Unwanted/malicious Internet content and it's filtering for certain user groups

Vanja Librić Radojević
Hrvatska akademska i istraživačka mreža CARNet
J. Marohnića 5, 10 000 Zagreb, Croatia
vlibric@carnet.hr

Ljerka Luić
b4b d.o.o.
Ulica grada Vukovara 271/VIII, 10 000 Zagreb, Croatia
ljerka.luic@b4b.hr

Abstract. *Access to the Internet has become largely available to almost every household. Using computers today is much easer then ever before, without special user knowledge. That means that Internet today is available to all generations of users without special knowledge about computer technology. In this work, we have presented more realistic way of functioning of this medium that largely confutes user perception of anonymity. In this work great emphasis is on the content that can be found on the Internet, including potential dangers that lurk there, with recommendations how to protect from them.*

The business world nowadays cannot exist without Internet, but even today there is large number of users, who do not use it in a proper way, which may lead to security problems for whole organization and/or to themselves. There are many possible measures for protecting users. Firstly there is education, then various antivirus and other security tools which protect from malicious software, to tools that enable filtering of certain types of on-line content. Those mentioned protecting measures are presented in the paper, with emphasis on real-life use of certain tools for content filtering in some organizations in the Republic of Croatia.

Keywords. malicious Internet content, filtering, security tools

1 Introduction

From the beginnings of Internet some thirty years ago until now, Internet has become so popular that most of the common citizens think of it as a regular communication media, like telephone or TV. In the dawn of Internet it was used mainly by the academic community members, who shared common interests, and since there was no available content that could be popular with common people, there was no computer crime on the Internet. It usually looked like small and safe neighborhood where everyone knew each other.

Nowadays situation is quite different, and more and more people cannot imagine their lives without Internet and its advantages. The entire part of, lets call it "real life" has migrated to Internet. So, today, we buy stuff on the Internet, we manage our bank transactions, we buy airplane tickets, place hotel reservations, etc. this is a whole new world of services together with their editors and supporting workers. [2]

In this work the focus is mainly on the content available on the Internet. Although prevails high-quality, educational and interesting content, we have concentrated our attention was concentrated on Internet content which are considered harmful for certain user groups, and we presented example on how that type of content can be filtered.

2 Content on the Internet

The most intriguing fact about Internet is that everyone can be author of its content, and for publishing of that content there is no need to have large sum of money or equipment, and even there is no need to have some extra user skills. The advantage of the Internet is a huge population of individuals which make a large virtual "country", and who are willing to share their knowledge and expertise in various areas of human life, and to expand common knowledge with others on the Internet.

The amount of content available on the Internet cannot be easily compared to any library or legacy

database system. On the other hand, we should be aware of the fact that sometimes is very hard to find right information, and that it can be very fatigued task to find just the right and interesting information that we are seeking. [1]

3 Rules of good behavior on the Internet

3.1 Netiquette

Internet is not just a computer network, but also a huge and informal community where people of different cultures, religions, nations, etc. meet each other. Because of that, Internet has its own code of conduct in order for community to function. Large amount of people communicate through the Internet, commonly in improper way. They usually don't know basic rules of Internet behavior, so the culture of communication should be learned.

The rules of good conduct on the Internet are called Netiquette, and they are similar to those in the real life. Aggressive and insulting behavior and misuse of privacy are unacceptable.

3.2 Supervisory authorities and sanctions

We are all aware of the fact that there is no Police or other legal authorities on the Internet, but since the problems which endanger computer security are numerous and dangerous, there are several supervisory authorities that are responsible to handle them.

The most common authorities are:

- CERT - Computer Emergency Response Team – is a national center which helps to solve computer security incidents in which at least on side is from Croatia. It also publishes reviews and distributes free security related programs. It alos publishes and distributes security advisories, documents and recommendations by which computer systems security is enforced
- Abuse Service – is a service which is established in every Internet service provider. It handles incident reports which are related to computer security incidents, Netiquette abuse and accepted terms of services.
- PKI CA - Public Key Infrastructure, Certification Authority is an authority which issues and signs digital certificates which are added to the digital documents. It serves to authenticate a person or computers which use some service(s), application or just communicates with other users through the Internet or in other digital means of communication.

We will go now in more details about the functioning of Abuse Services which are also more common services in our country which handle computer security issues. Every Internet Service Provider (ISP) has its own Abuse Service, which task is to receive and process security incident reports which are originated from users of that ISP. Abuse Services do have a goal to receive and process incident reports concerning the computer security incident and misuse of resources like:

- SPAM,
- Netiquette,
- Viruses / Worms / Trojans,
- Copyrighted materials,
- Commecial use (where is not allowed),
- Unauthorized access,
- Intrusions,
- DoS (Denial of Service),
- DDoS (Disributed Denail of Service).

We are presenting here statistical data of CARNet Abuse Service of number and type of computer incidents in 2008. in order to, at least partially, we can get a better picture of number and different types of computer security incidents in on of Croatias ISPs. We can presume that data from other ISPs are quite different because of different number of users and availability of services across the country. Fig. 1 represents statistical data in year 2008. from CARNets Abuse Service. [5]

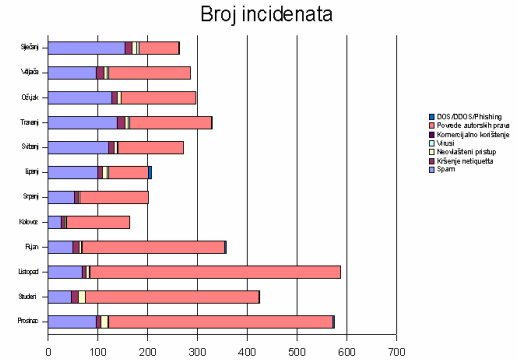

Figure 1. Statistical data about number ad types of incidents in CARNet

It is important to emphesize the fact the in May 2007, fourteen ISPs in Croatia have signed an Memorandum of installation of Task Force of Croatian Abuse Services, which task is to raise quality of computer security and acceptable use of Internet in our country. [10]

4 Dangers on the Internet

There are two main reasons why comes to computer security incidents. The first one is that there are errors in products themselves, and the other is because of bad configuration or inadequate maintenance of the systems. Computer code which utilizes mentioned errors are usually called *exploit code*.

Computer security incidents generally can be divided to: malicious software, unsolicited mails, and other problems, so we generally can say that there is vast possibilities of illicit activities; from plain

vandalism, through benign types of frauds, and real heavy criminal offenses. [2]

4.1 Malicious software and possible measures of protection

From vast number of malicious programs we should mention two types: Viruses and Worms, which usually bring down our computers to a point that they cannot be used anymore. Infected computer can secretly be used to distribute illegal materials and programs, SPAMs and to become part of coordinated attack to other computers (usually DDoS). When our computer is infected and that infection is discovered, it is possible that our ISP will ban our further connection to Internet until we clean our computer, in order to protect other users.

In order to protect ourselves from aforementioned programs we usually use special tools, calls anti-virus programs which for every virus or worm has definitions which can detect and usually also clean from infection.

Spyware and adware based incidents are usually connected with the web sites that we visit on daily bases. They infect our computer through the imperfections of used web browsers, or by using social skills to persuade users to skip certain security alerts. When they became in control of our browser and/or computer they can hide themselves in order that user is not aware of the problem, but usually there is some strange behavior of our browser that can be seen (opening of unwanted web sites, etc.)

Against this programs we can fight with specialized software which are called anti-spyware tools. Those tools can detect and usually eliminate them from our computers. It is necessary to update this tools on a daily basis to be sure about their detecting abilities. It is known that removal of spyware is not easy task, and sometimes there is need for several different tools to do it.

Although dialers were initially used as a legal programs, nowadays they are usually mentioned as computer incidents. They are used to place calls to expensive remote (usually over ocean) countries, without the knowledge of user, using in this way telephone circuits in countries that are usually not strictly enforcing laws. This usually means that even if it is clear that the users telephone line was abused, it is almost impossible to return its money back. To fight dialer programs we usually install anti-dialer tools or we ask our telco to put in place call restrictions to mentioned countries. It is also true that dialers are not so common nowadays since the dawn of broadband connectivity, which is not using classic telephone lines and modems anymore.

Trojan horses are specially engineered programs that rely on our trust, so they were named by that. They introduce themselves as an interesting program, sometimes it's a video clip or audio file, but in background there is much more going on. We, as a users, intentionally let those programs install on our computer in good faith, because we are convinced that there is nothing wrong going on.

SPAM is probably the most common incident. The fight against unwanted mails which take our time, and money, is still going on, although more then thirty years have passed from from first SPAM message. Today SPAM is increasingly represented in total traffic of emails, and it largely complicate business correspondation. In accordance with Spamhaus, a non-profit organization, 90% of all exchanged emails is SPAM., and that for around 80% of alls SPAMs only 200 spammers are held responsible. Billions of SPAM email are sent every day, mostly used for marketing purposes, virus distribution, and Internet impostures. When Internet has become massively used by both business and private suers, certain people have concluded that this is good medium to make money, and SPAM was born. [3]

4.2 Social engineering and measures of protection

Dangerous emails are not always apparent on first sight. They are created by the masters of social engineering – who manipulate users in order to go around security measures. Fraudsters are greatly abusing user ignorance and credulity.

We always need to have in mind that when we communicate with someone over the Internet whom we never met, everything we know about him is usually only facts that that person told us. We should take special care about commercials and similar emails which we received without our request (SPAM), persuading us to click on the links inside the message taking us to some web sites. It is recommended that we never follow those links, because they usually take us to the sites full of malware that can infect our computer and potentially can harm other Internet users. [3]

5 Unsuitable content

As we stated earlier, anyone on the Internet can publish some information which will become available to all Internet users. Because of that simplicity to transfer data to others, Internet has largely contributed to freedom of speech, but also made possible to publish and spread unwanted and harmful content.

Unsuitable content like child pornography or various Internet sites which spread interracial hate are banned by the law in almost all the world. Most of countries worldwide do have laws which describes what is forbidden to publish on the Internet, and if the server with that content is on their territory, there is possibility to remove that server from the Internet. For example, in Germany it is forbidden to publish web sites which promote Nazism, but, since there are no borders on the Internet there is possibility to put that

content ons servers in other countries, and users from Germany still can view those pages. But even in those cases the government is not defenseless. There is still possibility to place law suit in that country where the server and content resides.

In countries like China, where major part of available content on the Internet is considered as unsuitable, the censorship is maintained in other ways. Instead of placing law suits and requests to remove Internet sites that are considered offensive, they instead just filter them. There is a law in China which instructs ISPs to filter web sites that are considered unwanted. This censorship project in China is often called "Golden Shield", and the firewall that is "protecting" Chinas cyberspace is often called "Great Chinese Firewall". [7]

5.1 Potentially endangered user groups

Potentially endangered user groups are untrained users, beginners, and children. We shall take special care here about children who are becoming Internet users in large numbers every day. At the same time it is fascinating and scary to see their ability to use Internet, and on the other hand to be aware of their age, their frankness and naivety. For those reasons, often provocative by themselves they come in touch with content that is now suitable to their age. For that reason it is possible for the parent to place rules which pages children can visit and which they cannot. Internet content can be filtered based on several categories. It is also possible to have in place profile passwords, so that every user has its own control of activity logs for every PC in the home network. [7]

6 Content Filtering

Internet traffic filtering should be one of the ISP tasks in order to protect their users, because they can automatically stop a lot of potentially dangerous content. On one side, traffic is filtered to reduce SPAM, unauthorized access, etc., that is Internet content that can be harmful to end users and organizations which usually block incoming and outgoing traffic which is not acceptable by their corporate security policies.

To define a security policy is a first step to implement such solutions, and it's hard to find respectable organizations or companies that do not have such rules. Since the Internet is accessible to large number of users on their workplace, which bring a variety of services and content, that can become an issues for the company.

Consequences of Internet misuse at workplace is usually seen as a reduced workers productivity, congested Internet links and higher Internet costs. By using filtering and caching techniques, can reduce or solve most of mentioned issues. There are two available technologies to filter content: filter lists or software algorithms. Filter lists contain addresses of the web sites which are listed as restricted sites. Those lists are continually monitored and updated, in order to categorize web sites correctly. In this way the possibility to place a web site to wrong category is minimized. On the other hand, software algorithms use different methods to identify potentially dangerous or unwanted sites. Web sites are blocked based on their content, and not by the fact that they are listed on list of banned sites.

Both of this technologies have their advantages and shortcomings. The major dilemma when selecting form those technologies is accuracy. Filter lists usually are more accurate, primarily because of additional human checking, but they can also be outdated, because the number of web sites is growing rapidly. [12]

On the other hand, software algorithms use different methods by which they on-the-fly can detect some unwanted patterns or words that are programmed in advance. The major shortcoming of this method is it possibility to make mistakes, because it's possible to block a regular site, which is listing just few mentions of unwanted terms or words. On the other hand it can filter any site on the Internet. There are three possible ways to cache content:

- Client side caching – is a standard option available in many major web browser. It is caching content on the computer itself.
- Server side caching – enables possibility to reduce number of requests imposed to a web server. Almost all proxy servers can work this way, but then they are usually called differently: reverse cache server, inverse cache server, httpd accelerator, etc.
- Proxy cache servers – they are usually placed close to routers connecting to the "outside world" e.g. Internet. They are usually used to accelerate network traffic coming form the Internet, which in return means that we can use lower bandwidth (cheaper) on our Internet links. This type of systems can server many users, and they place web pages in their local cache, and then distribute them to other users requesting the same site, instead of pulling the same web page from the Internet.

Fig. 2 shows one of the possibilities to filter Internet content on the private network border (Intranet). [11]

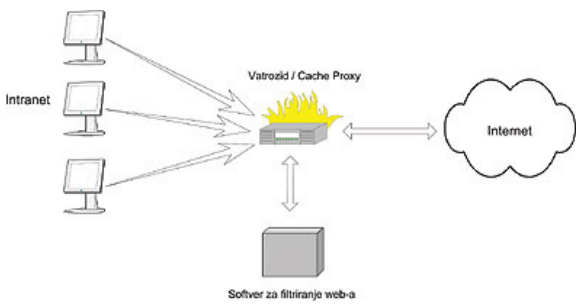

Figure 2. Internet Traffic Filtering

7 Implementation of unwanted and dangerous content filtering

From all above mentioned it is understandable the need to take some level of control and to protect users/customers from unwanted/dangerous content.
The usage of content filtering tools is possible in business and educational worlds. What is considered to be unwanted content in a organization and which categories of Internet content should be blocked to workers during the work hours is controlled by the employer through the security policy. Most of Croatian organizations don't use any type of content filtering, there are couple of examples where there is such policy, like Croatian largest petrol company INA, and Agrokor. It should be clearly stated that the process of content filtering can only partially protect users. There is also need for constant education and informing of new threats from the Internet, how to avoid them, and what to do if you encounter some of unwanted content or other Internet threats. [4]

In our example, we focus on most endangered user group – children, and show in details how one of the tools was implanted. This tools is called CA eTrust Secure Content Manager, and it can also be used to filter other user groups.

CARNet has, in cooperation of the Ministry of Science, Education and Sports, connected all primary and secondary schools Internet connectivity. In that way Internet has become available to wide population from children to school teachers, including all of this content. Since CARNet was aware that in this way different type of content from the Internet has become available to children, from the beginning of 2007, and under the resolution from the Ministry of Science, Education and Sports, content filtering was implemented. The goal of content filtering is to block unwanted content from children which are using school computers. [6]

CA eTrust Secure Content Manager operates on principle of blocking sites that are listed in some of restricted categories. Every Internet web site is categorized based on content which is providing, and those categories that are preselected are not available to the users. The task of web site categorization is done constantly and database of categorized sites is updated every few hours. It is also possible to administrator of the system to manually add or remove some sites that are found to be listed in wrong category. Every Internet site can be in one or more categories.

Here is the list of the categories that are blocked by CARNet in accordance with resolution from Ministry of Science, Education and Sports:

- Drugs,
- Gambling,
- Gambling Related,
- Gruesome Content,
- Hate Speech,
- Hacking,
- Malicious Sites,
- Nudity,
- Profanity,
- Pornography,
- School Cheating Information,
- Spam,
- Tobacco,
- Violence.

Fig. 3 shows the basic principles of connection and aggregation of traffic from the schools and redirection of such traffic towards the content filtering system. [5]

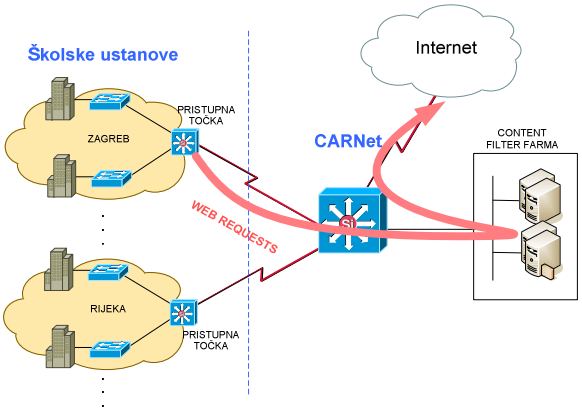

Figure 3. Access to Internet from the schools

Internet traffic form the schools is aggregated in some of the largest regional centers in Croatia: Dubrovnik, Zagreb, Osijek, Split, Rijeka, Pula, Zadar i Šibenik. Every center aggregates the traffic from the schools and redirects HTTP traffic from schools to the content filtering farm in order to be processed.
Fig. 4 shows the basic architecture of the Content Filtering system.

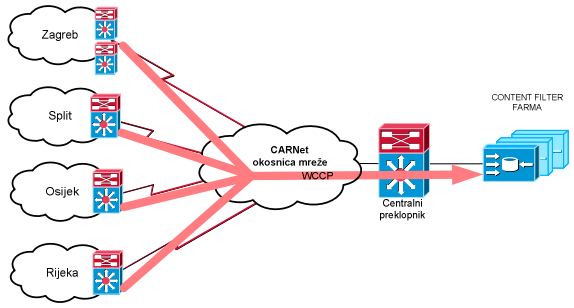

Figure 4. Content Filtering system architecture

WCCP protocol is used on the routers in regional centers to redirect traffic to the proxy cache cluster. By using multiple routers in a service group it is possible to achieve high level of redundancy, interface aggregation and load balancing. In order for the entire system to be reliable, scalable and efficient it was recommended that routers of the same hardware and software characteristics should be used.

In contrary, there can arise compatibility problems of WCCP parameters and impossibility to establish and maintain WCCP communication between routers or L3 switches on one side and proxy cache farm on the other side. In order to simplify implementation and troubleshooting of eventual issues, and to simplify monitoring and controlling of the WCCP system, it is recommended a design model with minimal number of access points in the system. In this way we can increase performance and security aspect of entire system. [8]

eTrust SCM uses some of the newest content management technologies in order to offer top level, flexible, and simple management of SPAM , viruses, and control of email and web page content which are accessed. eTrust SCM analyzes HTTP/FTP and SMTP traffic (in our case only HTTP traffic) which is passing through the network. Control Center service is responsible for traffic flow of certain type of data between Content Filtering Engine and Manager Console:

- rules and setup in distribution,
- quarantined objects that are sent to Quarantine Manager (only for SMTP filter),
- report data which is sent to the Reporter service.

Traffic filtering is based on a comprehensive set of rules and setups in Manager Console and by policy filters that are defined in Content Manager rules database. The sequence of event viewing is sorted by the sequence of the rules. When communication session starts, the traffic analysis is also started and the rules are checked in order to check if the traffic is of protocol that should be checked. [9]

8 Conclusion

The benefits that we gain from using the Internet are greatly bigger that the potential threats, and it is hard to imagine today modern life without the Internet. If we transfer manners from our real, everyday life to virtual world, even potential dangers should be far less threatening. As we saw, there are several tools to protect us as a Internet users, but we also think that the biggest protection is to have fair amount of "common sense".

The major Internet "problem" is availability of various content, and old saying is that opportunity makes the thief. Certain Internet users are very often more naive then in a real world. People usually feel that they are more protected behind their computers, they communicate more freely than they would be in a face-to-face situation, that dare to do things that they would never do in real life. Off course, most of this things they are willing to do because they think that their anonymity is guaranteed, and that often is not a case. Education is mandatory, because only well educated Internet user is really protected and then he can enjoy all of the wonders Internet has to offer.

Acknowledgments

Publication of this paper was supported by grant #036-0361983-3137 by the Croatian Ministry of Science, Education and Sports, to which the authors of this article are grateful for support.

References

[1] Berkman Center for Internet and Society at Harvard, http://cyber.law. harvard.edu/, 10.2.2009.

[2] Centar za informativnu dekontaminaciju, http://www.6yka.com/do/da,361, 13.3.2009.

[3] Coins DIG, http://coinsdig.com/business+news/anniversary+of+spam, 1.4.2009.

[4] GFK - Centar za istraživanje tržišta, http://www.gfk.hr/press1/internet2.htm, 1.2.2009.

[5] Hrvatska akademska i istraživačka mreža – CARNet, http://www.carnet.hr, 15.3.2009.

[6] Hrvatski povijesni portal, http://povijest.net/index.php/Svakodnevna-povijest/Internet-povijest.html, 21.3.2009.

[7] Incisive Media Ltd., http://www.vnunet.com/computing/analysis/2213355/china-cracks-insider-cyber-3925005, 3.4.2009.

[8] India eNews, http://www.indiaenews.com/technology/20080409/110041.htm, 9.4.2009.

[9] „KompletNet" – Internet marketing servis, http://www.grapnet.com/page_net.php?id=334&oid=312, 21.3.2009.

[10] Poliklinika za zaštitu djece Grada Zagreba, http://poliklinika-djeca.hr/index.php?option=com_content&task, 21.3.2009.

[11] Squid Foundation, http://www.squid-cache.org/, 5.4.2009.

[12] T-Com Hrvatska – Portal, http://www.tportal.hr/tehnologija/internet/fset.html, 13.4. 2009.