

raf

TK5105.875.I57 .M42 2006



0000039052

Internet surfing : online activity audit tool / Mohd Hafizi
Kamarudin.

INTERNET SURFING: ONLINE ACTIVITY AUDIT TOOL

MOHD HAFIZI B KAMARUDIN

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

**FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
KOLEJ UNIVERSITI TEKNIKAL KEBANGSAAN MALAYSIA
2006**

TESIS^ APPROVAL STATUS FORM

JUDUL: INTERNET SURFING: ONLINE ACTIVITY AUDIT TOOL

SESI PENGAJIAN: _____

Saya MOHD HAFIZI B. KAMARUDIN

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 meklumat TERHAD yang telah ditentukan oleh organisasi/badan di mama penyelidikan dijalankan)

_____ TIDAK TERHAD


(TANDATANGAN PENULIS)

Alamat tetap: No 2, Jalan Bunga Raya
Taman Kerian 34200
Parit Buntar Perak


(TANDATANGAN PENYELIA)

Pn. Siti Rahayu Bt Selamat
Nama Penyelia

Tarikh: 24/11/2006

Tarikh: 22 NOV 2006

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

^ Tesis dimaksudkan sebagai Laporan Projek Sarjana Muda (PSM)

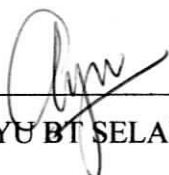
DECLARATION

I admitted that this project title name of

INTERNET SURFING: ONLINE ACTIVITY AUDIT TOOL

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

STUDENT :  Date : 24/11/2006
(MOHD. HAFIZI B KAMARUDIN)

SUPERVISOR :  Date : 24 NOV 2006
(SITI RAHAYU BT SELAMAT)

DEDICATION

To my beloved parents, Kamarudin B Saad, Noradiah bt Hj Othman

To brothers and sisters

To Noraini bt Othman and all my friends for this steadfast support.

ACKNOWLEDGEMENT

In the name of Allah the Almighty and most Merciful

Firstly, I would like to thank to Allah S.W.T for giving me the opportunity and time to finish my Project Sarjana Muda 1. I want to wish a special thanks to my supervisor, Pn Siti Rahayu for her co-operation during the implementation of the project. Your support and knowledge that she giving to me really help me in completing the task. I would like to thank the PSM committee for their hard work in giving the briefing and preparing the PSM report guidebook. Their co-operation and commitment in helping the student to finish Project Sarjana Muda 1 are very valuable for me to be the guidance to my future.

Lastly I would like to thanks to my parent and my family for their support and ideas, encouragement, and in many other aspect. Not to forget to all my friend end every person that always support behind me. Thank for their willingness to share idea and concern with me

ABSTRACT

Internet surfing: Online Activity Audit Tool is one of the important applications to whom that concerned about the usage of internet in a networking environment. The main purpose of this tool is to use URL as a source to investigate what been done by user when they are surfing the internet. A research is made on current system (proxy) and other thesis to find the information needed that can help in developing of the system. A System Development Life Cycle is selected as project reference process flow and to make sure implementation of system sunning in sequence. From the analysis, monitoring system and proxy just scanning the activity without any concern to capture and audit the activity. Using this tool, the usage of internet for inappropriate thing is a solution for implementing better firewall and proxy.

ABSTRAK

Internet Surfing: Online Activity Audit tool adalah satu applikasi yang penting kepada sesiapa yang mementingkan tentang pengawalan penggunaan internet di dalam satu- satu rangkaian. Tujuan utama applikasi ini adalah untuk menggunakan URL sebagai satu sumber untuk menyiasat aktiviti yang dilakukan oleh pengguna semasa mereka melayari internet. Satu kajian telah dilaksanakan terhadap system yang sedia ada iaitu proxy dan tesis tesis yang berkaitan untuk mencari maklumat bagi membangunkan applikasi ini. System Development Life Cycle (SDLC) telah digunakan sebagai rujukan bagi pengaliran kerja dan susun atur pembangunan untuk memastikan kelancaran projek. Daripada analisis yang dijalankan applikasi yang sedia ada hanya akan menyimpan dan mengaudit sebarang informasi tentang aktiviti pengguna. Diharapkan dengan applikasi Internet Surfing: Online Activity Audit tool dapat menyelesaikan masalah yang timbul.

TABLE OF CONTENTS

CHAPTER	SUBJECT	PAGE
	TESIS APPROVAL STATUS FORM	i
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENT	vii
	LIST OF TABLES	x
	LIST OF FIGURES	xii
	LIST OF ATTACHMENTS	xiv
CHAPTER I	1.0 INTRODUCTION	1
	1.1 Project Background	1
	1.2 Problem Statements	2
	1.3 Objective	3
	1.4 Scopes	4
	1.5 Project Significant	5
	1.6 Expected Output	5
	1.7 Conclusion	6
CHAPTER II	2.0 LITERATURE REVIEW AND METHODOLOGY	7
	2.1 Introduction	7
	2.2 Definition of Internet	8
	2.3 Monitoring Tool Overview	9
	2.4 Type of Monitoring Tool	12
	2.5 URL Information	14
	2.6 Project Methodology	16

	2.7	High Level Project Requirement	20
	2.8	Project Schedules	21
	2.9	Conclusion	22
CHAPTER III	3.0	ANALYSIS	23
	3.1	Introduction	23
	3.2	Problem Analysis	24
	3.3	Analysis of Current System	25
	3.4	Analysis of To-Be System	27
	3.5	Requirement Analysis	29
	3.5.1	Functional Requirements	30
	3.5.2	Software Requirements	31
	3.5.3	Hardware Requirements	32
	3.6	Conclusion	32
CHAPTER IV	4.0	DESIGN	33
	4.1	Introduction	33
	4.2	High Level Design	34
	4.2.1	Raw/Input Data	35
	4.2.2	Internet Architecture	37
	4.3	Network Architecture	37
	4.4	Logical Design	38
	4.5	Physical Design	42
	4.6	Conclusion	49
CHAPTER V	5.0	IMPLEMENTATION	50
	5.1	Introduction	50
	5.2	Software Development Enviroment Setup	51
	5.3	Software Configuration Management	53
	5.3.1	Configuration Environment Setup	54
	5.3.2	Hardware Setup	54
	5.4	Security Policies and Plan	56
	5.5	Development Status	56
	5.6	Conclusion	58
CHAPTER VI	6.0	TESTING	59
	6.1	Introduction	59
	6.2	Test Plan	60
	6.2.1	Testing Organization	60
	6.2.2	Test Environment	61
	6.2.3	Test Schedule	65
	6.3	Test Strategy	69
	6.3.1	Classes Of Test	69
	6.4	Test Design	70
	6.5	Test Case Result	75
	6.6	Conclusion	79

CHAPTER II	7.0	PROJECT CONCLUSION	80
	7.1	Observation and Weaknesses	81
		7.1.1 Weaknesses	81
		7.1.2 Strengths	82
	7.2	Propositions for Improvement	82
	7.3	Conclusion	83
		REFERENCES	85
		APPENDICES	1-58

LIST OF FIGURES

FIGURE	TITLE	PAGE
2.1	Connection to the internet	8
2.2	A basic internet process	9
2.3	A snapshot of Distinct Network Monitor Packet Analyzer	13
2.4	SDLC process flow	17
3.1	Internet Filter flow chart for current system	26
3.2	Proxy diagram for current system	27
3.3	Internet Surfing :Online Activity Audit flow chart	28
4.1	ERD of Database	35
4.2	A Basic Network Architecture	37
4.3	Internet Surfing: Online Activity Audit Tool design diagram	38
4.4	Physical Design for Internet Surfing Online Activity Audit	39
4.5	Context Diagram for Internet Surfing: Online Activity Audit Tool	41
4.6	Data Flow Diagram for Internet Surfing: Online Activity Audit Tool	43
4.7	User Login interface	44
4.8	Registration interface design	45
4.9	Main Page Interface Design	46
4.10	View Raw Data	47
4.11	View URL Detail Interface Design	48

LIST OF TABLES

TABLE	TITLE	PAGE
4.1	Example of Raw Data	34
4.2	User Account	36
4.3	Main Database	36
4.4	Keyword	37
5.1	Implementation Status of Modules	57
6.1	Unit Testing Schedule	62
6.2	System Testing Schedule	63
6.3	Login Admin Test Case	64
6.4	Client Server Connection Test Case	65
6.5	Save Data from URL Test Case	67
6.7	View Raw Data Test Case	67
6.8	Viewing Raw Data Detail Test Case	67
6.9	URL Classification Data Test Case	68
6.10	Adding Word To URL Classification Test Case	68
6.11	Login Admin Test Case Results	74
6.12	Client Server Connection Test Case Results	75
6.13	Save Data from URL Test Case Results	76
6.14	View Raw Data Test Case Results	76
6.17	Viewing Raw Data Detail Test Case Results	77
6.18	URL Classification Data Test Case Results	77

Adding Word To URL Classification Test Case Results

CHAPTER 1

Introduction

1.1 OVERVIEW

Audits are facts of life in organizations. Audit responsible for checking every little aspect in order to detect any abuse of source or anything that might harm the organization. It can be called a tool for an organization to manage any potential risk and result can be used to prevent the risk.

There are lots of activities been done by user in a network when they were surfing the internet. The activity is like downloading notes, searching for information, viewing source code for the site and others. In a place like cyber café where there are more than 20 computers, it hard to know what been done by the user for security concern and future reference.

Internet Surfing: Online Activity Audit is a system that can detect and record the activity that been done by user when they are online. This system record the IP address, date, activities and the computer name that been done by user in a log file.

The recorded activity in the log then is audit and used to generate a statistic for future reference and for security concern.

1.2 PROBLEM STATEMENT

In a huge computer network environment, lots of activities been done by user. The relationship between security and user community may be the most important. The reason for this is very simple. To prevent any security breach from happening that might cause risk to the organization. Remember behind the most security breaches are authorized user that no follow the rules. There are some problems that been trace in network environment that not be aware by some organization:

1. The usage control of online service.

- In some organization, the internet surfing service is open for those who working in the company. Some of the organization has made rules and regulation to the usage of the internet service. Even though the firewall can block the access to the website that are not necessary or may cause vulnerability to the company, but there is thousand more website that can be accessed by user. How to determine it for security concern? [1]

2. No standardization on internet surfing monitoring

- Right now, there is no standardized in monitoring the internet as a whole and none is being researched and developed. The only way to monitoring the internet now is to use existing public software and extend their functionalities. There are couple problems

with this approach. First, these public software are not intended for monitoring. Their usage eats up network capacity; thus allowing only a small amount of monitoring activities. Second, monitoring the internet is difficult and not many people are doing it. As a result, problems are not often reported and consequently solved infrequently. As a result, the internet performance is degrading. This phenomenon created by the lack of monitoring is referred to "gridlock" [2].

1.3 OBJECTIVE

The objectives of the systems are:-

- i. To develop an Internet Surfing: Online Activity Audit Tool.
- ii. To capture URL of activities been done by user while they are surfing the internet and record it in a log file
- iii. To provide audit log and generated a graphically view report based on the data recorded

1.4 SCOPES

This tool will be used by administrator at any network environment. Development of this tool is using java programming language integrated with PHP or JSP. The scope will be:

- i. Monitor online surfing activities (HTTP) by the user, capture it and record it in a log file.
- ii. Displaying the data by the site criteria, the report of the activities.

1.5 CONTRIBUTIONS

The development of this system brings lots of contribution to those who involved in monitoring the network. This ability to capture user activities is an advantage to this system where most of the current monitoring software does not cover on this activity. The main contribution of this software is:

1. **Make audit of internet usage easier.**

Sometimes it takes time to audit the usage of internet in an organization. Some of workers might be lying about what have been done by them when they are surfing the internet. This may cause the audit results not precise. By using this system, all the data of the activity can be captured without the knowledge of the user. The data also can determine any misuse of the internet and captured the person.

2. **Can detect any suspicious activities made by user.**

When someone is trying to attempt any suspicious activity for example trying to guess password of a site account, they will leave the trace in this system. The activities sometimes may harmful to the network. Using the trace we can detect any attempt to do illegal activities by user.

1.6 EXPECTED OUTPUT

The system that is going to be developed will be running on the web server of a network. The system will run all the time to capture all the surfing activities by user. When a user using Internet Explorer in a network, they will be prompted with a window that required them to enter their username and password. This means that the system is

just for network environment that strictly have authorized user to go outside the network to the internet. When the user entering the right password it will allow the user to use the internet and automatically it will start capturing what been done by user.

The data captured is the IP address, the user name, the site visited, the activities, the categories of activities. The data then will be saved inside the MySQL database. When an administrator wants to check the data there will be a webpage. The page will display a statistic of the activity by user. The page also can be accessed just within the intranet for security concern.

1.7 CONCLUSION

The relationship between the security and the user community of a company may be the most important. The reason for this is simple which to prevent any vulnerability to the company that may cause the company facing any risk. A company can spend thousand of ringgits on security but one single employee that doesn't aware about the security policies can cause the company to be in danger. One of the ways is by internet.

Lots of activities been done by user when they are surfing. Some are useful, some just for fun and some are danger to the company. The question now is how to determine what has been done by user when they online.

The reason for development of Internet Surfing: Online Surfing Activity Audit is to monitor what has been done by user when they are online. By using this system the activities by user can be audit and then can be used as reference to study for managing risk to the organization.

CHAPTER 2

Literature Review and Project Methodology

2.1 INTRODUCTION

This chapter focuses on literature review and project methodology. In this chapter, it will discuss and review about the approach and related research, reference and other findings about this system.

In project methodology section, selected approach or methodology to develop the system will be described. Every steps and activities will be stated for every stage of development. The requirements that are requisite in this system will be explained in high level project requirement and followed by Project Schedule and Milestones

2.2 DEFINITION OF INTERNET

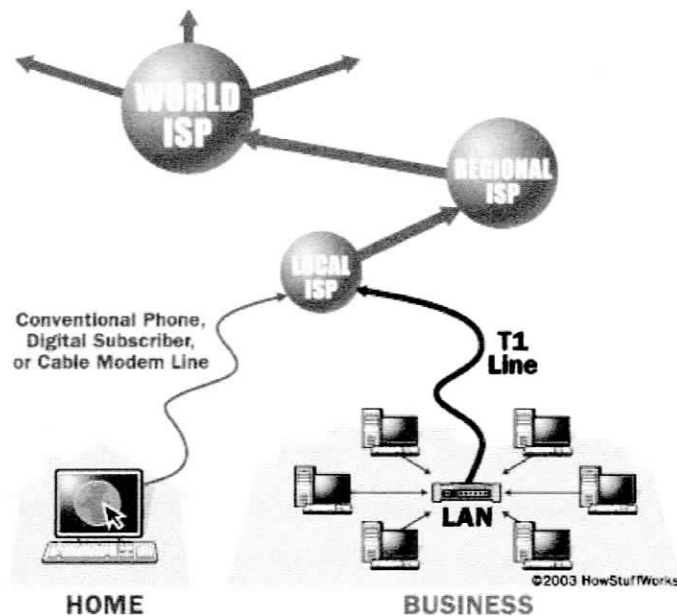


Figure 2.1: Connection to the internet [3]

The internet is a world wide connection of more than 171 millions computers that use the Internet Protocol (IP) to communicate. The Internet protocol was invented for the Advances Research Projects Agency (ARPA) of the U.S Department Of Defense. The goal is to create a centralized network that would continue to function if a bomb destroyed one or more of the network's node; information can still be reroute automatically so it could still reach its address. As a result of this bomb-proof design, any user can still communicate with other user, regardless of their locations.[xx]

Nowadays, the internet is become more and more popular. There lots of internet services provided to user. The eleven most popular internet services is:

1. Electronic Mail.(E-mail)
2. List Serv

3. Newsgroups
4. Chat
5. Instant Messaging.
6. Videoconferencing
7. File Transfer Protocol (FTP)
8. Multimedia Streaming
9. World Wide Web.
10. Rich Site Summary (RSS)
11. Blogging. [4]

When a person using the internet, there are lots of processes happen from requesting the information to the web until downloading the reply from the web server. For example, if a user wants to go to KUTKM web page, using an internet browser, he entered `http://www.kutkm.edu.my` in the URL field and search. The basic process behind this request is stated in the figure below:



Figure 2.2: A basic internet process [3]

When the URL is entered in the browser, browser will identified the URL and broke it into three parts:

- 1) The protocol : http
- 2) The server name: www.kutkm.edu.my
- 3) The file name: mainpage.html

The server communicated with a name server to translate the server name into IP address, which is use to connect to the server machine. The browser then formed a connection to the server at that IP address on port 80. Port 80 is the port that usually use by HTTP (Hypertext Transfer Protocol). Following the HTTP protocol, the browser sent a GET request to the server, asking for the file "http://www.kutkm.edu.my/mainpage.html". The server then sent the HTML text for the Web page to the browser. The browser read the HTML tags and formatted the page onto the user screen.

2.3 MONITORING TOOLS OVERVIEW

In a large network environment, it's difficult to monitor the activities done by user one by one. Some of the activities might be safe and some might be harmful to the network. At some company, there are policies on internet usage. This is for security concern and for saving resources. The policy is monitor on four general's area:

1) Awareness

- The awareness of user of the company policy. All users must know the policy of the internet use in the company.

2) Systems

- The monitoring of system configuration. The purpose is to make sure and determine whether the system has been configured according to company policy.

3) Employees

- This is the most important area where the most vulnerability come from this area. This monitoring is to make sure all employees use the computer according to policies

4) Computer use

- Monitoring on computer use where the computer is just for business use only.

To cover on those four area, it is very difficult to be done manually especially at the organization that have limited numbers of IT staff. This is where monitoring tools are use.

Monitoring is primarily intended to identify what has gone wrong or is about to go wrong. In general, monitoring systems can be thought of as having four components:

- Data collection and/or generation
- Data logging or storage
- Analysis, comparison, or evaluation
- Reporting and exception alerting

There are lots of monitoring tools available in the market today. All cover the monitoring in different way. In this project, the main concern is in the internet use. It is really important to know what have been done by user when there are surfing through the internet. It is because the internet is one of the major ways where lots of threat can

harm the network. Some user in an organization does not aware about the threat of the internet to the network.

2.4 TYPE OF MONITORING TOOLS

1) Ethereal

Ethereal is a network packet analyzer. A network packet analyzer will try to capture network packets and tries to display that packet data as detailed as possible. It works as network packet analyzer as a measuring device used to examine what's going on inside a network cable.

In the past, such tools were either very expensive, proprietary, or both. However, with the advent of Ethereal, all that has changed. Ethereal is perhaps one of the best open source packet analyzers available today. [5]

1. Distinct Network Monitor Packet Analyzer for IP Networks

Distinct Network Monitor is packets capture and network protocol analyzer software that translates complex protocol negotiation into natural language, pinpointing where errors occurred. Not only is easier to use than any other competing products, but it also translates the packet negotiation into natural language, something no other network protocol analyzer does. Network Monitor was developed for network professionals who need to quickly detect network errors rather than wading through pages of incomprehensible network traffic.