

TESIS^ APPROVAL STATUS FORM

JUDUL: WEB BASED HELPDESK WITH SPAM FILTERING

SESI PENGAJIAN: 2006

Saya ROJMAN B. SNIKON

(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 dan projek adalah hakmilik Kolej Universiti Teknikal Kebangsaan Malaysia.
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)


(TANDATANGAN PENYELIA)

Alamat tetap : 149, KG. HILIRAN MASJID,
21000, K. TERENGGANU

Tarikh : 20/11/2006

ZURINA SAAYA
Nama Penyelia

Tarikh : 20/11/2006

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

^ Tesis dimaksudkan sebagai Laporan Projek Sarjana Muda (PSM)

raf

TK5105.743 R67 2006



0000038844

Web-based helpdesk with spam filtering / Rosnan Snibon.

WEB-BASED HELPDESK WITH SPAM FILTERING

ROSNAN BIN SNIBON

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

**FACULTY INFORMATION AND COMMUNICATION TECHNOLOGY
KOLEJ UNIVERSITI TEKNIKAL KEBANGSAAN MALAYSIA**

2006

DECLARATION

I hereby declare that this project report entitled

WEB BASED HELPDESK WITH SPAM FILTERING

is written by me and is my own effort and that no part has been plagiarized

without citations.

STUDENT :  Date : 20/11/2006
(ROSNAN BIN SNIBON)

SUPERVISOR :  Date : 20/11/2006
(MISS ZURINA SAAYA)

DEDICATION

To all who helped me through the completion of this thesis..

ACKNOWLEDGEMENTS

First and foremost, I would like to show my deepest appreciation to my project supervisor, Cik Zurina Saaya, for guiding me through out the process of completing this PSM project.

Gratitude is to be expressed to the PSM committee members for providing guidelines and workshop in the aid of producing a report with quality and standards.

Last but not least, I would as well like to thanks my family and my fellow mates for the encouragement and useful thoughts that has been selfishly poured out to me through out the process of producing this report. Without the help from all, the completion of this report will be very much affected.

ABSTRACT

Web-Based Helpdesk with Spam Filtering is a relatively new enhancement of helpdesk system. Purpose of this project is to replace the traditional helpdesk technology which is using telephone technology. This project completed with spam filtering in order to prevent unwanted report (spam) from interrupt this application. It is a security part for this application. The main activities for this project are the users make a problem report to admin and admin will reply the solution to the users. This project use blacklist and white list filtering technique which is the report from the users will be filtered based on setting at blacklist and white list keyword. Only genuine report will be replied by admin. For the bulk (spam) report it will be deleted. The genuine and bulk (spam) report are stored in separate table. So, it will make the admin easy to know which report will be served. With this application or system, it can help the company in order to provide helpdesk services for their customers in an efficient way.

ABSTRAK

Web-Based Helpdesk with Spam Filtering merupakan satu pembaharuan baru berasaskan pada *Helpdesk* sistem. Tujuan projek ini adalah untuk menggantikan tradisional *helpdesk* iaitu dengan menggunakan teknologi telefon. Projek ini dilengkapi dengan *spam filtering* yang bertujuan untuk mengelakkan laporan yang tidak diperlukan (*spam*) daripada mengganggu sistem ini. Ia merupakan salah satu bahagian keselamatan untuk aplikasi ini. Antara aktiviti utama dalam projek ini adalah pelanggan membuat laporan berkenaan masalah yang dihadapi dan pihak *admin* akan memberi penyelesaian berdasarkan masalah yang diajukan. Projek ini menggunakan teknik penapisan jenis senarai hitam dan senarai putih dimana laporan daripada pelanggan akan ditapis berdasarkan aturan yang terdapat pada senarai hitam dan putih. Pihak *admin* hanya akan melayan laporan yang sah sahaja. Oleh itu, laporan yang tidak diperlukan (*spam*) akan dihapuskan. Tempat penyimpanan laporan bagi laporan yang sah dengan laporan yang tidak diperlukan (*spam*) adalah berbeza. Ini memudahkan pihak pengurusan syarikat untuk mengetahui laporan mana yang harus dilayan. Aplikasi atau sistem ini akan membantu syarikat didalam menyediakan perkhidmatan *helpdesk* dengan lebih baik kepada pelanggan.

TABLES OF CONTENT

CHAPTER	SUBJECT	PAGE
	ACKNOWLEDGEMENTS	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	xii
	LIST OF FIGURES	xiii
	LIST OF APPENDICES	xiv
CHAPTER I	INTRODUCTION	
	1.1 Project Background	1
	1.2 Problem Statements	2
	1.3 Objective	2
	1.4 Scope	3
	1.5 Project Significance	3
	1.6 Expected Output	4
	1.7 Conclusion	4
CHAPTER II	LITERATURE REVIEW AND PROJECT METHODOLOGY	
	2.1 Introduction	5
	2.2 Fact and Finding	6
	2.2.1 Spam Filtering	6
	2.2.2 Web-based Helpdesk Application	8

2.3	Project Methodology	10
2.3.1	Planning	11
2.3.2	Analysis	11
2.3.3	Design	11
2.3.4	Implementation	12
2.4	Project Requirement	12
2.4.1	Software Requirement	12
2.4.2	Hardware Requirement	13
2.5	Project Schedule and Milestones	13
2.6	Conclusion	14
CHAPTER III	ANALYSIS	
3.1	Introduction	15
3.2	Problem Analysis	16
3.2.1	Analysis of Current System	16
3.2.1.1	Business Flow	16
3.2.1.2	Description of Current Situation/Scenario	17
3.3	Requirements Analysis	17
3.3.1	Analysis Of To Be System	18
3.3.1.1	Function Requirement	18
3.3.1.2	System Flow	18
3.3.2	Software Requirement	19
3.3.3	Hardware Requirement	20
3.4	Conclusion	20
CHAPTER IV	DESIGN	
4.1	Introduction	21
4.2	High Level Design	21
4.2.1	Raw Input/Data	22

	5.5	Conclusion	53
CHAPTER VI		TESTING	
	6.1	Introduction	54
	6.2	Test Plan	55
		6.2.1 Test Organization	55
		6.2.2 Test Environment	55
		6.2.3 Test Schedule	56
	6.3	Test Strategy	57
		6.3.1 Classes of Tests	58
	6.4	Test Design	59
		6.4.1 Test Description	59
		6.4.2 Test Data	61
	6.5	Test Case Result	63
	6.6	Conclusion	65
CHAPTER VII		PROJECT CONCLUSION	
	7.1	Observation on Weaknesses and Strength	67
		7.1.1 Weaknesses	67
		7.1.2 Strengths	68
	7.2	Propositions for Improvement	68
	7.3	Conclusion	69
REFERENCES			70
APPENDICES			71

4.2.2	System Architecture	23
4.2.3	User Interface Design	25
	4.2.3.1 Navigation Design	26
	4.2.3.2 Input Design	27
	4.2.3.3 Output Design	28
4.2.4	Database Design	30
	4.2.4.1 Logical Database Design	27
4.3	Detailed Design	28
4.3.1	Software Specification	33
	4.3.1.1 Admin Login	33
	4.3.1.2 Setting Blacklist and White list Keyword	35
	4.3.1.3 View problem report and Reply Solution to users	36
	4.3.1.4 View spam report	37
	4.3.1.5 Client login	37
	4.3.1.6 User Make problem report	38
	4.3.1.7 Customer View Replied Solution From admin	39
	4.3.1.8 User and admin find solution From database	40
4.4	Conclusion	41
CHAPTER V	IMPLEMENTATION	
5.1	Introduction	42
5.2	Software Development Environment setup	43
5.3	Software Configuration Management	44
	5.3.1 Configuration environment setup	44
5.4	Implementation Status	51

LIST OF TABLES

TABLE	TITLE	PAGE
3.1	Software Requirements for This Project	19
3.2	Overall Hardware Requirements for This Project	20
4.1	Raw Input/Data of Web-Based Helpdesk Application	22
4.2	Web-Based Helpdesk Input Design	27
4.3	Web-Based Helpdesk Output Design	29
5.1	Implementation Status	52
6.1	Web-Based Helpdesk Test Environment	56
6.2	Test schedule according tasks, activities and duration to carry out testing activities.	57
6.3	Test design record according purpose, test description and expected result.	59
6.4	Test data record according test cases, test data and expected results.	62

LIST OF FIGURES

FIGURES	TITLE	PAGE
2.1	Every e-mail address has its own content-filter settings.	7
2.2	Users maintain their own whitelists and blacklists.	7
2.4	Interfaces for Administrator Site	9
2.5	Interfaces for Submitting a Trouble Ticket	9
2.6	The phases in the Prototyping Methodology	10
2.7	This Project Schedule and Milestones	13
3.1	Traditional Helpdesk Flow Model	16
3.2	ServiceDesk Plus Flow Model	17
4.1	Clients Server Architecture based on 2-tier Architecture	23
4.2	Web-Based Helpdesk Navigation Design	26
4.3	Web-Based Helpdesk Entity Relationship Diagram	31
4.4	Process Flow for Admin Login Function and Main Menu	34
4.5	Process Flow for Admin setting blacklist and white list Keyword	35
4.5	Process Flow for Admin View problem report and reply solution to user	36
4.6	Process Flow for Admin View spam report from user	37
4.7	Process Flow for User login function	38
4.7	Process flow for user make problem report to admin	39
4.7	Process flow for user view replied solution from admin	40
4.8	Process flow for user and admin solution from database	41

5.1	Software Development Environment for Web-Based Helpdesk with Spam Filtering	43
5.2	How to Install IIS	45
5.3	Set the filter file for php	46
5.4	Set the default file of php	46
5.5	Add the extension file for php	47
5.6	Select a setup type	48
5.7	Select a configuration type	48
5.8	Select a server type.	49
5.9	Select the database usage	49
5.10	Set the networking options	50
5.11	Set the Windows options	50

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
APPENDIX A	DATA FLOW DIAGRAM	71
APPENDIX B	USER INTERFACE DESIGN AND USER MANUAL	74
APPENDIX C	PROJECT GANTT CHART	93

CHAPTER I

INTRODUCTION

1.1 Project Background

A Web Based Help Desk with Spam Filtering is a system act as an assistance resource that troubleshoots problems with computers and similar products. This project is developed for company that offers IT trading and services. A typical help desk has several functions. It provides a central point to receive help on various computer issues. The user notifies the help desk of his or her issue, and the help desk issues a ticket that has details of the problem. If the first level is able to solve the issue, the ticket is closed and updated with documentation of the solution to allow other help desk technicians to reference. If the issue needs to be escalated, it will be dispatched to a second level. The second or higher levels handle more difficult problem that need to assign technician to solve the problem.

The Web Based Helpdesk that has done developed has a function to block spam from interrupt this application. This is one of the security parts. Security is the main issue to be considered when developing a application of web-based helpdesk. Spam is irrelevant or inappropriate messages sent on the internet to a large number of newsgroups or users. The end result from spam is the database bulk with useless content or server will attacked by virus, Trojans and malware. For this application, Black List and White List is used in the filtering function. This technique used to enable administrator to set-up the white list and blacklist keyword from a convenient interface.

By using this method, problem report will be blocked if it contains specific words or combinations of words that are categorized as with spam.

1.1 Problem Statement

Helpdesk has traditionally been used by Call Centers catering for customer enquires on pre or post sale issues. Telephone support was the main media for providing support until the advent of the Internet. The problems from the traditional helpdesk are usually customers get frustrated waiting on hold or navigating the automated telephone answering message tree. Also, when the admin answer phone calls from client, they must explain their problem, often to somebody untrained and the admin unable to resolve the problem immediately. So it will take much time to solve the problem. Another problem is the customer not sure that their problem will be deal with or given the priority it deserves if the telephone messages are left to the admin. Then, additional telephone calls need to be made into the call centre for progress reports if telephone calls are not made back to them. This company also cannot predict the volume of calls, especially during system outages and new product launches. The result by this problem is more staff is needed to support call centre or maybe the company will losing customers. All this adds up to additional costs of providing support through having spare capacity of staff, often under utilized telephone switch systems and associated equipment, phone call charges as well as the potential loss in revenue from unsatisfied customers.

1.2 Objective

The objectives of the application are as the following:-

- i. To develop a web based helpdesk that can operate in fully computing environment.
- ii. To develop a web based helpdesk that can filter problem report received to identify whether it is genuine report or spam by using Blacklist and White list filtering technique.

1.3 Scopes

As we know, every project is focusing on certain scope which has its own limitation or boundaries. The following are scopes for Web Based Helpdesk with Spam Filtering:

- i. This application contains five modules that are search, view, insert, delete and update functions.
- ii. The client side is for user to make problem report to administrator and check solution that was replied from administrator, search the solution from the helpdesk database, delete the previous solution that was own account, change own password and update account.
- iii. The administrator side is for admin check request from client, reply the solution to client, register new client, change own password, add new solution into the database, manage blacklist and white list settings and add new admin.
- iv. Filtering spam from interrupt this application by checking the particular keywords. These keywords can be in any part of the problem report such as the title and the body of the message text. Problem report will be separated with genuine report if it contains specific words or combinations of words that are categorized as spam using Blacklist and White list filtering technique.

1.4 Project Significance

This project will be used by company that offer IT trading and services. It will be company tools to manage maintenance services. This application will available when required because it will run 24 hours a day and 7 days a week. This system only needs one operator to monitor it compare to previous system that needs more than one operator. Therefore labor cost can be decreased. This application is completed with spam filtering that can prevent spam from interrupt this application. Spam will wasting people's time with unwanted message and also eats up a lot of network bandwidth.

1.6 Output

The output of this project is a web-based helpdesk application, which makes use a filtering technique to prevent spam from interrupt this application. When the clients want to make problem report, this application will filter the keywords for the title and description of the report. The keyword will be filtered in two ways. First way is filtered based on blacklist keyword setting and second is based on white list keyword setting. The report will be separated with genuine report if it contains any black list word. The black list and white list word is determined by administrator. The administrator will get two types of report namely, genuine and bulk (spam). Only genuine report will be replied by administrator. For bulk (spam) report, the administrator is not responsible to reply and will be delete them from the database. But, if in the table bulk (spam) have the genuine report, the administrator will move them into the table genuine report.

1.7 Conclusion

The main issues with the development of a web-based helpdesk application revolve around the security risks and measures. The objective of this project is to enhance the security of the web-based helpdesk especially on the spam filtering. Another objective of the project is to develop web based helpdesk that can operate fully computing environment. The next activities are making a literature review and project methodology.

CHAPTER II

LITERATURE REVIEW AND PROJECT METHODOLOGY

2.1 Introduction

This chapter will cover the area of literature review and project methodology. Literature review is the searching, collecting, analyzing and drawing conclusion from all debates and issues in relevant body of literature. For the purpose of this project, literature review will be carried out on two topics namely spam filtering and web-based helpdesk application. Several thesis, report, and journal were read and referred to support the fact and finding. With the literature review being done on the above mentioned topic, the overview of the current scenario for the above mentioned technologies is well studied and hence analysis can be done on those topics. With the understanding from the literature review, reader will be more depth with the purpose and direction of the development of the application.

Project Methodology is a way to use all available techniques, tools and approaches used to achieve the predetermined objectives. In this chapter a suitable project methodology will be chosen to adapt to the development behavior of this web-based helpdesk application. A suitable methodology will ensure the smoothness of the development of the application.

2.1 Fact and Finding

In this section, the fact and finding on the following two topics: Spam Filtering and web-based helpdesk application will be presented.

2.1.1 Spam Filtering

Spam is irrelevant or inappropriate messages sent on the Internet to a large number of newsgroups or users. The end result from spam is an inbox full of useless junk. Leveraging social networks in computer systems can be effective in dealing with a number of trust and security issues. Currently spam filters use the structure and syntax of email body along with training methods to classify email as spam or ham. These include techniques such as word statistics and Bayesian filters [1]. Bayesian filtering was proposed by Sahami et al (1998). The other sample for spam defending technique such as Black list and White list filtering, collaborative feedback based filtering, content-based technique, response based filtering. According to [2], most spam filtering systems use blacklists, where mail from a certain list of email addresses or matching a certain list of text patterns is rejected or otherwise filtered. Refer to [2], the way white list-based filtering works is by create a list of addresses of expecting people to receive mail from, and filter anything that is not from them into a separate low-priority mailbox that will be checked once a week or once a month or something.

To overcome the problem by spam, many ideas were introduced whether by develop software or hardware. Such as by [3]. According to [3], amavisd-new is a mail filter-it receives mail from mail gateway, scans the mail for viruses and spam, quarantines, rejects or discards offending items, and relays the rest to another mail server downstream for delivery. It's designed to let users adjust their content-filter settings and manage their quarantines from a convenient interface. Below are sample screen shoot for Maia Mailguard.

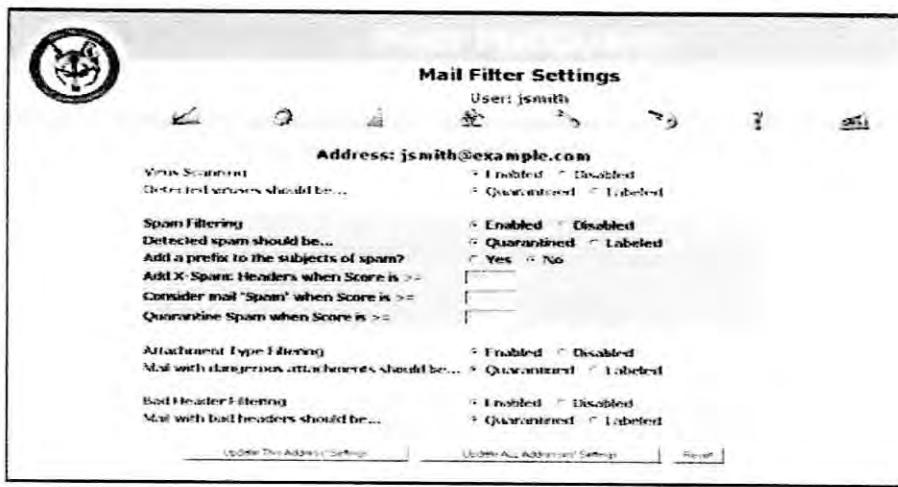


Figure 2.1 every e-mail address has its own content-filter settings.

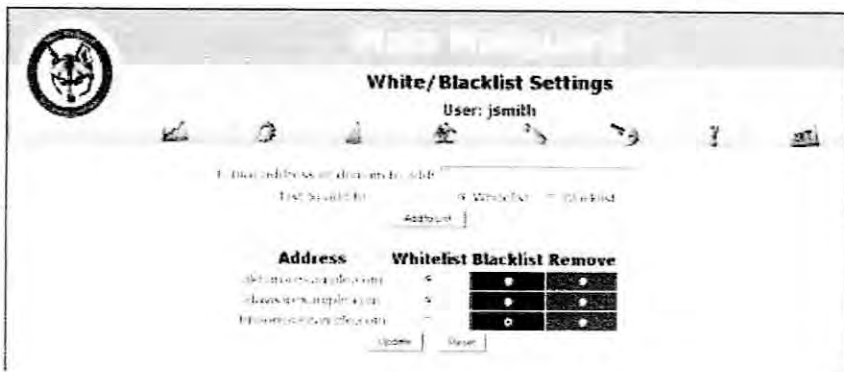


Figure 2.2 Users maintain their own whitelists and blacklists.

According to a recent survey, done by an email marketing service provider, seventeen per cent of permission-based email messages get incorrectly blocked or filtered by the Internet service providers (ISP). Even though never had spammed or have no intention of doing so in the future, legitimate emails might be filtered along with spammers sent emails. It has become important that have a general understanding on some filters currently in use to block the passage of spam emails.

2.1.2 Web Based Helpdesk Application

Web-based Helpdesk is a system act as an assistance resource that troubleshoots problems with computers and similar products. A typical help desk has several functions. It provides the users a central point to receive help on various computer issues. It increase efficiency and productivity among support center by allowing IT technicians, customers, or other desktop support member's connectivity from any network location. Because web interface components are thin-client applications, they can be accessed by a large number of concurrent users from most HTML web browsers.

A web-based system should be easy to use and at the same time the security requirements should not be sacrificed to the ease of use. However, the reality is usually the opposite where security is compromise to ease of use. The security which is desire in a web-based helpdesk application is difficult to be achieved at a reasonable cost. Creating a webpage is cheap, but to ensure the security of it is not. According to [4], an efficient help desk can be a strategic asset that offers tangible value and measurable ROI (Return of Investment).

One of the existing web-based helpdesk application is ServiceDesk Plus. It is a Help Desk module that enables end-users to submit trouble tickets via a browser or through email. It automates several key workflow tasks such as case routing, acknowledging requester/technician notifications and handling of SLA (Service Level Agreement) rules. As a result, support organization can now realize higher levels of productivity, keep everyone informed, provide superior service, improve problem resolution times, share knowledge effectively and keep end-users happy. ServiceDesk Plus being a 100% web-based Help Desk solution that enables end users to submit requests report from anywhere, anytime using just a browser. Below are sample interface for ServiceDesk Plus.

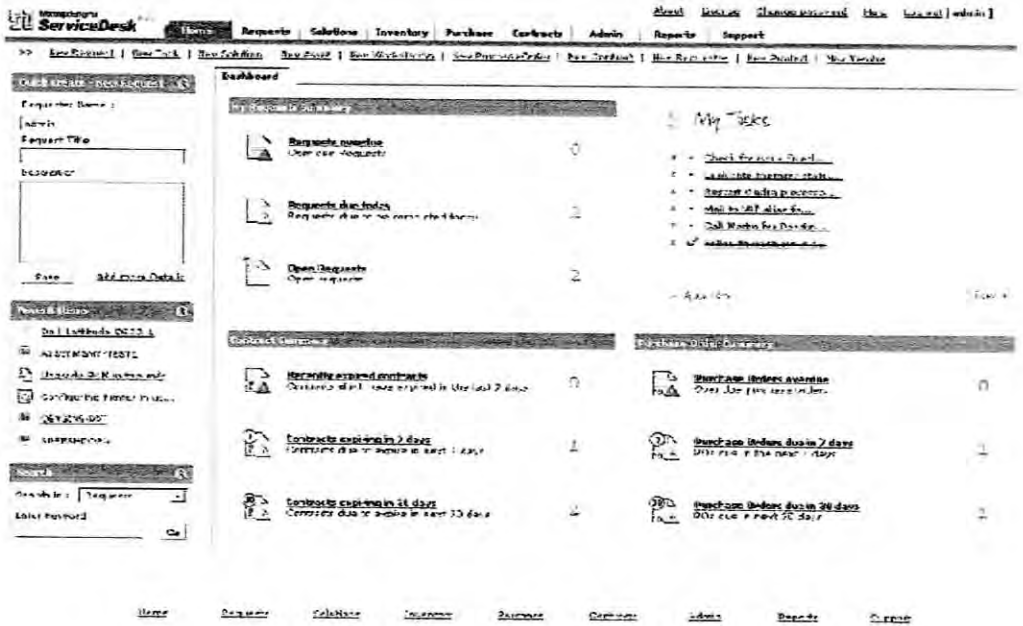


Figure 2.4 Interfaces for Administrator Site



Figure 2.5 Interfaces for Submitting a Trouble Report

2.2 Project Methodology

For the development of this web-based helpdesk application which is trying to make use of the content filtering concept for Spam Filtering. The prototyping