austria, Denmark, Italy, Luxembourg, Finland, Sweden, Spain and Greece have done little. For Greece it is mentioned: ''They plan to include a website and hotline but these have not yet been established''. Countries with initiatives and more activity on that matter are Belgium, Germany, Ireland, and England: The University of Namur has developed the MAPI—Movement Against Pedophilia on the Internet project, and the Belgian Police runs a hot-line, but this does not contain specific Internet safety awareness messages. ECO and FSM are industry initiatives in Germany (providing little safety information on their web sites). Some legal information for the public is given on two German Government sites. The Dept. of Justice Report Illegal and Harmful use of the Internet recommended the development of ''appropriate awareness measures as critical'' however there has not been a national campaign as yet in Ireland. Numbers of different Government-funded organisations include safety as part of their Internet Awareness program in England. However, there appears to be little co-ordination.

[13] What's happening across Europe? http://www.netaware.org/gb/website.html

One of the existing initiatives in the UK is Internet Watch Foundation (IWF), and this represents a leading example. Established in 1996, IWF is an independent body that works with law enforcement agencies and ISPs in the United Kingdom to remove illegal material from the Internet (particularly child pornography), and promotes the labelling and filtering of legal material that some may find offensive.

3.5. Summary

From the countries surveyed, Canada first began to deal with the 'Internet content problem', and seems to have gone far in its confrontation with it. People and organizations involved understand that a great deal of work remains to be done. For Canada the problem seems to have high priority, as the country wants to become the first fully connected country of the world, and will probably succeed.

In the US the Congress passed legislation regarding the mandatory use of filtering technologies, but at the time of the writing, there is still a strong movement against the decisions of the government. Some believe that the government decision for mandatory use of filtering policies in schools and libraries is not yet certain to be implemented.

In Australia there is enough concern on the problem and many initiatives are in progress. In fact in many schools already apply certain policies for the content issue. The European Union has not ignored the problem, with early discussions about the potential dangers on the Internet. The members of the EU, however, are not all at the same level, but that is natural as the use of technology is not widespread in all of them.

Despite their differences, all countries seem to be in favour of the Internet content hotlines: these are communications systems for receiving, processing, verifying, evaluating and acting upon complaints about Internet content. Many believe that hotlines can become the "Crime Stoppers" of the Internet. They are a link between the Internet user, the content owner/provider and law enforcement organizations. In a recent survey, researchers found that the majority of those who had wanted to complain did not do so because they did not know where to file their complaints (hotlines did not exist or the public was not aware of their existence). In addition to the UK, countries such as Australia, Austria, France, Germany, Ireland, the Netherlands, Norway and the US are operating Internet content hotlines today.

All these initiatives are suggestive of the extent of governments' attempts to inform the public, and of their quest for applicable, efficient solutions which meet the public acceptance.

A critical factor will be acquisition of public consensus and the approval of the designed actions, to avoid generation of other problems or reactions.

4. Techniques and solutions

Before referring on the solutions developed so far to confront the content filtering/blocking problem, we briefly refer to the techniques that can be used. These are the

use of keyword blocking, the use of negative and positive lists and the use of content labeling and rating systems.

Keyword blocking prohibits access to pages that contain the specified words (words can be tobacco, wine, drug, sex etc). It's use results in the exclusion of pages that should not be blocked and the opposite, unless this is a 'clever' software (rarely this is the case). This technique is inefficient and cannot successfully address the modern issues as it blocks any page that contains the word, even if it is written in medical or other context. If for example one wishes to block pages related with sex, he should write the word sex in all the existing languages to achieve true exclusion of all the pages that contain that word. The method is useless if the web developer adds an additional character, to words he suspects that can be searched from such programs in order to avoid blocking. As a result it can be easily bypassed.

The use of negative and positive lists is easy either from filters or browsers. The positive lists contain URLs and domains to which access is allowed to and negative lists contain the URLs that access is prohibited to. Use of white lists will be very restricting regarding the amount of available pages on the net.

Labeling and rating systems provide a way to categorize pages according to their content and provide that categorization to the user. He will have to decide what he wants to be blocked and what not. Its main problem is the size of the Internet and the rate of its daily expansion. That technique can prove to be inefficient if nobody offers to undertake the task of labeling. Today there are organizations that do that work and there are also forms that allow web developers self-rate the content they publish. While rating protects the free speech rights it will be inadequate until many people join the project.

Only by recognizing the defects of each technique, we can realize the difficulty of such a task and the limitations every solution has. The best possible solution will be the one that will incorporate and combine all the possible techniques eliminating their drawbacks and maximizing their positive parts.

The different solutions for access control/content filtering are Commercial Software, Rating Systems, Freeware Programs and Hardware Solutions. All of them are mainly structured on the use of lists. The way content control will be achieved is examination of the URL from which the content derives. There are three different approaches: (1) The proxy-cache server does the filtering on its own, if filtering and blocking requests are not massive and if the server provides such a possibility; (2) Plug-ins do the content control. That approach offers much more options and at the most cases there are ready site lists grouped by their content. These lists are updated frequently; (3) The last category includes the products that can operate as independent servers used only for filtering. By that way they can be used separately, especially when they are not used for caching, whereas the site lists are provided and updated automatically.

Many commercial filtering programs advertise that can block access to the harmful or improper pages on the Internet. In fact they claim they can do much more: monitor each user's activity and maintain detailed log files of all activity and violations (some offer the option of sending the log files to a specified supervisor person), control (block or allow) the use of programs, Usenet groups or news, specify

allowable times to access the Internet, block access to certain ports, allow different configuration for different users/groups of users, etc. Some of them can be installed at the server side (SmartFilter, [14] Bess, [15] WebSense, [16]) some are for clients only (NetNanny, [17] CyberSitter [18]) while some offer versions for home and for education (CyberPatrol [19]). This is not an exhaustive list and has no purpose to advertise or recommend the mentioned software.

The problem with commercial software is that the companies usually do not reveal the pages they block access to. What is more, in the EFF NRC Censorware white paper, is mentioned that: "No public documentation is offered of the blacklisting of privacy, anonymity and translation sites. Censorware hides blacklists in black boxes. All privacy and anonymity services, all websites that let a user receive material via en encrypted or private form, represent a threat to that control and are included in the blacklists of almost every category". Certain incidents of underblocking/overblocking allow to those against the use of filtering software (they call it Censorware), to claim that its use is inadequate.

This solution is not favorable for a school environment because of all the arguments mentioned above and one equally strong or stronger: the filtering is done by a third and not by the government or an educational authority. This gives to the company that implements the filtering, access to information only the authorized parties should have. That could be a major problem.

Definition of content with the use of Content Labeling and Rating systems consists of the attachment of a set of tags to each document or page on the Internet which specifies/describes the kind of the information displayed on it. Organizations that provide ratings for web sites are ICRA (Internet Content Rating Association), SafeSurf and the ESRB (Entertainment Software Rating Board). These organisations describe their activities as follows: ICRA's aim is to protect children from potentially harmful material on the Internet, SafeSurf defines its aim to supporting parents and ESRB defines its role as providing parents and consumers with objective information so that they can make informed decisions regarding computer and video games.

While the SafeSurf is the first such system ever implemented it seems that today the most famous rating system is the ICRAfilter (ICRA advertises in its homepage that the most visited Internet sites in the US have already adopted the ICRA labeling system to protect children surfing the Internet). We will in short describe ICRAfilter's characteristics, which in the most of the cases are identical with the other system's characteristics.

ICRAfilter is a browser-independent tool which is intended to provide parents with a means to filter their children's internet access according to ICRA labels, their

[14] SmartFilter Homepage—http://www.securecomputing.com/index.cfm?skey = 85
[15] Bess Homepage—http://www.n2h2.com/
[16] Internet Filtering by websense—http://www.websense.com/ index2.cfm
[17] NetNanny 4 Homepage—http://www.netnanny.com/home/home.asp
[18] CyberSitter Homepage—http://www.cybersitter.com/
[19] Cyber Patrol—Internet Filter Software for Home, Education,—http://www.cyberpatrol.com/

own "block" and "allow" lists or third party lists of web sites. It will be able to install on any PC running Windows 9x upwards. The real power of ICRA*filter* is that it also supports lists of web sites created by other organizations and companies extending the width of its solution. ICRA is talking to a number of potential list creators and there is special provision to have several lists available. The next associated event is programmed for March 21st when as well as releasing ICRA*filter*, ICRA will shortly be publishing help for webmasters responsible for large sites. One of the limitations of PICS labels is that they need to declare the domains they cover within the label (for example to label http://www.domain.com and http://subdomain.domain.com requires *two* labels). For large sites running on dedicated servers it will be possible to configure the server to include ICRA labels in the HTTP headers automatically. This means that extra work will not be necessary.

Today it seems that there is a lot of development under way and people work on the project. Despite that, everyone should remember that this method might proved to be of little efficiency if there is not someone who will undertake the task of labeling and rating of sites in Internet in a continual basis. The advantages of this method is that it protects the freedom of speech as it does not block access to sites but just categorizes and labels them and offers to the parents, teachers or the users themselves the decision of the pages to be blocked. The controversy of gaining access to sensitive data does not apply on that case and at the same time the user does not have to install additional software because the filter is embedded in the browser program.

In the Internet world, the community of freeware programs (GPL) has a strong presence. Some of its most well known programs for content filtering are SquidGuard, Squirm and Jesred. To speak precisely, these are redirectors used by squid to accomplish content filtering. Squid is a highly configurable proxy and cache server freeware program. Further information on that program can be found at the Implementation Issues paragraph, where we will explain the reasons why it is the preferred solution and the configuration options it offers. Other free programs are Netscape Proxy Server and Microsoft Proxy Server.

The use of proxies has a handicap or better a policy problem: proxy servers keep log files. These files contain all the requests a proxy receives, causing complaints and worries regarding the people that gain access to that data and the processes that can be put in place to monitor individual users' behavior. When the proxy is not configured as transparent the privacy problem mentioned previously, is eliminated since the activation of the proxy cache filtering service is activated by the user himself.

Some points that complicate the administrator's duties are the absence of a user-friendly administrative interface, as all the necessary configuration and modification is done by editing the configuration file. Moreover, some of the operation parameters can change only be recompiling the source code, which causes an administrative cost. Because the program is given free of charge, there is no company which officially supports the product, although a great number of resources can be found on the Net relevant with squid.

Hardware solutions appear to be more complete, with specialized systems of black box type. They are intended as big scale solutions for huge users' databases and lines of high capacity. They require minimal efforts to be installed, configured and operate

while full support is offered from their company. Along with that comes the usually high cost to obtain such a system.

We will refer to the case of NetCache. The complete device that supports the proxy service from the Network Appliance Company, started initially only as software. This was in fact based on one of the first proxies ever implemented for a number of platforms. At its current form, it consists of a device that uses its own file system, the WAFL (operating system microcode), embedded RAID disks system and special algorithms for caching. The installation process is relatively simple and directly oriented on the proxy service, in contrast with the servers based only on software that require installation of the computer, its operating system, the proxy server and after that the fine-tuning of the system. The administration of the device is done with the use of web pages and cgi forms, while upgrade of software can be done remotely. The device can be telneted, can send notification e-mail while it offers MIB for monitoring its operation through the SNMP protocol. The caching protocols include HTTP (1.0 and 1.1 with persistent connections), NNTP ___ SSL Tunnelling. It supports transparent caching, the ICP protocol for communication among proxies and also offers the choice to configure the cache contents, such as proxy headers, cookies and object types. Regarding the safety characteristics of the system, NetCache maintains log files of the requests served, the privileges of all groups of users, lists of access control, user authentication based on password or even on RADIUS server. The most important drawback of hardware solutions is their usual high cost.

5. Policies and architecture

In the first part of this section we will describe the policies that should be in place and the architecture scheme we choose to implement.

Currently in Greece there is no government decision regarding the policies mentioned here. The policies' part of our suggestions—which could also serve to initiate a public discussion around the issues related with content filtering and content blocking in Greece is:

- Creation of Acceptable Use Policies for the Greek School Network (computers and services). Its objective must be to inform users for the purposes of the creation of technological infrastructure of the school network and the actions or behaviors that are considered acceptable and those that are not. By that way the government and the educational authorities will make public statement of their policies to reach all concerned parts. A well-written Acceptable Use Policy focuses on the responsible use of the computer networks, Internet included, but also responsible use of the access and transmission of information within and from the schoolrooms. Most such policies include:
  - Description of the underlying philosophy and strategy implemented into the school network for the access to Internet;
  - Report of the educational uses and advantages of Internet;

  - List with the catalogue of the duties teachers, students (and the students' parents probably) have on the issues involved by the use of Internet;
  - A code of conduct in Internet (do not give away personal information, credit card numbers, etc);
  - Description of the consequences of a possible violation of the AUP;
  - Description of what is considered acceptable and what not acceptable use of the school network and of the Internet;
  - Statement that denounces the school's responsibility under certain circumstances;
  - Reminder that the access to the school network and to the Internet, and the use of computer networks is a privilege;
  - Statement that the AUP is in accordance (does not violate) national telecommunications rules.
- Specification of the pages considered as 'improper', 'harmful', 'illegal', or 'void of educational content'. In many countries and in the cases where schools or homes use commercial software, one of the most important problems that raise opposition is the withdrawal of the kind of filtered pages.
- Choice of the suitable software program (squid in our case), which can meet the technical standards we set up.
- A critical point is the responsible, with no exaggerations informing of the teachers. They will have to explain the use of the decided policies to the students and through them to the society. Therefore their understanding of the problem and the proposed confrontation, along with their support can guarantee the success. Without their sincere participation success will prove hard.
- Control of the time and opportunities students will have access to the computers (when in lesson hour, they will be supervised from their teachers). If there is acceptance for such a movement, a program could be used to lock out some programs and allow only a limited number of approved ones. By that way the teacher will not have to worry about the exchange of files using ICQ, ftp or other similar programs.
- Control of installation of software on school computers. Only authorized persons should install software. The common problems are license issues, installation of 'Trojan horse' programs (programs with malicious or harmful code), worms, etc. Currently there are programs on the Internet that allow the administrator to lock certain programs and allow the use of few. Another aspect of the problem is that there is always the possibility of a break down and it is not always certain that the people capable to fix these problems will be available.
- Availability of educational material to constitute the positive counteroffer. This is another critical point: when students are free to do whatever they want in front of a computer screen and the subject of the lesson is outdated or indifferent to them, then the chances are higher that they will begin searching for other sites (games, cracks for programs, pages with sexual content, etc). It should be clear that the provision of the network is not the main objective: The main intension of schools is the use of all possible ways and new technologies for educational purposes, and

therefore the use of the school network and equipment for specific cognitive purpose.

- The age of students should be taken in consideration. While some state that for the younger students (ages 6–11) the problem is not that big, others claim that these ages are the most vulnerable. Our proposal is to implement the same policies (blocking of porn sites, of sites related to violence or drugs, or pages of aggressive content) regardless of the student age.
- Provision of the cachemaster's mail address for communication between him and the school community. If for example users run into a page that they think should not be inhibited, they can inform the cache master to remove it from the blacklist.
- Implementation of the adopted policies in a pan-Hellenic level for the school network to avoid increased maintenance complexity for the administrator in a day-to-day basis.
- If the school provides email accounts for all students and teachers (not true currently in Greece but visible in the very near future), many things should be considered: whether teachers should know the students' passwords or they will be allowed to use their email accounts totally free, whether for each student's account the responsible can be exclusively her/his father/mother, whether the state can claim that email accounts are provided for educational purposes and at certain cases will be examined by the school supervisor, etc.

The thoughts mentioned above are reflected in Fig. 1:

In the first level stands the Acceptable Use Policy Agreement, which is today considered as indispensable for all connected schools and libraries. At the second level, with their priorities signaled by the order of their writing, stand the outlined steps.

The Greek School Network's [20,21] topology has hierarchical structure and consists of the following levels: Backbone Network, Distribution Network, Access Network and Local Network Units as illustrated in Fig. 2.

The backbone network is the Greek Research and Technology Network (GRNET), which provides Internet Services to the Greek Academic and Research community. The specifications of the backbone network provide the possibility to create an efficient closed educational private network, which ensures Quality of Service, security and integration in the educational procedures.

The distribution network is the part of the network that interconnects points of presence (nodes) with the backbone network. The topology has such design in order to preserve the operational cost in low levels, which is particularly critical in large geographical region networks. These points of presence (nodes) are distinguished in two categories: *Prefectural Nodes* which are points of presence interconnected directly with the corresponding point of presence of the backbone network and *Regional Nodes* which are points of presence interconnected indirectly with the

[20] Greek School Network Homepage—http://www.sch.gr

[21] Greek Research and Technology Network Homepage—http://www.grnet.gr

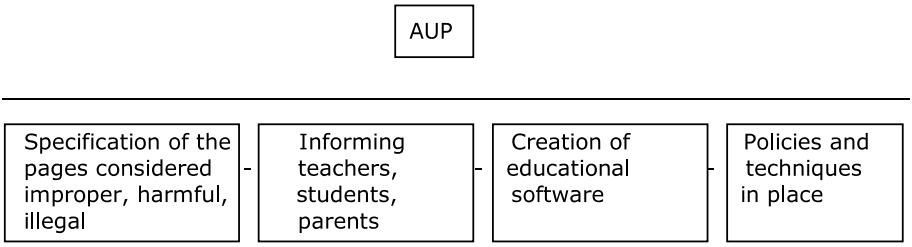

Fig. 1. The priorites for the Greek School Network.

Fig. 2. The Greek School Network topology.

backbone network through connections with the nearest prefectural node Winds of Aiolos, 2000.

The proposed architecture is in favor of the use of a proxy-cache server for the school network because the proxy server on a network—especially when it is configured as transparent—is the ideal point for the implementation of control on the requested objects: it concentrates the users requests and is the only entrance point of web traffic. As we will be using squid as a transparent proxy-cache server that solution fits perfectly in our case. We will use plug-ins (squid provides the possibility to use external programs) such as Squidguard, Squirm or Jesred to accomplish efficient and quick access control. That approach offers much more options and at the most cases there are ready site lists frequently updated and grouped by their content.

From squid's main page we read: [22] "Squid is a high-performance proxy caching server for web clients. Unlike traditional caching software, Squid handles all requests in a single non-blocking, I/O-driven process. It keeps metadata and especially hot objects cached in RAM, caches DNS lookups and implements negative caching of failed requests".

[22] Squid main page http://www.squid-cache.org

The extent of Squid's usage in educational networks all over the world and the level of acceptance it enjoys from the global networking community (especially researchers, scholars and educators), is obvious from the number of networks implemented on: the JANET network (a private, government funded network for education and research in England with 90–120 million hits per day), the GARR network (Italy), the DFN network (German), the Swiss network SWITCH, in the Dutch network SURFnet and at the academic network of the USA, the NLANR.

We are going to use Squidguard as the redirector of our choice because it is faster and opens less redirectors than it's competitor programs. The example presented in SquidGuard's page [23] to outline the program's speed, is that for a 2000 URLs list to filter and a 11,000 Urls database Squirm needs 2 min and 25 s, Jesred needs 1 min and 45 s, while for SquidGuard it takes only 9 s (tests were made on a 233 Pentium). As for the database size, it is extremely important the fact that it can hold many URLs, allowing by that way less generic regular expressions since actual sites can be entered, which minimizes the errors.

We will configure squid to fit our needs and we will try to be as efficient as possible. Additionally we could promote and incorporate into our solution, which is mainly based on the use of lists to allow or block access to URLs and domains, the use of filtering software.

6. Implementation issues

Squid is free software, licensed under the terms of the GNU (General Public License). The most important resource for squid's performance is physical memory, so fast disks are important for high-volume caches. Currently the last stable version of the software is version 2.4STABLE2 as of Monday August 27. Some new options were added to squid's configuration file after that last release. The most important of them will be presented. All new parameters intend to improve the performance, the security or the flexibility of the program and are grouped to better point their use.

Additional configuration parameters are:

The *maximum_object_size_in_memory* is a parameter which specifies the size limit for objects kept in memory, the *refere_log* parameter allows the creation of a logfile that contains http referer headers, the *extension_methods* parameter makes squid deny unknown methods (except standard http requests) unless they are included to this list, the null storage type allows squid to keep some statistics when it runs as a proxy without caching anything, although it would notuse any disk space, the *store_dir_select_algorithm* parameter selects which of the two currently supported

[23] squidGuard Homepage—http://www.squidguard.com

algorithms for selecting cache directories will be used (least-load and round-robin with *least-load* as default), and the *ie_refresh parameter which* enables a hack for Microsoft Internet Explorer versions that do not send 'no-cache' request headers when the user presses the reload button.

The following parameters where hardcoded in previous releases and are made now configurable:

The *dns_retransmit_interval* parameter that specifies the initial retransmit interval for internal DNS queries and the *dns_timeout* parameter that specifies the amount of time to wait for an answer from an internal DNS query (after this amount of time Squid gives up and returns an error).

Security-related parameters are:

The *authenticate_ip_ttl_is_strict* parameter when enabled allows squid to deny requests when a user appears to change IP addresses within the *authenticate_ip_ttl* time.

Parameters that test and improve performance options are:

The parameter *high_response_time_warning* makes squid print a warning to cache.log and syslog if the median response time goes above this limit, the *high_page_fault_warning* parameter makes squid print a warning to cache.log and syslog if the page fault rate goes above this limit, the *high_memory_warning* parameter makes squid print a warning to cache.log and syslog if the memory usage goes above this limit. The disk storage type is created to improve Disk I/O performance (disk uses external-child-processes to perform all cache disk I/O operations Shared memory is used for reading and writing data buffers).

In this section we will shortly describe some selective configuration options that are available and can be implemented to fine-tune the system. For every such configuration option, a suggested value will be indicated or further explanation will be provided.

Interesting information is the way related sites are grouped. The SquidGuard's blacklists database contains the following categories: porn, aggressive, drugs, violence, ads, audio-video, hacking and warez. The lists to be blocked are collected with the use of a program (it is described as a dump robot) that uses Berkeley DB. The most critical sections for a school environment appear to be the first four. An interesting configuration option is the monthly automatic update of that list.

The configuration file can include a number of different options. One of them is time 'spaces', which can be defined in order to grant or deny access according to hours of the day, days of the week or months of the year. We could define an access-granted-time for schools working hours writing 'weekly mtwtf 07:30–14:30'. The schools would then be allowed to access Internet at Monday (m), Tuesday (t), Wednesday (w), Thursday (t) and Friday between 07:30 in the morning and 2:30. In the previous time period, Saturday (or some range of hours everybody would agree access is needed) could be added to allow teachers and users use the network for research or educational purposes. The same can be done with IP ranges or user ids. That will result in a configuration file that will contain a number of different variables and access lists to operate with.

Another option is selective logging. Logging of blacklist matches is one possibility that could allow the administrator check what sites are blocked at a day or week period.

7. Statistical data from the Greek School Network

To show the impact and the results our system has up to now we present here some statistical data. We will describe in a short and abstractive way what happens when the users try to access web pages we do not allow access to: When a user from within the Greek School Network tries to visit a "forbidden" URL, his is redirected to one specific web server. That server displays a page with the message "The page you are trying to visit is forbidden. You can contact the cachmaster if you feel that page should not be blocked", among others. Finally, a CGI script creates a record in the log file with the network source address, the date, and of course the web page URL. From that point on we will refer to the number of forbidden pages that were blocked from the squid program as hits.

Fig. 3 shows the distibution of hits per day. The collected data refers to the period from November 9 until March 30 (9/11/2001–30/03/2002).

The number of distinct sites for this period is 63,173 while the top 50 pages (with the most hits) are shown in the following table. The percentage of pages with a hit number greater than 20 are only the 20% of the 63,173 blocked pages.

Those with a hit number greater than 1000 are the 2% of total pages.

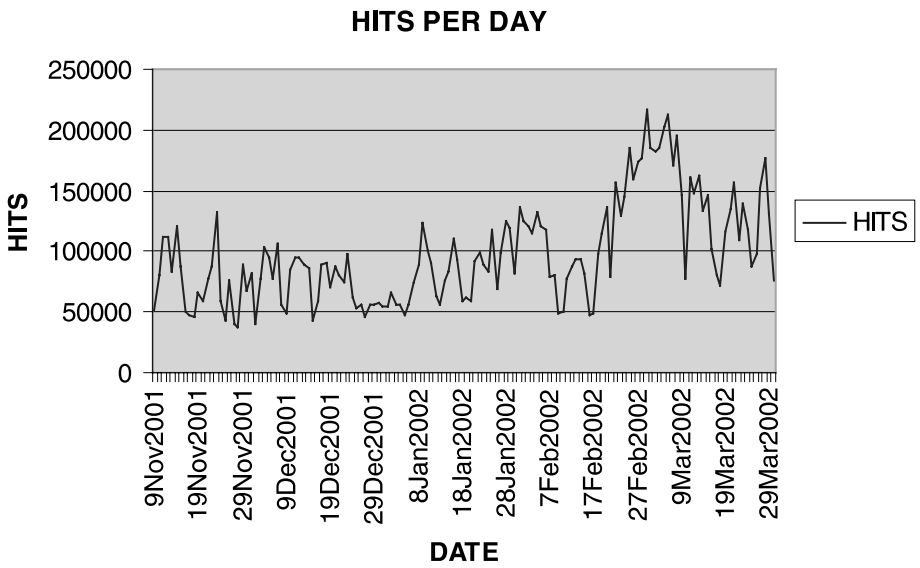

Fig. 3. The number of hits per day from 9/11/2001 to 30/03/2002.

HITS	URL
720471	stats.virtuagirl.com
629299	impression.7search.com
436604	banner.coupe.de
389365	go.163.com
347065	crd1.bannerbank.ru
338767	www.bannerchange.de
325018	cgi.sexswap.com
304809	www.addfreestats.com
237386	www.sextrade.de
201883	classic.adlink.de
201019	www.valuesponsor.com
191178	stats.hitbox.com
177498	ads.xxxbanner.de
170887	www.popupmoney.com
165387	www.camxchange.de
125968	www.kingsolomons.com
111116	www.sexcount.de
106634	www.bigmouse.com
90193	banner.visit-x.net
74907	www.adultlink.de
72379	go2.163.com
71838	top.gaysexlist.com
68823	go4.163.com
67422	top.list.ru
66381	go6.163.com
62712	www2.x-check.de
61647	www.terra.es
61483	jmm.livestat.com
61371	go5.163.com
60074	www2.playerdome.net
59145	www.hobbyhure-xchange.de
56891	www.inet-cash.de
55334	west.adlink.de
52914	count.paycounter.com
50899	www.bannermeister.de
50240	www.jp-kogalsex.com
49506	www17.smutserver.com
49498	www.intergal.com
48629	www.fransexe.qc.ca
48417	go1.163.com
48323	counter.mtree.com
45767	mirrorsearch.speedbit.com
45576	adserver.ign.com

HITS	URL
42958	go7.163.com
42951	v0.extreme-dm.com
42489	go3.163.com
41527	www.maximumcash.com
39579	www.easywarez.com
38586	go8.163.com
38068	service.bfast.com

8. Conclusions

We firstly feel the need to state that we focus our interest on the educational networks or networks established for specific purposes. By no means the thoughts that are presented here, apply for adults or people who pay to establish access to the Internet. Our propositions do not apply on any other field or domain of public activities like libraries and universities. We believe that filters should not be uniformly applied because that would be against freedom of expression, though and inquiry especially by scholars and university students.

The rapidly evolving nature of the Internet virtually ensures that no filtering technology can be a hundred percent perfect. Thus, even when a school uses a particular solution, children should be taught how to respond if they still manage to access something that it is inappropriate for them.

If a school district employs monitoring, its Acceptable Use Policy should explain what it would be doing, and the procedures a student should follow if he or she encounters a site that would be considered inappropriate. Certain kinds of network management products may provide basic information on how students and staff are using the network.

The practices are quite the same across the different nations. This happens because the filtering problem is relatively new and those who started first to deal with it lead the way. The content filtering issue we discuss herein has never caused a central debate either from parents for their kids or from schoolteachers or even from the government in Greece. We work on that issue because we believe that every country should be aware of the possible dangers and be as prepared as possible to deal with them. If we know and understand what happens in the rest of the world and in countries that have a lead over Greece in technological issues we will save ourselfs time and trouble.

The biggest percentage of the pages on the net is written in English, which is nowadays an international language. Therefore the most interesting sites and definitely those that get read from the biggest audience are the same even for the countries whose language is other than English. That means that the problems are the same since their source is the same. This is why many countries seem to adopt the same strategies despite their natural differences.

For the most educational institutions, a convenient solution has been the installation of filtering software on the proxy server. This is the ideal place to do filtering for a network because it is the one point through which all network communications pass. Proxy servers also can produce data logs that can help monitor system usage and performance, whether they are used for content monitoring or not. Filtering on a proxy server can have an impact on network performance because of the need to match a URL against what may be a long list of blocked sites. The use of a caching server, which, in simple terms, permits a recently viewed site to be viewed again without having to send a request out beyond the network, can help speed access and reduce the bandwidth that would otherwise be needed.

Despite the development of new software products, rating systems, and industry-supported online safety campaigns, the parental and school guidance will remain the basic factor of the solution of the problem. Another way out is the creation of an enclosed, protected environment where children could enjoy technologic advance accessing material only one or two decades ago scientists could not access.

Today it is technologically feasible to group the human knowledge obtained up to date, categorize it according to the targeted student ages and thematic areas and provide it to the students of all the schools all around the world. When educational authorities will be capable to undertake such a task and provide its results to the students, the objectionable, harmful or illegal content we desperately try to lock out, will be of much less interest than it is today.

The problem we are dealing with is not simple. More than 20 million of the 35 million domain names registered in the world (February 2001), are ".com" domains. Web server surveys have shown there were more than 27 million web servers in operation as of January 2001. [24] One study estimated the Web to have approximated 800 million pages in February 1999. [25] While the above results are not accurate for 2001, they are indicative of the order of magnitude of the sites on the Internet. If this is to change, obviously it would not shrink.

The software that blocks access to certain sites is not perfect (lists may include pages that should not be included or vice versa). These lists cannot be human reviewed as demonstrated from the previous paragraph. The common way to cope with that issue is the provision of the cachemaster's email address, where users will be able to report the pages they think that should not be blocked. That will solve the problem as it is estimated that such cases are limited.

However even then, the solution will not be perfect. It can never be since the ideal solution—content labelling and categorizing with possible alterations according to reader's age or other factors—is hard even for a human: What a person could claim he can 'say' a certain page is proper for a sixteen years old teenager whilst it is not proper for a younger child? What his background should be? It definitely cannot be a computer scientist. What his criteria will be when deciding on the appropriateness of

[24] Worldwide Domain Statistics, NetNames Global Domain Name database http://www.domain-stats.com/

[25] Web Server Survey http://www.netcraft.com/survey/Reports/0101/

the content? And how his 'rules' will be coded and executed by a dumb robot—if we suppose that such rules do exist?

Current technologies create (more or less) lists with forbidden sites. These lists contain pages with 'forbidden' words or sentences. If the database is big and gets updated often then the mechanism is good. If the mechanism incorporates interaction with the user or the administrator it is even better.

Sincere participation and plain collaboration of all the people related directly or indirectly with the educational process can ensure and safeguard the success of such a project. Without that support the project will be imperiled, since reactions can emerge from the school community and interfuse to the society after short time. That might have been the case of the United States of America where legislation came at a time when people had not accepted the idea of mandatory filtering in schools and libraries.

Even if we find the ideal solution and implement it in our school networks, no one can guarantee that this will be the ideal solution or even an appropriate one, in one or two years time. People with wide knowledge of the Internet and the emerging technologies should monitor the developments that take place and adjust their policies to the new era.

Implementation of content/access control does not solve radically the problem of students using the school networks to gain access to objectionable or illegal material. There is a number of ways that can be used to exchange that material and email and ICQ are only the two most common programs to do that. While solutions for that issues do exist (there are programs that allow only the execution of authorized programs, while the rest are locked out), the governments and the school communities should decide if they wish to implement so hard constraints.

When content filtering will be implemented, there will be a quick transition from the previous state of immunity to a controlled status. That could cause an increased number of complaints. Summarizing, even the strongest arguments against the use of filtering software (overblocking/underblocking issue, subjective criteria or possible use of filtering technologies as censor ware tools) are not strong enough to stop the ongoing trend for use of such technologies in schools as the demand for safe networking experiences for minors is a reality in all technologically advanced societies.

9. Future work

Our future plans include the improvement of the current design and architecture to best fit the continually emerging technologies and therefore the successful address of the new problems that almost certainly will come up. As with all aspects of technology, the tools and solutions for managing Internet content will continue to evolve, as will do too the methods that unscrupulous Internet users and merchants can use to lure children to places and materials that could be considered inappropriate for minors.

An interesting topic seems to be the use of DansGuardian program (its author recommends its use with SquidGuard). DansGuardian is a freeware filtering proxy

for Linux, FreeBSD and OpenBSD that uses Squid to do all the fetching. It's filtering has four steps: It checks the content of the pages against a configurable denied phrase list, it implements PICS filtering, it checks the MIME type of the requested file and checks this against a configurable denied MIME type list, and finally it checks the file extension of the request against a configurable denied file extension list.

Another priority would be the examination of the available ways to implement virus checking and it's effectiveness. The viralator program [26] is a perl script that scans http downloads on Linux servers for viruses, after passing through the squid proxy server. It works in conjunction with virus scanners like McAfee, AntiVir etc. The evaluation of such techniques will be an important target.

Reference

Adamopoulos, N., Bouras, C., Ganos, P., Karaiskakis, D., Paraskeyas, M., 2000. Winds of Aiolos: How connect the Greek Schools in Internet. Towards the e-learning Community, Bolton International Conference, Bolton, UK, 19–20 October 2000, pp. 27–37.

[26] Viralator Homepage—http://viralator.loddington.com