DESIGNING OF WEB FILTERING POLICIES

Ahmed Salah Eldeen and Akram Mustafa Elhaj Adam

ABSTRACT: *At present, the internet plays an integral part in business and education. However, without setting appropriate controls in place, universities are likely to be faced with a broad range of issues. One of which is a misuse of the internet by students if they try access to inappropriate web content during exams. The objectives of this research are to design the web filtering policies, use a tool having the capability of denying access to the content of all websites except the exam site, and to compare and validate the Smoothwall Express over Monowall. The research assisted in assuring the quality in respect of final achievement for the students.*

المستخلص: في الوقت الحاضر، تلعب شبكة الإنترنت دورا أساسيا في مجال الأعمال والتعليم. ومع ذلك، من دون وضع الضوابط المناسبة في المكان فان الجامعات مثل جامعة السودان المفتوحة التي لها الأسبقية في تقديم أول امتحان على الانترنت في السودان في عام 2014 من المرجح أن تواجه مجموعة واسعة من القضايا؟ واحدة منها هو إساءة استخدام الإنترنت من قبل الطلاب اذا حاولوا الوصول إلى محتوى غير مصرح به على شبكة الإنترنت خلال اثناء الامتحانات. في عام 2014 حضرت أول امتحان على الانترنت في جامعة السودان المفتوحة لبرنامج الماجستير في تقنية المعلومات- فرع الفاشر. ولقد لاحظت أن جميع مواقع الانترنت كانت مفتوحة خلال ساعات الامتحانات. وهكذا، كان هناك إمكانية للبحث عن إجابة على الانترنت طالما كانت المواقع غير المصرح الوصول إليها مفتوحة. الهدف من هذا البحث تصميم سياسات تصفية المواقع علي شبكة الإنترنت، مما سيساعد في ضمان أن نتائج الامتحانات سوف تعكس جهد الطالب الفعلي ؟ وسوف يتحقق ذلك من خلال استخدام أداة "*Smoothwall Express* " التي لديها القدرة على منع الوصول إلى محتوى جميع المواقع باستثناء موقع الامتحان. إحدى الطرق التي يمكن استخدامها لمنع الوصول الي صفحات الويب غير المصرح بها هي استخدام نظام التصفية على شبكة الإنترنت. وهو برنامج يمكن أن يفحص صفحات الويب لتحديد ما إذا كانت بعض أو معظم الصفحات لا ينبغي أن يتم عرضها للطلاب. هذا البحث سوف يساعد في ضمان الجودة فيما يتعلق بالتقييم النهائي للطلاب. على العكس من ذلك، سيكون احتمال عدم التطابق بين نتيجة الامتحان النهائي وجهد بعض الطلاب الفعلي حاضراً إذا كان الوصول الي المواقع غير المصرح بها ممكناً.

KEYWORD Design, Policy, filtering, OUS

INTRODUCTION

The Internet is a network of global exchanges - including private, public, business, academic and government networks - connected by guided, wireless and fiber-optic technologies. The terms Internet and World Wide Web are often used interchangeably, but they are not exactly the same thing; the Internet refers to the global communication system, including hardware and infrastructure, while the Web is one of the services communicated over the Internet. Communication systems were first developed for radio communication. However, as computing advanced, peer-to-peer (P2P) communication was gradually delivered and enhanced. During the last two decades, the Internet has influenced and upgraded networking to global standards. Billions of Internet users rely on multiple application and networking technologies.

Internet Protocol (IP) is the Internet's primary component and communications backbone. Because the Internet is comprised of hardware and software layers, the IP communication standard is used to address schemes and identify unique connected devices. Prominent IP versions used for communications include Internet Protocol version 4 (IPv4) and Internet Protocol version 6 (IPv6).

At present, the internet plays an integral part in business and education. However, without setting appropriate controls in place, universities such the Open University of Sudan (OUS) which it has the precedence in introducing the first online exam in Sudan in 2014 is likely to be faced with a broad range of issues? One of which is a misuse of the internet by students if they try access to inappropriate web content during exams.

Web filtering is required to stop the students of the OUS from accessing unauthorized websites during exam hours if they intended to do so. Web Filtering is basically done through URL Filtering by either allowing or dis-allowing access to the specific web page.

The policies of the websites filtering could be set by the IT department personnel. So, there is a software application running on a local Firewall that contains of the URL filtering feature which enables to block individual websites.

In 2014, OUS conducted the first online exam for the master program in Information Technology in El Fasher branch. It is observed that all the websites were opened during exam hours. Thus, there was a possibility to search for an answer on the internet as long as unauthorized websites were accessible.

In addition to the normal invigilator function, this has led to necessity of utilizing from web filtering products, which OUS could deploy extensively as part of the proactive management process.

The objectives of this research are defined as below:

i. To design the web filtering policies, which will assist in ensuring that an exam results will reflect the actual student's effort. This will be achieved through using a tool (*Smoothwall Express*) has the capability of denying access to the content of all websites except the exam site.

ii. To compare and validate the *Smoothwall Express* over Monowall.

One of the methods that can be used to block unauthorized web page is by using web-filtering system. Web-filtering system is a program that can screen the web pages to determine whether some or most of the pages should not be displayed to the students. Then, the filter checks the origin of the contents of the web pages according to the set of rules provided. In this research, by using the web filtering system, it will block any web pages that contained unauthorized content.

Around a year 2000, Smoothwall[13] software has been founded as an open source project exclusively to develop a Smoothwall application to be used as a firewall to prevent malicious web activities. It is software that enables the control of incoming and outgoing packets in a network. Today, the project is funded and supported by Smoothwall Limited, a company established in 2001 to handle the proprietary version. Strangely enough, most of the functions performed by *Smoothwall Express* can also be carried out by advanced routers and monitoring tools providing that policy and rules are well layered out, implemented and monitored regularly. To compensate for extortionately costly firewall, router and packet filter software from other vendors, Smoothwall provides customers with a choice between open and proprietary version.

The proposed Firewall (*SmoothWall Express*) is a Firewall distribution based on the GNU/Linux operating system, designed for ease of use, is configured via a web-based GUI.

The *SmoothWall Express* enables us to easily build a firewall to securely connect a network of computers to the internet.

Almost any Pentium-class PC can be used, for example, an old low specification PC long redundant as a user workstation or server. The *SmoothWall Express* creates a dedicated hardware firewall, offering the facilities and real security associated with hardware devices.

The *SmoothWall Express* comes pre-configured to stop all incoming traffic that is not the result of an outgoing request. The rules file that implement this policy are part of the system configuration and should not normally be edited by other than the configuration procedure.

In respect of system and hardware specifications, requirements may vary depending on traffic throughput and processing requirements, which themselves vary according to the number and size of protected networks. *More details on this will be discussed in chapter three.*

This research will assist in assuring the quality in respect of final achievement for the students. On the contrary, the probability of mismatching between the final exam result and the actual student's effort will be presented if unauthorized websites are accessible.

This Chapter was organized into two subject areas which relate to this research. First, it will explain about the literature review, through clarifying the definition of design, filtering, and policies. Then, it will present the review of related literature and studies. This part represents a search of the literature from the sources such as books, magazines, journal and Internet.

Specialized computers known as routers are responsible for directing packets appropriately. Each router is connected to several communication links, which may be cables (fiber-optic or electrical), short-range wireless, or even satellite. On receiving a packet the router makes a decision of which outgoing link is most appropriate for getting that packet to its ultimate destination. The approach of encapsulating all communication in a common format (IP) is one of the major factors for the Internet's success. It allows different networks, with disparate underlying structures, to communicate by hiding this non-uniformity from application developers.

Routers identify computers (hosts) on the Internet by their IP address, which might look like 192.0.2.166. Since such numbers are hard to remember, the domain name system (DNS) allows mnemonic names (domain names) to be associated with IP addresses. A host wishing to make a connection first looks up the IP address for a given name, then sends packets to this IP address. For example, the Uniform Resource Locator (URL*) www.example.com/ page.html* contains the domain name ''*www.example.com.*'' The computer that performs the domain-name-to-IP-address lookup is known as a DNS resolver, and is commonly operated by the Internet service provider (ISP) the company providing the user with Internet access.

During connection establishment, there are several different ways in which the process can be interrupted in order to perform censorship or some other filtering function. The next section describes how a number of the most relevant filtering mechanisms operate. Each mechanism has its own strengths and weaknesses and these are discussed later. Many of the blocking mechanisms are effective for a range of different Internet applications, but in this chapter, concentration was make on access to the Web, as this is the current focus of Internet filtering efforts.

The goals of deploying a filtering mechanism vary depending on the motivations of the organization deploying them. They may be to make a particular Web site (or individual Web page) inaccessible to those who wish to view it, to make it unreliable, or to deter users from even attempting to access it in the first place. The choice of mechanism will also depend- upon the capability of the organization that requests the filtering where they have access to, the people against whom they can enforce their wishes, and how much they are willing to spend. Other considerations include the number of acceptable errors, whether the filtering should be overt or covert, and how reliable it is.

The most important mechanisms available to implement a filtering regime are: TCP/IP Header Filtering, TCP/IP Content Filtering, and HTTP Proxy Filtering

Each mechanism has different properties of who can deploy systems based around them, what the cost will be, and how effective the filtering is. These properties are Positioning of System and Scope of Blocking, Error Rate[5], Detectability, Circumvent-ability, Reliability, Cost and Speed, and Insertion of False Information.

There are three basic approaches to filtering Internet content:

a) Allowing through known ‘good’ content (inclusion filtering)
b) Blocking known ‘bad’ content (exclusion filtering)
c) Examining content and blocking when it fails acceptability tests

A policy is a deliberate system of principles to guide decisions and achieve rational outcomes. A policy is a statement of intent, and is implemented as a procedure or protocol. Policies are generally adopted by the Board of or senior governance body within an organization whereas procedures or protocols would be developed and adopted by senior executive officers. Policies can assist in both subjective and objective decision making. Policies to assist in subjective decision making would usually assist senior management with decisions that must consider the relative merits of a number of factors before making decisions and as a result are often hard to objectively test e.g. work-life balance policy. In contrast policies to assist in objective decision making are usually operational in nature and can be objectively tested e.g. password policy.

The term may apply to government, private sector organizations and groups, as well as individuals. Presidential executive orders, corporate privacy policies, and parliamentary rules of order are all examples of policy. Policy differs from rules or law. While law can compel or prohibit behaviors (e.g. a law requiring the payment of taxes on income), policy merely guides actions toward those that are most likely to achieve a desired outcome.

Policy or policy study may also refer to the process of making important organizational decisions, including the identification of different alternatives such as programs or spending priorities, and choosing among them on the basis of the impact they will have. Policies can be understood as political, management, financial, and administrative mechanisms arranged to reach explicit goals.

A variety of organizations, institutions, companies, and countries seek to restrict Internet access from within their premises and territories. For example, companies may seek to improve employee productivity by restricting access to leisure sites; libraries and schools may seek to avoid exposing children to sexually-explicit content, or be required to do so; countries may seek to control the information received by their citizens generally. Common among nearly all these applications is the public unavailability of the filtering lists that, by the design of filtering systems, users cannot and do not know the set of specific sites blocked. In some cases users might ask for a specific site and be told of its unavailability due to filtering, but in other cases such unavailability may be conflated with unremarkable network blockages a Web site might be unreachable for any number of reasons and the failure to view it at a particular moment cannot reliability be attributed to active filtering?

Literature review and discussion of previous studies, which are relevant to the research are: Ignoring the Great Firewall of China[1], Government Mandated Blocking of Foreign Web Content[2], Failures in a Hybrid Content Blocking System[3], Web Sites Sharing IP Addresses: Prevalence and Significance[4], Internet Filtering Accuracy Review[5,12], Benjamin Edelman[6], On Anonymity in an Electronic Society: ACM Computer. Survey[7], Online Contribution Practices in Countries that Engage in Internet Blocking and Censorship Human Factors[8], Faulty filters[9,10], Image of Internet Police: Jingjing and Chacha Online[11], and Tools and Technology of Internet Filtering[12].

From the literature review surveyed, it is observed that the error rate in most of the filtering regimes is presented due to the fact that many of the filtering product are based on exclusion filtering or/and content filtering approaches. Exclusion filtering is based on black lists (or blocking lists) of known objectionable sites and is a more common. Exclusion filtering adopts the “innocent until proven guilty” philosophy and allows through all content it does not know to be unacceptable. Sites and content that have not yet been classified by the product vendor pass through the filter, and the sheer scale of the Internet means that much “undesirable” content could pass through unscathed. These products rely on their black lists being comprehensive and up-to-date. Vendors of exclusion filtering products cannot guarantee that their users will not come across “undesirable” content, but they can block much such material, particularly if it is coming from well-known and established sites.

RESEARCH METHODOLOGY

The proposed Firewall (*SmoothWall Express*) is an open source Firewall distribution based on the GNU/Linux operating system, designed for ease of use, is configured via a web-based GUI. The *SmoothWall Express* enables users to easily build a firewall to securely connect a network of computers to the internet.

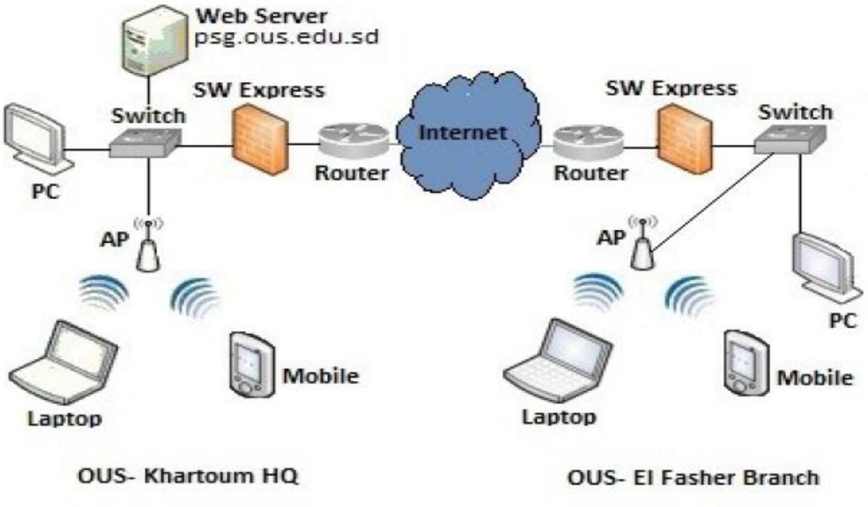

Fig 1: SmoothWall Express Firewall

Almost any Pentium-class PC can be used, for example, an old low specification PC long redundant as a user workstation or server. The *SmoothWall Express* creates a dedicated hardware firewall, offering the facilities and real security associated with hardware devices.

The *SmoothWall Express* comes pre-configured to stop all incoming traffic that is not the result of an outgoing request. The rules file that implement this policy are part of the system configuration and should not normally be edited by other than the configuration procedure.

Implementation of Smoothwall does optimize centralized management and monitors such as: Logging incoming and outgoing traffics such as URL, Assigned addresses, Device privilege, Simplified GUI interface usability, IP address ownership and location, Web Proxy, IP tools such as Pinging, Trace Routes, SSH, Remote login, VPN connections, Network performance attribute such as bandwidth, Network Intrusion Detection, Access Control List (ACL), Patches for Fixes, updates and previous installation history on the dashboard, Backup, Dashboard language preference, Email protocol such as POP3, QOS services, and DHCP service for LANs.

Requirements may vary depending on traffic throughput and processing requirements, which themselves vary according to the number and size of protected networks.

The IP address that the *SmoothWall Express* NIC will use is required in this test for example: 192.168.72.142

The network mask used in conjunction with the IP address to define the network that this NIC belongs to, for example: 255.255.255.0 is also necessary for the experiment, beside the Password with some restrictions to be applied.

The admin account is used to access *SmoothWall Express* via a web browser and carry out routine configuration and management.

Designing Policies to Smoothwall Express

In the Default security policy dialog box there is three policies available as follows:

- **Open:** *SmoothWall Express* allows all outgoing requests.
- **Half-open:** The default policy, *SmoothWall Express* allows most outgoing requests and blocks potentially harmful requests.
- **Closed:** *SmoothWall Express* blocks all outgoing requests. Anything to be allowed must be explicitly enabled.

Taking into account that only one website (*www.psg.ous.edu.sd*) is allowed to access, so the **Closed** option suits our requirements which it represents an inclusion filtering (white list) approach. Therefore, the policy that should be selected at the time of installing *Smoothwall Express* is **Closed** (the default policy is half-open). In a later stage, we shall explicitly enable access to only *www.psg.ous.edu.sd*.

Comparison between *SmoothWall Express* and Monowall

Monowall[14] is a complete embedded firewall software package that, when used together with an embedded PC, provides all the important features of commercial firewall boxes (including ease of use) at a fraction of the price (free software).

Monowall is designed to run on a 16MB flash card, and it has the smallest footprint of the firewalls[15]. Because of this, Monowall only provides the bare bones features for a firewall. Still, given it's so small, it's a rather impressive distribution.

RESULTS AND DISCUSSION

Smoothwall Express and Monowall have been installed and tested to see how effectively it carried out the task of filtering. The assessment of ease of use was largely subjective, based on experiences in installing, using, and de-installing the products under test. Relative performance scores were given for both *Smoothwall Express* and Monowall, such as ease of installation. The numeric scores are out of a maximum of 10, with higher numbers indicating a better performance. Other capabilities were just checked and given a yes/no score. No attempt was made to combine these subjective ratings to give an overall score.

Filtering effectiveness was tested by installing the *Smoothwall Express* and Monowall under test and then attempting to access all of the Web pages on the test list. This list includes 14 sites, and includes both some of sites that could be expected to be blocked and site that should be passed through (http://psg.ous.edu.sd). The test were evaluated and scored according as in table (1)

Determining Effectiveness

The effectiveness tests aim to determine how well the *Smoothwall Express* and Monowall does its job. A truly effective filter meets the following criteria:

i. It blocks all undesirable Internet content.
ii. It allows access to http://psg.ous.edu.sd.
iii. It securely tracks all attempts to access undesirable content.

The first two criteria were tested using carefully constructed lists of Web containing both sites with content that could be expected to be blocked (Top 10 search engines, skype.com, facebook.com, and youtube.com) and site with content that should pass through (psg.ous.edu.sd). Thus, the test list was as follows:

https://www.google.com; https://www.bing.com; https://www.yahoo.com; https://duckduckgo.com; https://archive.org; https://www.youtube.com; https://www.facebook.com; http://www.skype.com; http://www.ask.com; http://www.aol.com; http://www.wolframalpha.com; http://www.chacha.com; http://uk.ask.com; http://psg.ous.edu.sd

Evaluations

Smoothwall Express and Monowall were installed then evaluated for both usability and the results of the effectiveness tests.

The result of the effectiveness tests (Fig(1) and Fig(1)) pointed out to some weaknesses on Monowall, when identification to any other website failed; given that the Internet consists of tens of billions of pages with millions more added every day, so undesirable content will still be accessible.

Consequently, one can say that *Smoothwall Express* suits the requirement of denying access to the content of all websites except the exam site through applying appropriate policies which enable us to achieve our objectives successfully.

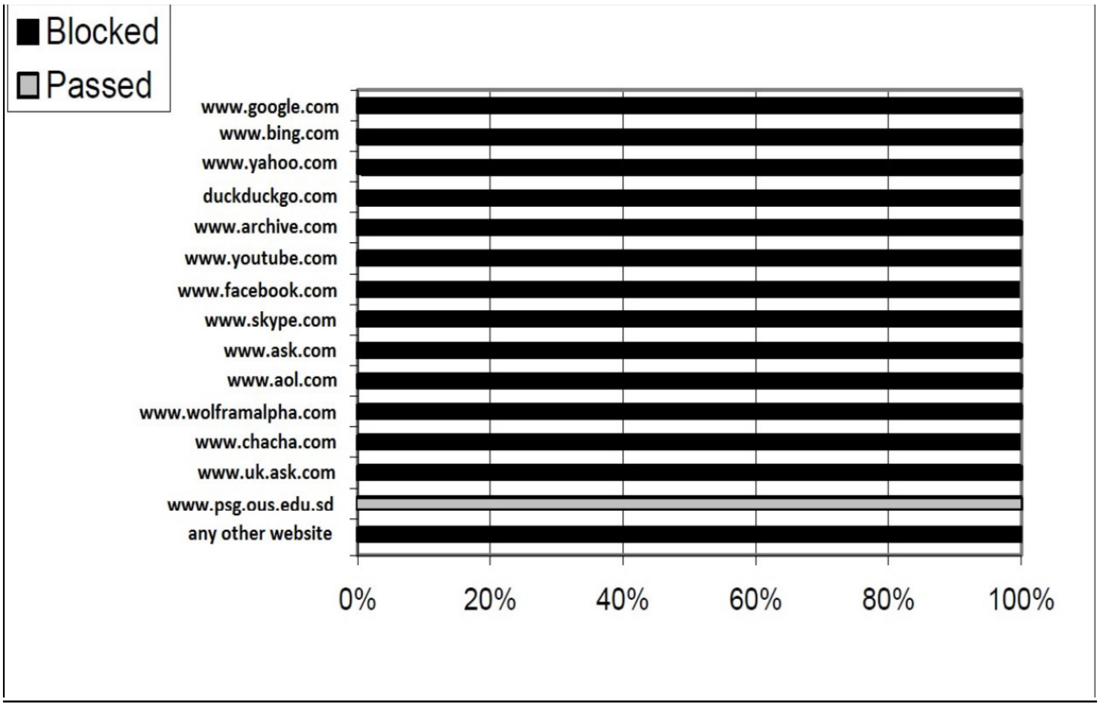

Fig 4.1: Smoothwall Express Effectiveness

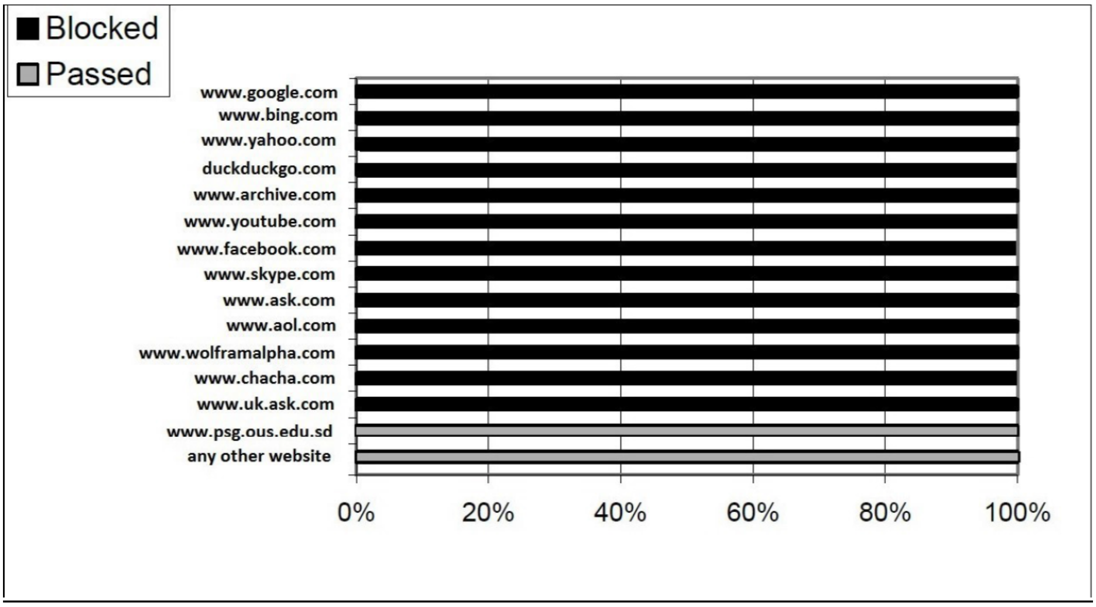

Fig 4.2: Monowall Effectiveness

Table 4.1: Smoothwall Express vs Monowall Usability Assessment

Category	***Score or Result-*** *Smoothwall*	***Score or Result-*** *Monowall*
Ease of Installation		
Easy to install	9	8
Easy to uninstall	10	10
Ease of Configuration		
Simple configuration	10	10
Flexible	10	3
Detrimental to system stability?	NO	NO
Ease of Use		
Easy to use	10	10
Side effects	10	10
Documentation		
Troubleshooting and support	10	10
Claims/Capabilities		
URL blocking	Yes	NO
IP blocking	Yes	Yes
Tracking all access?	Yes	Yes

Based on usability assessment (Table(1)) it's found that there is a big difference between the *Smoothwall Express* and Monowall in respect of flexibility, due to the fact that Monowall needs to identify all possible addresses of the websites that students may try accessing during exam hours, in order to search for an answer. On the contrary, *Smoothwall Express* has the feature of blocking all outgoing packets, whilst anything to be allowed must be explicitly enabled.

At the end of the day, locking down a network is a trade-off between flexibility and control. Wherever you choose to draw the line, it is important to be aware of the gray areas in your network traffic, the general level of technical expertise amongst your users, and how life can be made more difficult for those seeking to undermine acceptable usage policies.

CONCLUSION

The aim of this research was to design a web filtering policy, which will assist in ensuring that an exam results will reflect the actual student's effort? This has been achieved through using

Smoothwall Express that has the capability of denying access to the content of all websites except the exam site.

REFERENCES

[1] Richard Clayton, Steven J. Murdoch, and Robert N. M. Watson, ‘‘Ignoring the Great Firewall of China’’ in 6th Workshop on Privacy Enhancing Technologies (Cambridge, England, 28–30 June 2006).

[2] Maximillian Dornseif, ‘‘Government Mandated Blocking of Foreign Web Content,’’ in Security, E-Learning, E-Services: Proceedings of the 17, ed. J. von Knop, Haverkamp, and E. Jessen (DFN-Arbeitstagung u¨ber Kommunikationsnetze, 2003).

[3] Richard Clayton, ‘‘Failures in a Hybrid Content Blocking System,’’ in Fifth Workshop on Privacy Enhancing Technologies (Dubrovnik Cavtat, Croatia, 30 May–1 June 2005).

[4] Ben Edelman, ‘‘Web Sites Sharing IP Addresses: Prevalence and Significance,’’ Berkman Center for Internet and Society (September 2003).

[5] Internet Filtering Accuracy Review, CERTUS Consulting Group LLC, Oct. 2001

[6] Jonathan Zittrain and Benjamin Edelman, “Empirical Analysis of Internet Filtering in China”, Berkman Center for Internet & Society, Harvard Law School, November, 2002

[7] Edman, et al, “On Anonymity in an Electronic Society: ACM Computer Survey”, 2009, vol. 42(1), Article 5

[8] Shklovski, et al., “Online Contribution Practices in Countries that Engage in Internet Blocking and Censorship Human Factors”, 2011, pp.1109-1118.

[9] Faulty filters. EPIC, Dec, 1997

[10] http://ncac.org/resource/internet-filters/ (accessed December, 2015).

[11] Xiao Qiang, ‘‘Image of Internet Police: Jingjing and Chacha Online,’’ China Digital Times”, http://chinadigitaltimes.net/2006/01/image_of_internet_police_jingjing_and_chacha_online_hon.php (accessed February 19, 2007).

[12] Steven J. Murdoch and Ross Anderson, “Tools and Technology of Internet Filtering”, www.sec.cs.ac.uk (accessed December, 2015).

Websites

[13] http://www.smoothwall.net (accessed December, 2015).

[14] mOnOwall.ch (accessed December, 2015).

[15] http://www.techradar.com/news/software/applications/7-of-the-best-linux-firewalls-697177/2 (accessed December, 2015).