BORANG PENGESAHAN STATUS TESIS*

JUDUL: INTERNET FILTERING PROXY VERSION 3

SESI PENGAJIAN: 2007/2008

Saya NEO LEE SIEN

(HURUF BESAR)

mengaku membenarkan tesis (PSM/Sarjana/Doktor Falsafah) ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

_____	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
_____	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
/	TIDAK TERHAD	

(TANDATANGAN PENULIS)

Alamat tetap: NO 2, JALAN NASIB 3, TAMAN NASIB, 86000 KLUANG, JOHOR

Tarikh : 9/11/07

(TANDATANGAN PENYELIA)

ROBIAH BINTI YUSOF

Nama Penyelia

Tarikh : 9/11/07

CATATAN: * Tesis dimaksudkan sebagai Laporan Projek Sarjana Muda (PSM)

** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

INTERNET FILTERING PROXY VERSION 3

NEO LEE SIEN

This report is submitted in partial fulfillment of the requirements for the Bachelor of Computer Science (Computer Networking)

FACULTY OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
UNIVERSITI TEKNIKAL MALAYSIA MELAKA
2007

DECLARATION

I hereby declare that this project report entitled

INTERNET FILTERING PROXY VERSION 3

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : _____ Date: 9/11/07

(NEO LEE SIEN)

SUPERVISOR : _____ Date: 9/11/07

(PUAN ROBIAH BINTI YUSOF)

DEDICATION

Specially dedicated to my beloved parents,
Mr. Neo Theng and Mrs. Tan Ah Eng

For my lecturer and supervisor, Madam Robiah Binti Yusof at Universiti Teknikal Malaysia Melaka (UTeM)

And last but not least to all my loving friends and siblings who have encouraged, guided and inspired me throughout my journey of education

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to express my deepest gratitude to Madam Robiah Binti Yusof, my supervisor for her invaluable guidance and constructive suggestions and advices throughout this project which really help me progress through out the project. Thanks for all the information and consideration.

Here, I would also like to express my sincere thanks to all my course mates, friends and others colleagues for their support and help in one way or another. In addition, I also wish to express my deepest gratitude to the lecturers and tutors who help me a lot when I faced the problems during development of Internet Filtering Proxy system.

Last but not least, I wish to express my deepest appreciation and heartfelt thanks to my beloved family for their understanding, motivation, support and sacrifices so that I attend and succeed in this project.

ABSTRACT

This project paper is about the Internet Filtering Proxy Version 3 for the person who concerned about the quality of the information that children or user access when using the Internet. This system is enhancing from previously student system. This system can be used by an administrator or parent to monitor the internet usage and cyber activities. The most effective way to handle this concern is by closely monitoring the websites that users use. The main purpose of this system is to add additional function which is filter based on combination two categories and start the proxy interface and MySQL database automatically. This system is not only monitoring the websites that surfed but also can block the websites that are inappropriate for users. All the information about the blacklist websites that need to filter will be saving in the MySQL database. To develop this system, a research from current and previous system is crucial to the developer. This system used Prototyping Methodology as a project reference process flow. This Internet Filtering Proxy Version 3 is build to solve the current problem of previous versions and can prevent the user from accessing inappropriate material on the Internet.

ABSTRAK

Projek ini adalah mengenai Internet Filtering Proxy Version 3 untuk orang yang ingin memastikan kualiti penggunaan Internet yang diguna oleh pengguna ataupun kanak-kanak. Sistem ini diperbaiki daripada sistem yang telah dibangunkan oleh pelajar sebelum ini. Sistem ini boleh digunakan oleh pentadbir atau ibu bapa untuk memantau penggunaan Internet. Cara yan paling efektif untuk menghindari pengguna daripada melayari laman web yang tidak wajar ialah memantau apa yang telah mereka lakukan semasa menggunakan Internet. Tujuan sistem ini ialah untuk melakukan penambahan fungsi baru untuk proses penapisan ke atas penggunaan Internet. Penambahan baru bagi fungsi sistem ini ialah melalui penggunaan teknik penapisan yang berdasarkan gabungan dua kategori dan penambahan fungsi untuk memulakan proxy interface dan MySQL secara automatik. Segala maklumat yang dimasukkan oleh pengguna akan disimpan di dalam pangkalan data. Sistem ini juga boleh menghalang pengguna dari melayari laman web yang tidak wajar. Prototyping Methodology telah digunakan sebagai panduan untuk membangunkan sistem ini. Sistem ini dibangunkan untuk menyelesaikan masalah yang ada dalam sistem dahulu.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DDR-RAM	-	**Double Data Rate Random Access Memory**
DFD	-	**Data Flow Diagram**
FRS	-	**Functional Requirement Specifications**
IE	-	**Internet Explorer**
IP	-	**Internet Protocol**
JDK	-	**Java Development Kit**
JRE	-	**Java Runtime Environment**
LAN	-	**Local Area Network**
NIC	-	**Network Interface Card**
RAM	-	**Random Access Memory**
SRS	-	**System Requirement Specifications**
SQL	-	**Structured Query Language**
TCP/IP	-	**Transmission Control Protocol/Internet Protocol**
URL	-	**Uniform Resource Locator**
UTP	-	**Unshielded Twisted Pair**
WWW	-	**World Wide Web**

LIST OF ATTACHMENTS

CHAPTER I

INTRODUCTION

The first chapter of this report will discuss about the background, problem statements, objectives, scopes, significances and expected output for the proposed project. This "Internet Filtering Proxy Version 3" is basically developed by Java language and MySQL. Administrators or parents can use this filtering system to prevent children from accessing inappropriate materials. The system can filter the web sites based on IP Address, port, keyword, URL and combination of two categories.

1.1 Project Background

The Internet is one of the greatest inventions of all times. However, the Internet today contains an enormous amount of information that can be offensive to adults, and unsuitable and sometimes downright dangerous to children. Therefore the network administrators or parents must play the active role to restrict their children from surfing those inappropriate websites and keep them use the Internet wisely.

Internet Filtering Proxy monitors the entire Internet traffic and blocks the access to websites according to the chosen filter. The requested web pages are displayed only if they comply with the specified filters. As a result, Internet filter proxy software gives network administrator or parents the ability to control content displayed, block websites by category and set up admin passwords. Through this,

network administrator and parents can control and monitor their children's cyber activities.

The current Internet filtering proxy built by previous student still have some weakness, such as cannot filter based on combination of two categories, cannot start database and proxy interface automatically, and system lagging while handle a lot of surfing activities since the databases system was build with MS Access. Therefore some enhancements will be done from the previous student's software by adding some new functions, which are filter the websites based on combination techniques, start database and proxy interface automatically and build the databases system with MySQL.

This enhanced software not only given the network administrators or parents the ability to monitor the websites that the users visited, but it also blocks the websites that are inappropriate for children especially pornographic sites. This software can block the inappropriate websites based on port number, URL, IP address, keyword or combination of two categories as its additional feature where it's interface and MySQL database can run automatically when the computer startup.

1.2 Problem Statement(s)

Today the highest number of daily search engine requests is the pornographic site and there is 4.2 millions of pornographic websites allocated in the Internet (12% of total websites). This means that most of the internet users use this technology to surfing these pornographic websites. This global phenomenon had cause the tension of most of parents nowadays while their children run the cyber activities. Therefore it is important to have the internet filtering software to control and prevent the users from accessing these materials.

Since there are many inappropriate websites over the Internet, therefore we cannot assume the users remember all these websites address. As the additional functions in this system, users can filter the websites based on combination of two

categories besides IP address, keyword, URL and port number. This ways let the users easier to filter the websites that want instead of remember their websites address.

The existing Internet filtering proxy system occurs a lagging problem while handles a lot of web traffic. The system cannot run its function properly after the users browse the websites few times. As a result, this system needs to restart again to enable it run its functions properly. Due to this problem, the system is not an efficient and effective Internet filtering proxy system.

For the current system that done by previous students, there is still got some weakness. One of the problem is the Internet Filtering Proxy interface and its database system cannot run automatically while the computer startup. As a result the system cannot work its functions properly to monitor and block the websites that users surf until they started the interface and database manually. However, it is important to make sure the system's interface and database can run automatically while the computer startup to protect and prevent the users from surfing the inappropriate websites all the time.

1.3 Objectives

The objectives of this project are to make some improvement from the previous projects, which are:

(1) **To integrate the existing system with the combination filtering technique.** Since the existing system cannot filter the website based on the combination of two categories, therefore the additional functions let the users set the websites they need to filter more specific and all of this information will be saved in the databases.

(2) **To develop the MySQL database for the system.**

All the information about the blocker URL, port number, IP address and keyword will be saving in the database. MySQL database will be developed to handle more databases and prevent the system from lagging.

(3) **To enable the proxy interface and MySQL database start automatically.**

The system's proxy interface and its database can run automatically when the Windows startup instead of start them manually.

1.4 Scopes

The scope of this project is use to make some enhancement from previous project that acts as the web filter and web cache proxy in the Internet. This system will have the filtering function to filter all the inappropriate websites based on combination techniques as additional features. Administrators can clearly monitor the websites that surf by the users and specify the websites as black list in this application. The website that classified in the black list will saved in the database system and will be use to compare with the user's requested website. The requested web pages will be displayed only if they comply with the specified filter list.

The targeted users for this project will be the parents or network administrators to monitor and control their children from surfing the inappropriate websites. This proxy system is purposely designed for the usage of the home use. Therefore, the targeted users can deploy this application to their home environments.

Beside that, this system will be developed in Java language for the part of scripting and MySQL for the part of database. The Java scripting integrates with MySQL database system can be used to handle a large scale of users and databases to prevent the system from lagging. The completed proxy system also will be tested by using two types of web browsers, which are Internet Explorer and Mozilla Firefox.

Lastly, this system's proxy interface and database must be start automatically when Windows startup instead of start them manually. As a result, this feature let the proxy system run their functions effective and efficient. The users can using this proxy system in monitor and restrict their loved children to surfing the inappropriate materials.

1.5 Project Significance

This proposes project has its significance and advantages. This project will be developing to be one of the Internet filtering proxy applications that user friendly, easy to use and easy to install for every stages of the age. This project is not only designed for the expert administrator use, but it also suitable for the use of beginner administrator and parents who is not very expert in the area of computer and network administration.

The project is bringing a lot of the benefits to the users especially for those who are very concerned about the impact of the Internet contents. This category of users may be the parents as well as educators or network administrator that responsible to assure the quality of their children cyber activities. This system will acts as assistance to parents in monitor and control their children from using this technology into the inadequate way. This project is importance to every user especially the parents in the process of guiding their children in learning using Internet since the Internet today is full of the inappropriate websites.

Beside that, this system is also has the filtering function to block the inappropriate websites based on the chosen filter. This system allows the users to filter the unwanted websites not only based on one category separately but also can based on any combination of two categories. As a result, the users can classify the inappropriate websites as the black list in more specific ways. The classified websites will saved in the database system and will be use to compare with the user requests all the time.

1.6 Expected Output

This project has some expected outputs when it is finished developing. This system will be able to filter the inappropriate websites based on any functions of the current system and also the combination of two categories in this system. So this system allows the users to filter unwanted websites according to the more specific categories.

The proxy interface and its MySQL database system will run automatically when the computer startup compare to the previous system that need to start manually. It is very important to make sure that the system interface can be running in automatic way when the computer startup since sometimes the user maybe forget to compile and start the system and it will not be able to run the its functions properly.

1.7 Conclusion

As a conclusion, this proposes project is an enhancement work from the previous student project. Filter based on the combination techniques is the additional function in this system which will work together with the current system. Therefore this project will make sure that all the additional functions can function properly. The additional function is to make sure the system can filter the websites more specify and perfectly when the users using the web browser.

The weaknesses that cover in this project is to make the proxy interface and its MySQL database system can run automatically when the computer started. This enhancement makes the existing system become more powerful and useful in aspect of network security.

The next chapter of this report is Literature Review and Project Methodology. This chapter is important in this project since it acts as a guideline for the developer to develop this project. The method of develop this project are consists compatibility of step development that will be done from preliminary until the end of this project.

CHAPTER II

LITERATURE REVIEW AND PROJECT METHODOLOGY

This chapter will discuss about the Literature Review and Project Methodology of Internet Filtering Proxy Version 3. The literature review will be researched after the scope and objective had been determined. Moreover, this chapter also describe about the requirements and methodology for the project.

2.1 Introduction

This chapter will discuss and study more about Internet Filtering Proxy issues and development process includes the development tools and requirements that used to improve this project and the system features. To develop a successful project, it needs a lot of information, facts and research about the project as a guideline. That information is used to make sure that the project will be execute properly all the time and the project result will be delivery on time. The research is concentrated on title, objectives, problem statement, scopes and system requirement of this proposes project.

All of the needed information and facts can be gathering from many resources, such as books, journals and articles from the internet and also study about existing product. The related information that collected can be use as a guideline and reference to develop this project. Beside that, the facts that gathered also can be use to support and make sure the project is workable and beneficent.

prevent its employees from accessing a specific set of Web sites. Beside that, proxy technology is often seen as an alternative way to provide shared access to a single internet connection.

2.2.2 Filtering Combination Function

The best filter programs will use a combination of filtering techniques, including URL filtering, keyword filtering, IP address filtering and port number filtering to filter the undesirable Internet content. Moreover the truly effective Internet blocking and filtering software is able to filter the website based on all the possible ways that Internet content can be distributed.

A Web page is not a single monolithic entity, but is an entity that composed of a number of independent components; each of them will with their own URL, which is fetched separately and independently by the web browser. Each of these components is directly accessible through its own URL and may also be a candidate for filtering. For example, a filter may block access to *http://www.playboy.com/* but may not block access to pictures that are used on the Playboy home page, such as *http://www.playboy.com/sex/dec99.gif*. However the current system of Internet Filtering proxy overcome this kind of problem in filtering function, therefore combination of URL and keyword filtering technique not need to apply in this project.

In fact, the domain names that used in URLs are actually translated to numeric IP addresses before they are used to access its content. Users can always use the numeric form of a URL to access the inappropriate sites that they want, however this way may deceive a filter that just looking for the matched domain name form. For example, the URLs http://www.playboy.com/centrefolds/ and http://206.251.29.10/centrefolds/ are completely equivalent and interchangeable. As a result, the combination of URL and IP address filtering techniques is not suitable to use since domain name and IP address are interchangeable.