

ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล.

ระบบกรองคำหยาบผ่านเว็บเพจบนพร็อกซีเซิร์ฟเวอร์

WEB-BASED DIRTY WORDS FILTERING
ON PROXY SERVER SYSTEM



611752797
112925056

วัน เดือน ปี.....	22 พ.ค. 2550
เลขทะเบียน.....	0.3.3.5.2
เลขเรียกหนังสือ.....	คท. ๐๓.๓๗๙.๖ ๒๕๔๙
"ห้องสมุดคณะเทคโนโลยีสารสนเทศ สจล."	

รายงานนี้เป็นส่วนหนึ่งของวิชาโครงการพัฒนาระบบงาน
หลักสูตรวิทยาศาสตรมหาบัณฑิต สาขาวิชาเทคโนโลยีสารสนเทศ
คณะเทคโนโลยีสารสนเทศ
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคเรียนที่ 1 ปีการศึกษา 2549

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้ภายในห้องสมุดเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

**WEB-BASED DIRTY WORDS FILTERING
ON PROXY SERVER SYSTEM**



**A SYSTEM DEVELOPMENT PROJECT OF THE REQUIREMENT FOR
THE DEGREE OF
MASTER OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY
FACULTY OF INFORMATION TECNOLOGY
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG**

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2006

FACULTY OF INFORMATION TECHNOLOGY

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับใช้เฉพาะในหน่วยงานที่ออกหรือรับมอบนี้ไปโดยไม่อาจไปอ้างใช้เพื่อการดำเนินการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ชื่อหัวข้อ	ระบบกรองคำหยาบผ่านเว็บเพจบนพรีอ็อกซีเซิร์ฟเวอร์
นักศึกษา	นางสาวณัฐณี วิทยาปัญญา
รหัสนักศึกษา	47066403
ปริญญา	วิทยาศาสตรมหาบัณฑิต
สาขาวิชา	เทคโนโลยีสารสนเทศ
แขนงวิชา	วิทยาการสารสนเทศ
ปีการศึกษา	2549
อาจารย์ที่ปรึกษา	รศ. ดร. โชติพัทธ์ ภรณ์วลัย

บทคัดย่อ

เนื่องจากปัจจุบันองค์กรต่างๆ องค์กรมีการใช้งานอินเทอร์เน็ตกันอย่างแพร่หลาย แต่จากการศึกษาลักษณะการใช้งานอินเทอร์เน็ตของผู้ใช้งานภายในองค์กรพบว่า ได้มีการใช้อินเทอร์เน็ตไปในทางที่ไม่เหมาะสมจำนวนมาก ทำให้เกิดการผลเสียด้านการเสียเวลาในการทำงาน ดังนั้นดิฉันจึงได้เกิดแนวคิดที่จะทำการพัฒนาระบบกรองคำหยาบผ่านเว็บเพจบนพรีอ็อกซีเซิร์ฟเวอร์ขึ้น โดยจะทำงานผ่าน Web application ซึ่งจะช่วยทำหน้าที่อำนวยความสะดวกแก่ผู้ดูแลระบบคอมพิวเตอร์ในการปรับตั้งค่าของ Squid Proxy Server โดยจะสามารถปรับตั้งค่าให้ Squid Proxy Server ทำการไม่อนุญาตให้เข้าถึง URL IP Address ที่ผู้ใช้งานร้องขอ ปรับตั้งค่าให้โปรแกรมไม่อนุญาตให้เข้าถึง Web Site โดยพิจารณาจากคำที่อยู่ใน Request URL นอกจากนี้โปรแกรมกรองเว็บเพจยังสามารถใช้ปรับแต่งค่าของ Protocol ICAP (Internet Content Adaptation Protocol) เพื่อใช้ในการกรองเนื้อหาของเว็บเพจที่ประกอบไปด้วยคำ และประโยคที่ไม่เหมาะสมออก ผลจากการทำงานระบบนี้จะช่วยทำให้เพิ่มแบนด์วิดท์ให้กับอินเทอร์เน็ต ทำให้คนภายในองค์กรมีเวลาในการทำงานมากขึ้น และยังเป็นการป้องกันไวรัสคอมพิวเตอร์ได้อีกทางหนึ่งด้วย

Title	Web-Based Dirty Words Filtering on Proxy Server System
Student	Ms. Natthini Witthayapanya
Student ID.	47066403
Degree	Master of Science
Programme	Information Science
Academic Year	2006
Advisor	Asst. Prof. Dr. Chotipat Pornavalai

ABSTRACT

As the internet has been used widely in many organizations these days, misunderstanding of how to use it correctly would lead to the more time consuming in business. Hence, the new program of web application has been developed in order to help for managing Proxy Server Squid by scanning with keywords in the URL IP Address that is the result of the requirement of system administrator to block Web Site as it has been found that keywords which are in the request URL are not appropriated. Furthermore, this program can also setting Protocol ICAP (Internet Content Adaptation Protocol) in order to filtering keywords in the web content to shown in web content. These would reduce time consuming in business as well as to protect the virus from the unknown source.

กิตติกรรมประกาศ

การเดินทางครั้งนี้ดิฉันคงไม่สามารถมาถึงยังจุดหมายปลายทางได้ หากปราศจากเสบียงอาหารที่ส่งมาอย่างไม่เคยขาดจาก คุณพ่อ และคุณแม่ ขอขอบคุณท่าน อ.โชติพัชร ซึ่งเปรียบดังแสงสว่างที่คอยส่องแสงช่วยนำทางดิฉันเสมอ ยามที่การเดินทางของดิฉันพบกับความมืด ขอขอบคุณเพื่อนๆ IS18.1 ที่เป็นเพื่อนร่วมเดินทางที่แสนดี สุดท้ายนี้ดิฉันขอขอบคุณทุกคนที่มอบความรักความหวังดีตลอดการเดินทางให้แก่ดิฉัน ถึงแม้ดิฉันจะไม่ได้กล่าวนามของท่านเหล่านั้นในโอกาสนี้ แต่ทุกๆท่านนั้นจะจารึกอยู่ในความทรงจำของดิฉันตลอดไป

ณัฐณี วิทยาปัญญา



สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา.....	1
1.3 ทฤษฎีหรือแนวความคิดที่ใช้ในการศึกษา.....	1
1.4 ขอบเขตการศึกษา.....	2
1.5 ขั้นตอนการศึกษา.....	2
1.6 ประโยชน์ที่คาดว่าจะได้รับ.....	2
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....	3
2.1 Proxy server.....	4
2.1.1 ประวัติความเป็นมาของ Proxy Server.....	4
2.1.2 การทำงานของ Proxy Server.....	4
2.1.3 คุณลักษณะของ Proxy Server.....	6
2.1.4 Protocol HTTP (Hypertext Transfer Protocol).....	7
2.1.5 Protocol ICP (Internet Control Protocol).....	7
2.1.6 Cache.....	8
2.1.7 Squid.....	10
2.2 Protocol ICAP (Internet Content Adaptation Protocol).....	10
2.2.1 Architecture ของ ICAP.....	11
2.2.2 รายละเอียดของ ICAP Server และ ICAP Client.....	12
2.2.3 Mode การประมวลผลของ Protocol ICAP.....	12
2.2.4 ICAPURIs.....	14

สารบัญ (ต่อ)

	หน้า
2.2.5 ICAP Header.....	15
2.2.6 การ Implement Protocol ICAP.....	16
2.3 การทำงานของ Proxy Server ร่วมกับ Protocol ICAP.....	18
2.4 ทฤษฎีที่ใช้ในการสร้างโปรแกรมกรองเว็บเพจ.....	20
2.4.1 Web Server Apache.....	20
2.4.2 Perl Script.....	21
บทที่ 3 การออกแบบระบบงาน	
3.1 ความต้องการของระบบ.....	24
3.2 โครงสร้างของระบบ.....	24
3.3 การทำงานของระบบ.....	25
3.4 แผนภาพแสดง Use Case.....	27
3.5 แผนภาพแสดง Activity.....	29
บทที่ 4 การพัฒนาระบบงาน	
4.1 เครื่องมือที่ใช้พัฒนาระบบ.....	37
4.2 ขั้นตอนการพัฒนาระบบ.....	38
4.3 หน้าจอของระบบ.....	41
4.4 การทดสอบการทำงานของโปรแกรม.....	49
บทที่ 5 สรุปผลการศึกษาและข้อเสนอแนะ	
5.1 บทสรุป.....	58
5.2 ข้อดีและข้อเสียของระบบ.....	58
5.3 ข้อเสนอแนะ.....	59
บรรณานุกรม.....	60
ภาคผนวก.....	61
ภาคผนวก ก. การติดตั้งโปรแกรมและการใช้งาน.....	62
ประวัติผู้เขียน.....	63

สารบัญตาราง

ตารางที่

หน้า

2.1 ส่วนดีของภาษา Perl..... 22



สารบัญรูป

รูปที่	หน้า
2.1 แสดงภาพรวมของ Proxy Server.....	5
2.2 แสดงตัวอย่างของค่า Response ที่ได้รับมาก่อน.....	9
2.3 แสดง Request ที่เป็น If-Modified-Since.....	9
2.4 แสดง Response ตัวเลข 304- Not-Modified.....	9
2.5 แสดงการทำงานและส่วนประกอบของ Service Protocol ICAP.....	12
2.6 แสดงขั้นตอนในการทำงานใน mode Request Modification.....	13
2.7 แสดงขั้นตอนในการทำงานใน mode Response Modification.....	14
2.8 แสดงโครงสร้างของ ICAP URIs.....	15
2.9 แสดง ICAP URI.....	15
2.10 แสดง Request Header.....	15
2.11 แสดง Response Header.....	15
2.12 แสดง Header เมื่อ ICAP Client ส่ง Request ไปที่ ICAP Server.....	16
2.13 แสดง Header เมื่อ ICAP Server ส่ง Response ไปที่ ICAP Client.....	16
2.14 แสดงตำแหน่งของ Vectoring point เมื่อ set ค่าในแบบต่างๆ.....	18
2.15 แสดงการทำงานของ Proxy Server ร่วมกับการทำงานของ ICAP Service.....	19
2.16 แสดงการทำงานของ Protocol ICAP ใน mode respmod.....	19
2.17 แสดงการแก้ไข file httpd.conf.....	20
2.18 แสดงการทำงานของตัวแปลภาษาแบบอินเตอร์พรีเตอร์.....	23
2.19 แสดงการทำงานของตัวแปลภาษาแบบคอมไพเลอร์.....	23
3.1 แสดงโครงสร้างระบบ.....	24
3.2 แสดง flow chart การทำงาน โดยรวมของ Proxy Server Squid และ ICAP Server.....	26
3.3 แสดง Use Case ของ ระบบ.....	28
3.4 แผนภาพแสดง Activity diagram Log on.....	29
3.5 แผนภาพแสดง Activity diagram Configuration URL to Block.....	29
3.6 แผนภาพแสดง Activity diagram Configuration IP Destination to Block.....	30
3.7 แผนภาพแสดง Activity diagram Configuration Word in Request URL to Block.....	30
3.8 แผนภาพแสดง Activity diagram View Service Squid.....	31
3.9 แผนภาพแสดง Activity diagram Log Counter Word.....	32

สารบัญรูป (ต่อ)

รูปที่	หน้า
3.10 แผนภาพแสดง Activity diagram Configuration squid and ICAP Server.....	33
3.11 แผนภาพแสดง Configuration Word Content to Filter.....	34
3.12 แผนภาพแสดง Activity diagram View Service ICAP Server.....	35
3.13 แผนภาพแสดง Activity diagram Configuration Synonym to Dictionary.....	35
3.14 แผนภาพแสดง Activity diagram View Service ICAP Server.....	36
4.1 แสดงโปรแกรม GEdit.....	37
4.2 แสดงการ set ค่าที่ squid ICAP client.....	38
4.3 แสดงการเข้าถึงเว็บเพจโดยตรง.....	39
4.4 แสดงการปรับแต่งค่าที่กรองที่ file valspeak.l ที่ ICAP Server.....	39
4.5 แสดงเว็บเพจที่มีการกำหนดให้เข้าถึงเว็บเพจผ่านทาง proxy server.....	40
4.6 แสดงหน้าจอ logon.....	41
4.7 แสดง error message ถ้าใส่ Username Password ผิด.....	41
4.8 แสดง Menu ของ โปรแกรม.....	42
4.9 แสดงหน้าจอ Configuration Squid and ICAP Server.....	42
4.10 แสดงหน้าจอ Service Squid.....	43
4.11 แสดงหน้าจอ Service ICAP Server.....	43
4.12 แสดงหน้าจอ URL to Block.....	44
4.13 แสดงหน้าจอ IP Destination to Block.....	45
4.14 แสดงหน้าจอ Word in request to Block.....	45
4.15 แสดงหน้าจอ Word Content on Web to Filter.....	46
4.16 แสดงหน้าจอ Dictionary Synonyms.....	46
4.17 แสดงหน้าจอ Dictionary Verbs Tense.....	47
4.18 แสดงหน้าจอ Log Counter Word.....	48
4.19 แสดง list URL to Block.....	49
4.20 แสดง Error message ที่ไม่สามารถเข้าเว็บที่ Block URL นั้นได้.....	49
4.21 แสดง List IP Destination to Block.....	50
4.22 แสดง Error message ที่ไม่สามารถเข้า IP Address นั้นได้.....	50
4.23 แสดงคำมีที่ Block อยู่ใน Word in Request URL to Block.....	51

สารบัญรูป (ต่อ)

รูปที่	หน้า
4.24 แสดง Error message ที่ไม่สามารถเข้าเว็บนั้นได้เมื่อมีคำที่อยู่ใน Word in Request URL to Block.....	51
4.25 แสดงการเก็บหน้าจอก่อนการเพิ่มคำที่ให้กรอง.....	52
4.26 แสดง Feature Dictionary Synonym.....	53
4.27 แสดงการเพิ่มคำเพื่อทดสอบการกรองคำ.....	54
4.28 แสดงคำที่เพิ่มจาก Dictionary Synonym.....	54
4.29 แสดงผลลัพธ์ของการเพิ่มคำโดยจะแสดงที่ Browser ที่แสดงผลที่ Client.....	55
4.30 แสดงการเก็บหน้าจอก่อนการเพิ่มคำที่ให้กรองประโยค.....	56
4.31 แสดงการเพิ่มคำเพื่อทดสอบการกรองคำ.....	56
4.32 แสดงการเพิ่มคำเข้ามาที่โปรแกรม.....	56
4.33 แสดงผลลัพธ์ของการเพิ่มคำโดยจะแสดงที่ Browser ที่แสดงผลที่ Client.....	57

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบันระบบอินเทอร์เน็ตได้เข้ามามีบทบาทเป็นอย่างมากในองค์กรต่างๆ จนเกือบจะแทบทุกองค์กรเลยก็ได้ที่มีระบบอินเทอร์เน็ตให้คนในองค์กรใช้ เพื่อเพิ่มความสะดวกในการทำงานเพื่อศึกษาหาความรู้ แต่พบว่าในบางครั้งคนในองค์กรได้ใช้อินเทอร์เน็ตไปในทางที่ไม่เกี่ยวข้องกับการทำงาน ผลก็คือ ทำให้เสียเวลาในการทำงานเป็นอย่างมาก ดังนั้นจึงได้มีการหาทางในการป้องกันไม่ให้คนในองค์กรใช้อินเทอร์เน็ตไปในทางที่ไม่เกี่ยวข้องกับการทำงาน โดยจะทำการสร้างโปรแกรมที่ใช้ในการกรองเว็บเพจ โดยจะทำการกำหนดค่าต่างๆ เช่น การกำหนดให้ทำการ Block URL IP Address ที่ไม่ต้องการให้ User เข้าถึงเว็บไซต์ กำหนดการ Block Request URL ถ้าพบว่าใน Request URL นั้นมีค่าที่ไม่เหมาะสม นอกจากนี้แล้ว โปรแกรมยังมีการกรองอีกชั้นหนึ่ง โดยจะเป็นการกรองในระดับเนื้อหาของเว็บเพจเพื่อไม่ให้เว็บเพจนั้นแสดงค่าที่ไม่เหมาะสมออกมาการกรองในระดับนี้จะช่วยให้การกรองมีประสิทธิภาพมากยิ่งขึ้นจากและจะเกิดผลกับองค์กรคือ ทำให้คนในองค์กรมีเวลาในการทำงานมากยิ่งขึ้น

1.2 ความมุ่งหมายและวัตถุประสงค์ของการศึกษา

1. เพื่อการสร้างโปรแกรมกรองเว็บเพจ โดยจะสามารถปรับแต่งค่าในการกรองของเว็บเพจผ่านทาง Web Application
2. เพื่ออำนวยความสะดวกให้กับผู้ดูแลเครือข่ายสามารถใช้งานได้ง่ายยิ่งขึ้น

1.3 ทฤษฎีหรือแนวคิดที่ใช้ในการวิจัย

1. การทำงานของ proxy server และการปรับแต่งค่าต่างๆ เพื่อกรอง URL IP และช่วงเวลา
2. การทำงานของ protocol ICAP server และการปรับแต่งค่าต่างๆ เพื่อใช้ในการกรองเนื้อหาของเว็บเพจ

1.4 ขอบเขตของการศึกษา

1. การปรับแต่งค่าผ่านทาง Web Application โดยจะสามารถกำหนดค่า เวลา URL IP
2. การปรับแต่งค่าผ่านทาง Web Application โดยจะสามารถกำหนดค่าที่ต้องการกรองได้

1.5 ขั้นตอนของการศึกษา

1. ศึกษาหลักการทำงานของ proxy server squid
2. ศึกษาการทำงานของ protocol ICAP
3. ศึกษาการปรับแต่งค่าต่างๆ ของ proxy server squid เพื่อใช้ในการกำหนดสิทธิในการเข้าถึงเว็บไซต์
4. ศึกษาการปรับแต่งค่าต่างๆ ของ protocol ICAP ที่ใช้ในการกรองค่าบนเว็บเพจ
5. ศึกษาการทำงานร่วมกันของ proxy server squid และ protocol ICAP
6. ทดลองปรับแต่งค่า เพื่อทำการกรองค่าบนเว็บเพจ
7. ศึกษาการเขียน โปรแกรมที่เป็นแบบ web application เพื่อที่จะเพิ่มค่าปรับแต่งที่ใช้ในการกรองเว็บเพจ

1.6 ประโยชน์ที่คาดว่าจะได้รับ

1. เป็นการจำกัดสิทธิของคนในองค์กรทำให้ไม่ใช้อินเตอร์เน็ตไปในทางที่ไม่เหมาะสม
2. ทำให้เว็บไซต์ที่ user ใช้ไม่มีค่าที่ไม่เหมาะสม
3. เข้าใจหลักการทำงานของ Proxy Server
4. เข้าใจหลักการทำงานของ Protocol ICAP
5. ผู้ดูแลระบบสามารถทำงานได้สะดวกโดยทำการปรับแต่งค่าผ่านทางเว็บเพจ

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

เนื่องจากในปัจจุบันได้มีการใช้อินเตอร์เน็ตในองค์กรกันอย่างแพร่หลาย แต่ในบางครั้งการใช้อินเตอร์เน็ตไปในทางที่ไม่เหมาะสม ทำให้เสียเวลาในการทำงานจึงทำให้เกิดแนวคิดในการโปรแกรมกรองเว็บเพจขึ้นมาโดยจะเลือกให้ตัวกรองเว็บเพจนั้นทำงานอยู่บน Proxy Server เพื่อการจัดการที่ง่ายเนื่องจากทำการปรับแต่งค่าที่ Proxy Server เพียงเครื่องเดียวก็ทำการกรองเว็บเพจได้ โดยในขณะที่ Client ร้องขอเว็บเพจนั้นเราสามารถที่จะตรวจสอบว่า URL ที่ร้องขอมานั้นอนุญาตให้ใช้ได้และสร้างตรวจสอบเนื้อหาของเว็บเพจด้วยว่ามีค่าที่ไม่เหมาะสมอยู่หรือไม่ โดยในการตรวจสอบนี้จะใช้ Protocol ICAP ช่วยในการตรวจสอบและกรองค่าที่ไม่เหมาะสมออกไป และเพื่อเป็นการง่ายต่อการใช้งานจึงได้พัฒนา โปรแกรมกรองเว็บเพจที่เป็นการปรับแต่งค่าผ่านทาง Web Application ดังนั้นทฤษฎีที่เกี่ยวข้องใน โครงงานนี้จะแบ่งออกเป็น 4 ส่วน คือ

ส่วนที่ 1 เป็นการศึกษาในส่วนของ Proxy Server จะมีรายละเอียดเกี่ยวกับการทำงานของ Proxy Server และการทำงานของ Protocol ที่เกี่ยวข้อง เช่น Protocol HTTP (Hypertext Transfer Protocol) Protocol ICP (Internet Cache Protocol) ส่วน Proxy Server ที่ใช้ในการศึกษานี้จะใช้ Proxy Server ของ Squid เนื่องจากเป็น Proxy Server ที่มีประสิทธิภาพสูง ใช้กันอย่างแพร่หลาย และเป็น Open source เป็น free Proxy Server โดยจะศึกษาถึงการทำงานต่างๆ และการกรองเว็บเพจของ Proxy Server Squid และนอกจากนี้แล้ว Proxy Server Squid ยังสนับสนุน Protocol ICAP อีกด้วย

ส่วนที่ 2 จะเป็นการศึกษาในเรื่องการสร้างตัวกรอง เว็บเพจ โดยจะทำงานพัฒนาต่อจาก Squid ซึ่ง Squid สามารถกรองได้ในระดับ URL IP Address Request URL สิ่งที่จะพัฒนาต่อคือให้สามารถกรองเว็บเพจระดับเนื้อหา โดยจะศึกษาที่ Protocol ICAP ซึ่งเป็น Protocol ที่ใช้ในการกรองเนื้อหาของ เว็บเพจ ที่ได้รับความนิยมมาก เพื่อนำมาใช้ในการพัฒนาโปรแกรมกรอง เว็บเพจต่อไป

ส่วนที่ 3 จะกล่าวถึงการทำงานของ Proxy Server Squid ร่วมกับ Protocol ICAP

ส่วนที่ 4 จะกล่าวถึงทฤษฎีที่ใช้ในการสร้างโปรแกรมกรองเว็บเพจที่เป็นแบบ Web Application โดยจะกล่าวถึง Web Server ของ Apache และ Perl Programming

2.1 Proxy server

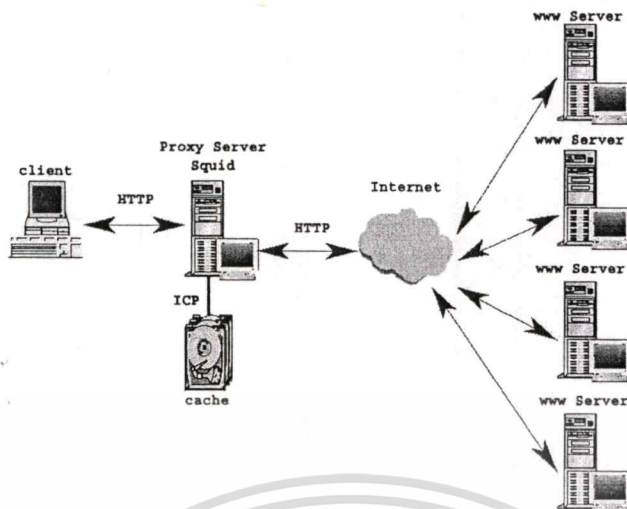
2.1.1 ประวัติความเป็นมาของ Proxy Server

web ได้เริ่มต้นขึ้นในปี ค.ศ. 1990 Proxy Server แต่เดิมนั้นเรียกว่า gateways ในครั้งแรก WWW gateway ถูกเขียน โดย Tim Berners-Lee จากทีม World Wide Web ที่ CERN ทีมของ gateway ที่นี้จะอ้างถึงอุปกรณ์ที่ทำงาน Forward Packet ระหว่าง Network ในบางครั้งจะเป็นการ Convert ระหว่าง Protocol เดียวกันหรือต่าง Protocol กันในปี ค.ศ. 1993 ทีม Web Proxy Server ได้ตัดสินใจที่จะร่วมทีมกับ WWW gateway เพื่อทำการแยกให้ชัดเจนระหว่าง Internet/firewall gateways(Proxies) กับ Information gateways(gateway)

Internet/firewall gateways ได้เปลี่ยนชื่อเป็น Proxy Server เพื่อเป็นการสื่อถึงการกระทำที่เกิด Information gateway (หมายถึง เครื่องมือของ CGI หรือ Application ของ Web Sever API การเข้ามาแทนที่ Standalone ของ Server ที่มีวัตถุประสงค์ที่เป็นพิเศษ) ขึ้นกับ Client ในทางตรงกันข้ามกับการกระทำที่เกิดกับ Server

2.1.2 การทำงานของ Proxy Server

proxy Server จะเป็นเครื่องคอมพิวเตอร์ที่อยู่ตรงกลางระหว่างเครื่อง Client และเครือข่ายอินเทอร์เน็ต มีหน้าที่ในการเก็บหน้าเว็บเพจที่เครื่อง Client ได้เคยเข้าถึงเว็บไซต์นั้นมาแล้ว โดยจะทำการเก็บหน้าเว็บเพจไว้ที่ cache ซึ่งอุปกรณ์ที่ใช้เก็บของมุล คือ ฮาร์ดดิสก์ โดยที่ถ้าในครั้งต่อไป ถ้า Client เรียกใช้งานหน้าเว็บเพจเดิมอีกก็จะทำการ update ให้เป็น version เดียวกันกับ origin Server แล้วก็จะส่งเว็บเพจนี้ไปให้ Client ประโยชน์ของการใช้ proxy Server คือ ทำให้เพิ่ม bandwidth ในการเข้าถึงเว็บเพจ โดยที่ถ้าเป็นเว็บเพจที่เคยมีอยู่ใน cache แล้วก็จะสามารถส่งเว็บเพจนั้นไปให้ Client ได้เลยโดยไม่ต้องไปร้องขอเว็บเพจจาก origin Server มาอีก โดยจะแสดงภาพรวมของการทำงานดังรูปที่ 2.1



รูปที่ 2.1 แสดงภาพรวมของ Proxy Server

ตัวอย่างการทำงานของ Proxy Server ในกรณีที่ Client ต้องการที่จะเข้าถึงเว็บ <http://www.squid-cache.org> จะมีขั้นตอนการทำงานดังนี้

1. เมื่อ Client ต้องการที่จะเข้าถึงเว็บ <http://www.squid-cache.org> ก็จะส่งคำสั่งร้องขอไปยัง Proxy Server

2. Proxy Server ก็จะทำการตรวจสอบว่าใน Proxy Server ว่า URL ที่ Client ร้องขอนั้นมีอยู่ใน Cache ของ Proxy Server หรือไม่

- ในกรณีที่ Proxy Server พบว่ามีเว็บเพจที่ร้องขอนั้นอยู่ใน Cache ดังนั้น Proxy Server ก็ จะทำการร้องขอเว็บเพจนั้นไปที่ Origin Server ของ Squid บนอินเทอร์เน็ตเพื่อนำค่าที่อยู่ใน Header กลับมาตรวจสอบกับ Cache ที่มีอยู่ว่าเป็น Version เดียวกันหรือไม่ ถ้าเป็น Version เดียวกัน Proxy Server ก็จะส่งเว็บเพจที่มีอยู่ใน Cache ไปให้กับ Client ที่ร้องขอ แต่ถ้าในกรณีที่ เป็นคนละ Version กัน Proxy Server ก็จะทำการร้องขอเว็บเพจที่ Origin Server ของ Squid ให้ส่ง เว็บเพจมาให้ แล้ว Proxy Server ก็จะทำการสำเนาเว็บเพจนั้นไว้ใน Cache (เพื่อว่าถ้ามีการร้องขอ เว็บเพจนี้ในคราวต่อไป จะได้ส่งเว็บเพจที่มีอยู่ใน Cache ไปให้) พร้อมทั้งส่งเว็บเพจนั้นมาให้ Client

- ในกรณีที่ Proxy Server พบว่าไม่มีเว็บเพจนั้นอยู่ใน Cache ก็จะทำการร้องขอไปที่ Origin Server ของ Squid ให้ส่งเว็บเพจมาให้ แล้ว Proxy Server ก็จะทำการสำเนาเว็บเพจนั้นไว้ ใน Cache (เพื่อว่าถ้ามีการร้องขอเว็บเพจนี้ในคราวต่อไปจะได้ส่งเว็บเพจที่มีอยู่ใน Cache ไปให้) พร้อมทั้งส่งเว็บเพจนั้นมาให้ Client

2.1.3 คุณลักษณะทั่วไปของ Proxy Server

- Transparency มีการ Filter เป็นเบื้องหลังบน Proxy และจะต้องไม่มีผลกระทบต่อผลลัพธ์สุดท้ายคือ User จะได้รับ Response ที่เหมือนกันระหว่างการ Connect โดยตรงหรือการใช้ Proxy Server

- Client สามารถกำหนดได้ว่าจะใช้หรือไม่ใช้ Proxy Server
- ที่ Server ปลายทางไม่ได้รับผลกระทบระหว่าง Proxy

Transparency ในความหมายของ Proxy Server หมายถึง User จะไม่ต้องการรู้สถานะของ Proxy Server ในการเข้าถึง Web จะต้องเป็นไปอย่างราบรื่น บ่อยครั้งที่ User ไม่ทราบถึงเส้นทางการทำงานของ Proxy Server โดยทางพฤติกรรมแล้ว User จะมองเห็นเสมือนกันว่าเป็นการรวมกันของ Proxy และ Original Server เพราะว่า Web ที่เข้าถึงโดย Proxy จะเหมือนกับการเข้าถึง Server โดยตรงและ Cache ที่ Proxy Server จะช่วยให้เพิ่มประสิทธิภาพในแง่ความเร็วของการเข้าถึง Web อีกด้วย ในทางตรงกันข้าม ถ้าการทำงานของ Proxy มีการควบคุมและกรอง จะมีการนิยามของคำว่า non-transparent Effect และ User อาจจะสามารถบอก Proxy ถึง Error Message ที่สร้างโดย Proxy Server

Non-Transparent Proxy Server (เป็น Technique ชั้นแรกของ Proxy Server) มีการทำงานคือ Proxy Server จะเข้าสู่ Client ได้โดยตรงโดยใช้ Program Client และการ Request URL จะปรากฏที่ท้ายสุดของ URL บน Proxy Server

ดังนั้น Proxy Server จะทำการเข้าถึง เว็บเพจ ได้ URL ทั้งหมดที่อ้างอิงนี้ จะตรวจสอบย้อนหลังด้วย URL ที่แสดงให้เห็นที่ Proxy Server แต่ในทางตรงข้ามแล้ว URL ที่ User เห็นจะเป็น URL ที่แตกต่างจาก Proxy และ Proxy Server สามารถรู้ได้ แต่ในเวลาต่อมาเมื่อ Client มีการสร้าง Proxy-ware ก็ได้เปลี่ยนแปลงวิธีการเป็น Transparent Proxy Server อย่างซ้ำๆ

2.1.4 Protocol HTTP (Hypertext Transfer Protocol)

Protocol HTTP จะเป็น Protocol ที่สำคัญและเกี่ยวข้องโดยตรงกับ Proxy Server ใน World Wide Web นั้นจะประกอบด้วย Protocol หลายตัวประกอบกัน แต่ Protocol HTTP จะเป็น Protocol อันดับแรกที่ใช้กันในการส่งผ่านเอกสารบน Web

Protocol HTTP เป็น Protocol ที่ใช้ในการ Request หรือ Response โดย Client จะส่ง Request ไปที่ Server และ Server ก็จะส่ง Response มาที่ Client จากหลักการพื้นฐานนี้สามารถนำมาใช้กับ Proxy Server คือ Client จะส่ง Request ไปยัง Proxy Server และ Proxy Server ก็ส่ง Request นั้นต่อไปยัง Server หรือส่งต่อไปให้กับ Proxy อีกอีกเรียกว่า Request Chain ส่วน Response ที่ได้กลับมาจะมีเส้นทางเดียวกับการ Request หรือ Response ที่ได้ อาจจะเป็นแบบที่ Proxy Server ส่ง Resource ใน Cache กลับไป โดยไม่ต้องส่งต่อไปที่ Server

Protocol HTTP ประกอบด้วย Method ที่เกี่ยวกับ URL และ Version ของ Protocol ที่อยู่ในส่วนของ Header Method ที่ใช้มากคือ Method GET ใช้ในการเข้าถึง เว็บเพจ ส่วน Method POST ใช้ใน Submit form ของหน้า เว็บเพจ Protocol HTTP ถูกออกแบบให้ใช้งานง่ายโดยเป็น

1. ASCII text Protocol ง่ายต่อการ debug ไม่เหมือนกับที่เป็นแบบ binary
2. มีส่วนที่ขยายได้ คือ Protocol HTTP อนุญาตให้มี Header ขนาดที่ใหญ่ สามารถยืดหยุ่นในการ configuration ค่า
3. มีความเข้ากันได้ของต่าง Version คือ ในการออกแบบ Application ที่มีคนใช้เป็นจำนวนมากนั้น การเข้ากันได้กับ Protocol ต่าง Version เป็นเรื่องที่สำคัญเป็นอย่างมาก Protocol HTTP จะสามารถทำงานที่ Version เก่ากว่าโดยจะ ignore feature บาง feature ที่ Protocol Version ปัจจุบันไม่รู้จัก
4. ความเร็วในการทำงาน Protocol HTTP เป็น Protocol ที่ทำงานได้อย่างรวดเร็ว เพราะมีการทำงานไม่ซับซ้อน มีคำสั่งไม่มากนัก

2.1.5 Protocol ICP (Internet Control Protocol)

Protocol ICP เป็น Protocol ที่ใช้สำหรับการ Query ของ Proxy Server ในการเรียก เว็บเพจ ใน Cache โดยทั่วไปแล้วจะใช้กับการที่ Proxy Server เรียกใช้ Cache จาก Proxy Server อื่นๆ แต่ก็สามารถใช้ได้กับ Client ที่ Query ใน Proxy Cache

Protocol HTTP จะสร้างบน Protocol TCP ส่วน Protocol ICP จะสร้างบน Protocol UDP การใช้แบบ Protocol UDP นี้จะเป็นการช่วยลด Overhead ที่เกิดขึ้น รวมทั้งการ สร้าง Connection ของ TCP Protocol UDP เป็น Protocol แบบ Connectionless ไม่ต้องการสร้าง Tree-way Handshake เหมือนกับ TCP

Protocol TCP เป็น Protocol ที่มี Reliable Transport โดยจะ เก็บค่าด้วย network packet ที่มีการรับและ Acknowledge ที่เป็นจริงจากผู้ส่ง ถ้า packet ไม่ได้รับ Acknowledge มันจะทำการส่งซ้ำจนกระทั่งได้รับ Acknowledge

แต่ Protocol ทำงานตรงข้ามคือ จะเป็น Protocol แบบ Unreliable Transport มันจะไม่รับประกัน packet ที่ส่งไปและไม่ค่นึงถึงการที่ส่งไปแล้วเกิด Packet lost แต่มีข้อดีคือ Protocol UDP เป็น Protocol ที่ใช้งานง่าย และทำงานได้เร็วมาก เหมาะกับ Application ที่ต้องการความเร็วในการ Response แต่ไม่จำเป็นที่จะต้องมี Reliable Transport สามารถยอมรับได้ถ้ามีการสูญหายของ packet บาง packet

วัตถุประสงค์หลักของ Protocol ICP จะต้องทำงานอย่างรวดเร็วในการหา เว็บเพจ ใน Cache ของ Proxy Server โดยต้องการรู้ให้แน่นอนว่า Cache ที่มีอยู่หรือไม่ เพื่อจะได้ตัดสินใจได้ว่า Proxy Server จะต้อง Connect HTTP เพื่อหา เว็บเพจ ที่ Client Request

ส่วนวัตถุประสงค์อื่นของ Protocol ICP คือ การกระทำหมายถึง ความเร็วของ Proxy Server แต่ละเครื่องที่ Query ถ้าใช้ Resource บน Proxy Server หลายๆเครื่องความเร็วนี้จะแสดงถึงความเร็วของ Proxy Server และ Network ว่าเป็นอย่างไร

การทำงานพื้นฐานอย่างง่ายของ Protocol ICP ในเรื่องการกระทำ หมายถึง การ Request URL จาก Resource ที่ Query Response คือ "HIT" แสดงถึงการพบ URL ที่ Request ใน Cache หรือ "MISS" แสดงถึงการที่มิได้พบใน Resource ใน Cache ของ Proxy Server

ถ้า Lack Response ของ Proxy Server จะเป็นข้อมูลที่จะช่วยให้รู้ว่า Proxy Server Down อยู่ หรือ Overload หรือ เกิดจากการคับคั่งหรือ Down ของ Network ในกรณีนี้จะไม่มี Protocol ICP Response ที่ได้รับคือจะไม่มีการ Request URL ในกรณีที่ Response ทั้งหมดเป็น "MISS" Proxy Server จะส่งทั้งหมดเป็น 1 Response โดยทั่วไปแล้ว Proxy Server จะมี Response ที่รวดเร็วมากที่สุด ของ ICP Query ที่เป็น Response แบบ "MISS" เพราะว่าการที่ Response อย่างรวดเร็วนั้นจะทำให้ load น้อยและ/หรือ ทำให้การเชื่อมต่อของ Proxy Server ใน Network เร็ว

2.1.6 Cache

Cache มีต้นกำเนิดมาจาก Proxy Server หมายถึง ขั้นตอนในการเก็บสำเนา เว็บเพจ ขั้นตอนในการเข้าถึง เว็บเพจ ที่มีคนเรียกใช้ เว็บเพจ นั้นแล้ว ด้วยที่เก็บข้อมูลที่เป็น Media (โดยทั่วไปแล้วจะใช้ Disk แต่ในบางครั้งอาจเป็น Main Memory ที่เก็บ Cache ในระยะเวลาสั้นๆ) ที่ Proxy Server จะทำให้ประสิทธิภาพดีขึ้น เพิ่มศักยภาพ สามารถเรียก เว็บเพจ ที่มีอยู่แล้วใน Cache ด้วยระยะเวลาอันสั้น และช่วยลด Bandwidth ใน Network

ข้อดีของ Cache คือ ทำให้ประสิทธิภาพดีขึ้น โดยการลด Bandwidth Cache เป็นส่วนสำคัญที่สุดของ Proxy Server ยังมีในส่วนที่ Proxy Server สามารถทำได้อีกหลายอย่าง เช่น ในเรื่องของ Security และ Monitor ที่มีการจัดเตรียมไว้ให้ Proxy Server จะสามารถรองรับ traffic ของจำนวน Web ที่มีอัตราการเพิ่มขึ้นเป็นจำนวนมาก

ข้อเสียของ Cache คือ การใช้ Cache อาจเป็นการเสี่ยงที่จะได้รับข้อมูลเก่าที่เก็บอยู่ใน Cache แต่สามารถแก้ปัญหานี้ได้ Proxy Server จะมีการตรวจสอบค่า Header ของ เว็บเพจ โดย Protocol HTTP ด้วย Condition GET

Condition GET เป็น Feature ของ Protocol HTTP ประโยชน์คือ ใช้ในการ Request เว็บเพจ จาก Cache ของ Proxy Server จะเป็นแบบ Single Request เมื่อ เว็บเพจ ที่ Request นั้นมีอยู่ใน Cache พร้อมทั้งเป็น เว็บเพจ ที่ Up-to-Date ถ้าไม่เป็นเช่นนั้น สิ่งที่เป็นผลลัพธ์จะเป็นสิ่งที่ไม่ต้องการ (จะทำการ Request ทั้ง 2 ครั้ง ครั้งที่ 1 คือ การตรวจสอบค่า Up-to-Date ครั้งถัดมาคือ รายละเอียดของ เว็บเพจ ถ้ามีการเปลี่ยนเนื้อหาภายใน เว็บเพจ ที่ Condition GET เราจะใช้ค่า Last-Modified: ที่ Response ที่ Header ในตอนที่ได้รับ เว็บเพจ เมื่อ เว็บเพจ นั้นถูกเก็บอยู่ใน

Cache คำนี้ (วันและเวลาของ last modification จะถูกเก็บไว้ใน เว็บเพจ ที่เป็น Cache) จะส่งไปในรูปแบบของ If-Modified-Since เป็น Request Header ไป ตรวจสอบที่ เว็บเพจ จริง ว่ามีการเปลี่ยนแปลงที่ เว็บเพจ นี้หรือไม่ ตัวอย่างของ คำ Response ที่ได้รับมาก่อนแสดงดังรูปที่ 2.2

```
HTTP/1.0 200 Ok
Server: Netscape-Enterprise/2.0
Date: Sat, 19 Apr 1997 10:11:25 GMT
Last-Modified: Fri, 18 Apr 1997 15:00:34 GMT
Content-type: text/html
Content-length: 6510
```

รูปที่ 2.2 แสดงตัวอย่างของคำ Response ที่ได้รับมาก่อน

Condition GET ที่ Request นั้น ใช้ timestamp จาก Last-Modified: Header และส่งต่อไปยัง Request ที่เป็น If-Modified-Since: Header แสดงดังรูปที่ 2.3

```
GET /people/ari/ HTTP/1.0
User-agent: Mozilla/3.0
Accept: text/html, image/gif, */*
If-Modified-Since: Fri, 18 Apr 1997 15:12:00 GMT
```

รูปที่ 2.3 แสดง Request ที่เป็น If-Modified-Since:

ถ้าเว็บเพจไม่มีการเปลี่ยนแปลง Server จะ Response ตัวเลข 304 Not-Modified ซึ่งเป็น Response Status Code และจะไม่ส่งเนื้อหาของเว็บเพจมาให้ Response แบบนี้จะเป็นการช่วยเพิ่มศักยภาพและลด Bandwidth ของ network จะไม่มีการส่งเว็บเพจจริงๆ ที่มีอยู่ แต่จะเป็นการใช้เว็บเพจจาก Cache ตามตัวอย่างข้อความดังนี้แสดงดังรูปที่ 2.4.

```
HTTP/1.0 304 Not-Modified
Server: Netscape-Enterprise/2.0
Date: 20 Apr 1997 15:45:12 GMT
```

รูปที่ 2.4 แสดง Response ตัวเลข 304 Not-Modified

ถ้า เว็บเพจ มีการเปลี่ยนแปลง เว็บเพจ Version ใหม่จะส่งกลับมาและมี Response เป็น 200 Ok โดยจะต้อง ส่ง เว็บเพจ จริงๆ ไปให้ Client

2.1.7 Squid

Squid ต้นกำเนิดมาจาก module cache ของ Harvest Research Project ที่ปรับปรุงโดย Duane Wessels สนับสนุนการทำงานของ Protocol FTP gopher wais และ HTTP และยังสนับสนุน Protocol SSL สำหรับ security connection ของ proxy โดยจะเก็บ object ที่ใช้งานบ่อยๆที่ RAM และมี database ที่เก็บ object อยู่ที่ directory บน disk

Squid เป็นที่รู้จักกันดีของ web cache ที่มีคุณภาพสูง มีความเชื่อถือได้ และมีประสิทธิภาพด้วย low level web-Protocol และ cache เป็นแง่ดีในการทำงานคือทำงานในส่วนที่ใกล้กับ level ของ hardware โดยไม่ได้ออกแบบให้ทำงานใน high level เหมือนกับ เนื้อหา filter

Proxy Server Squid มีการ set ค่า configuration ได้หลายค่า โดยจะมีหลาย option (โดยทั่วไปแล้วจะอยู่ที่ /usr/etc/squid.conf) เช่น Access Control List (ACLs) เป็น parameter ที่เกี่ยวกับการทำงานของ Squid ที่เป็นในส่วนของ เวลาของ account URL hostname ที่เป็นส่วนของ URL Client hostname เป็นต้น จะแสดงได้โดยการ configuration Squid และ ACLs โดย proxy administrator ของ สามารถที่จะกำหนดการ access ของแต่ละช่วงเวลาของการทำงานที่เวลา 10:00 น.-12:00 น. ของ account แต่ละ account หรือแต่ละ hostname ของ Client ได้

ข้อดีของ Squid อีกข้อหนึ่งก็คือ สนับสนุนการใช้ cache ร่วมกัน ซึ่งถ้า proxy ที่เราใช้อยู่ไม่มี object ที่ต้องการใน hard disk ตามหลักจะต้องติดต่อกับ origin Server เพื่อไปดึงข้อมูลมา แต่ใน concept นี้ proxy ที่เราใช้ยังสามารถไปดึงข้อมูลจาก proxy อื่นๆ ได้ แต่ได้แก่เพียง origin Server ที่อยู่ใกล้กับ origin Server ของคุณ แต่ถ้าหากการเข้าไปที่ origin Server แท้จริงเร็วกว่า concept นี้ก็ไม่ใช่ว่าความคิดที่ดีเลย แต่ปัญหานี้ได้ถูกแก้แล้ว โดยการตั้งค่าการติดต่อ Server อื่นๆ ให้ไปหา object ใน Server อื่นๆ พร้อมกับ Server ของตนเอง

2.2 Protocol ICAP (Internet Content Adaptation Protocol)

มาตรฐาน ICAP สำหรับ Internet Content Adaptation Protocol เกิดขึ้นในเดือน ธันวาคม ปี 1999 ที่ ICAP Forum จุดมุ่งหมายของการออกแบบ Protocol นี้เพื่อเพิ่มเป็น Service ที่ใช้เกี่ยวกับการกรองเนื้อหาของ เว็บเพจ การ scan virus การแปลภาษา การใส่โฆษณาบน เว็บเพจ โดยจะทำให้สิ่งต่างๆเหล่านั้นบน Server โดยส่วนมาแล้วจะอยู่บน web Server

ICAP สร้างบน Protocol TCP อนุญาตให้ข้อมูลของ internet base ผ่านเข้าสู่ web Server เมื่อมีการเพิ่ม Service นี้จะทำงานได้ Service นี้จะอนุญาตให้มี traffic ที่มากผ่านสู่ web Server เหมือนกับ raw HTTP ข้อมูลที่ส่งไปยัง Server นี้จะอนุญาตให้มี traffic ที่มากผ่านเข้าสู่ web Server ทำให้สามารถใช้ได้กับ network ที่มีขนาดใหญ่ โดยที่ Service นี้จะไม่ไปรบกวนการทำงานของ web Server

เนื่องจาก ICAP เป็น HTTP-based port ที่ ICAP ใช้จะเป็น port เดียวกันกับที่ HTTP ใช้ คือ port 80 โดยส่วนมากแล้ว port 80 จะใช้สำหรับ fire wall ดังนั้นการติดตั้ง ICAP จะไม่ต้องไปเปลี่ยน policy ของ security

Protocol ICAP เกิดขึ้นเนื่องจากปัญหาการที่ Internet มีการเติบโตอย่างรวดเร็วจะเกิดผลต่างๆ เช่น การที่มี traffic ที่เป็นจำนวนมาก ทำให้เกิดปัญหาที่ bandwidth การ Access

Protocol ICAP เป็น Protocol แบบ Request และ response เหมือนกับ Protocol HTTP แต่ ICAP ไม่ใช่ HTTP ทั้งหมด ICAP เป็น application Protocol ที่ทำงานบน HTTP การออกแบบของ ICAP จะทำที่ application-layer ซึ่งอยู่บน HTTP อีกทีหนึ่ง

ICAP เป็น “Protocol ที่อนุญาตให้ ICAP Client ส่งข้อความผ่านทาง Protocol HTTP ไปที่ ICAP Server จากการ transform หรือ กระบวนการอื่นๆ ที่ Server จะ execute และส่งกลับไปในรูปของข้อความที่ส่ง response กลับไปที่ Client จะมีการ modified ข้อความเสมอในทั่วไปแล้ว มีการปรับปรุงข้อความจะขึ้นอยู่กับ Request หรือ response ของ HTTP ” จากคำกล่าวนี้เองที่ทำให้ Protocol ICAP เหมาะที่จะใช้กับ Web content Filter นอกจากนี้ Protocol ICAP ยังสามารถใช้กับการ translation language Mark Up language translation(PDA/Cell Phone) การ Scan Virus การใส่โฆษณาบน เว็บเพจ

2.2.1 Architecture ของ ICAP

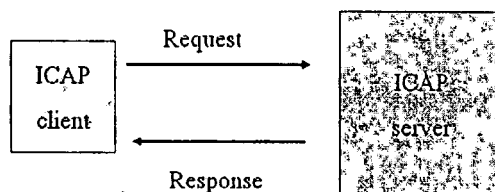
ความคาดหวังของ Service ของ Protocol นี้คือ ใช้งานร่วมกันได้ มีความง่าย และง่ายต่อการใช้งาน Protocol นี้จะต้องติดต่อสื่อสารโดยใช้กระบวนการ และจะต้องเข้ากันได้อย่างสมบูรณ์ ด้วยการ ติดตั้งที่มาตรฐานของอุปกรณ์ network และมาตรฐานของ application

ใน Commercial Web sites จะประกอบด้วย e-commerce file FTP Server และ database ถึงแม้ว่าการสื่อสารนั้นจะไม่จำกัดแต่เฉพาะ HTTP แต่ส่วนใหญ่จะใช้ HTTP เป็นหลัก ICAP Proxy จะใช้งานง่ายกับ HTTP post ซึ่งคือ “Client Request” และคำขอที่ Origin Server response จะ encapsulate ในส่วนแรกของ HTML body

การออกแบบของ Protocol ICAP จะทำที่ application-layer ซึ่งอยู่บน HTTP อีกทีหนึ่ง โดยที่การทำงานของ Protocol ICAP จะเป็นการทำงานของ ICAP Service ซึ่ง ICAP Service จะมีอยู่ 2 ส่วน คือ

- ICAP Server
- ICAP Client

ICAP Client และ ICAP Server จะเป็นเหมือนสิ่งที่ย่อยขอรับบริการ และคอยให้บริการ แสดงดังรูปที่ 2.5



รูปที่ 2.5 แสดงการทำงานและส่วนประกอบของ Service Protocol ICAP

2.2.2 รายละเอียดของ ICAP Server และ ICAP Client

- ICAP Server ทำงานเหมือนกับ HTTP Server คือ จะคอยทำงานตอบสนองกับ Request ที่ ICAP Client ร้องขอมา โดยจะทำหน้าที่ในการวิเคราะห์เนื้อหาของเว็บเพจที่ ICAP Client ส่งเข้ามาว่ามีค่าในเว็บเพจที่ต้องการจะให้กรองหรือไม่ ถ้ามี ICAP Server ก็จะทำการกรองค่าเหล่านั้นในเว็บเพจให้ แต่ถ้าไม่มี ICAP Server ก็จะทำการส่งเว็บเพจนั้นให้ ICAP Client ต่อไป ดังนั้นในการ set ค่าปรับแต่งค่าของค่าที่ต้องการจะกรองในเว็บเพจที่ต้องการปรับแต่งค่าที่ Configuration file ของ ICAP Server

- ICAP Client เป็นโปรแกรมที่ติดต่อกับ ICAP Server มีวัตถุประสงค์ในการส่ง Request ไปยัง ICAP Server จะเป็นการกำหนดค่าให้ ICAP Server ประมวลผลอย่างไร ดังนั้น ค่าที่ปรับแต่งที่ file ของ ICAP Client จะเป็นการกำหนดการประมวลผลของ ICAP Server

2.2.3 Mode การประมวลผลของ Protocol ICAP

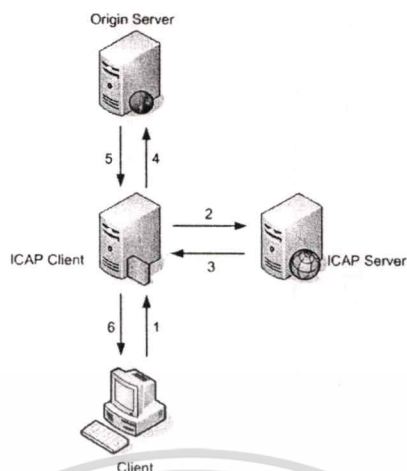
การประมวลผลเนื้อหาของเว็บเพจของ Protocol ICAP สำหรับ HTTP Service ซึ่งเป็นการกระทำที่เกิดขึ้นระหว่าง ICAP Client และ ICAP Server โดยที่ ICAP Client จะส่ง Request ไปที่ ICAP Server จะมี mode การทำงาน 2 mode คือ

- Request Modification
- Response Modification

รายละเอียดและขั้นตอนการทำงานของ Request Modification และ Response Modification

Request Modification

Request Modification (reqmod) เป็น mode ที่อนุญาตให้ modified Request ที่ส่งจาก Client ไปที่ ICAP Server ในทิศทางที่ไปสู่ Origin Server ซึ่ง ICAP Server จะทำการ modified โดยจะขึ้นอยู่กับ Service Request ของ ICAP Client มีขั้นตอนในการทำงาน แสดงดังรูปที่ 2.6



รูปที่ 2.6 แสดงขั้นตอนในการทำงานใน mode Request Modification

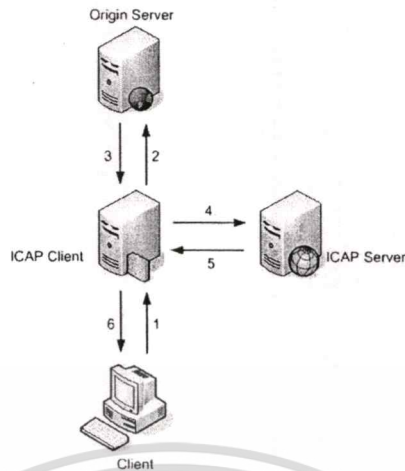
1. Client จะทำการส่ง Request ในการขอ Object จาก Origin Server ไปที่ ICAP Client
2. ICAP Client ส่ง Request ต่อไปที่ ICAP Server
3. ICAP Server จะทำการ execute ด้วย ICAP Service จาก Request ที่ได้รับ และจะส่ง modified Request ซึ่งถูก ICAP Server execute แล้วออกไป (ICAP Server disallow Access จะแสดง Error Message disallow ให้กับ Client ที่ Request เว็บเพจ) หรือ ส่ง response ของ Request กลับไปที่ ICAP Client

ถ้าในขั้นตอนที่ 3 มีการส่ง Request กลับ

4. ICAP Client ส่ง Request ในการร้องขอเว็บเพจไปให้ Origin Server
5. Origin Server ส่ง response ให้กับ Request ที่ส่งมาจาก ICAP Client ส่ง reply (จาก ICAP Server หรือ Origin Server) กลับไปที่ Client
6. ICAP Client ส่ง reply กลับไปที่ Client ที่ร้องขอเว็บเพจ (โดยอาจจะเป็น reply จาก ICAP Server หรือ Origin Server)

Response Modification

Response Modification (respmod) เป็น mode ที่อนุญาตให้ modified Request หลังจากการ return กลับจาก Origin Server โดย ICAP Server จะกระทำกับ Request Object เพียงอย่างเดียวมีขั้นตอนในการทำงาน แสดงดังรูปที่ 2.7



รูปที่ 2.7 แสดงขั้นตอนในการทำงานใน mode Response Modification

1. Client จะทำการส่ง Request ในการขอ Object จาก Origin Server ไปที่ ICAP Client
2. ICAP Client ส่ง Request ต่อไปที่ Origin Server
3. Origin Server จะส่ง response ออกมาให้จาก Request ที่ร้องขอ ไปให้ ICAP Client
4. ICAP Client จะส่งเว็บเพจที่ Origin Server reply กลับมาไปที่ ICAP Server
5. ICAP Server จะทำการ execute ด้วย ICAP Service ที่ Origin Server reply กลับมา และส่งเว็บเพจที่ ICAP Server ทำการ modified reply กลับไปที่ ICAP Client
6. ICAP Client จะส่ง reply กลับไปที่ Client

2.2.4 ICAP URIs

ICAP URI (Uniform Resource Identifiers) เป็นการระบุ Request ของ ICAP Client โดยที่ ICAP URI จะเป็น absolute URI ที่ระบุถึง complete Hostname และ Path ของ resource ในการ Request โดยที่จะมีโครงสร้าง แสดงดังรูปที่ 2.8

ICAP_URI = Scheme ":" Net_Path ["?" Query]

Scheme = "icap"

Net_Path = "://" Authority [Abs_Path]

Authority = [userinfo "@"] host [":" port]

รูปที่ 2.8 แสดงโครงสร้างของ ICAP URIs

ตัวอย่างของ ICAP URI แสดงดังรูปที่ 2.9

`icap://icap.example.net:2000/services/icap-service-1`

รูปที่ 2.9 แสดง ICAP URI

2.2.5 ICAP Header

ในการส่ง ICAP Message ของ Protocol ICAP จะมี Header ที่คล้ายกับ Protocol HTTP โดยที่จะมี Header 2 แบบ คือ Request Header และ response Header

Request Header

Request Header ของ Protocol ICAP จะเหมือนกับ Protocol HTTP คือ ICAP Request Header ที่ start status line จะประกอบด้วย Method , complete URI ของ ICAP resource และ Version ของ ICAP โดยในตัวอย่าง Header นี้ Version ของ ICAP เป็น Version 1.0 จะแสดงดังรูปที่ 2.10.

`RESPMOD icap://icap.example.net/translate?mode=french ICAP/1.0`

รูปที่ 2.10 แสดง Request Header

Response Headers

Response Header ของ Protocol ICAP จะเหมือนกับ Protocol HTTP คือ ICAP Response Header จะต้องมี start ที่ status line Header ที่จะแสดงถึง Version และ status code ของ Protocol ICAP แสดงดังรูปที่ 2.11

`ICAP/1.0 200 OK`

รูปที่ 2.11 แสดง Response Header

ตัวอย่าง Header ของ Protocol ICAP ใน mode Response Modification เมื่อ ICAP Client Request ไปที่ ICAP Server แสดงดังรูปที่ 2.12

```
RESPMOD icap://icap.example.org/satisf ICAP/1.0
Host: icap.example.org
Encapsulated: req-hdr=0, res-hdr=137, res-body=296
```

```
GET /origin-resource HTTP/1.1
Host: www.origin-server.com
Accept: text/html, text/plain, image/gif
Accept-Encoding: gzip, compress
```

```
HTTP/1.1 200 OK
Date: Mon, 10 Jan 2000 09:52:22 GMT
Server: Apache/1.3.6 (Unix)
ETag: "63840-1ab7-378d415b"
Content-Type: text/html
Content-Length: 51
```

This is data that was returned by an origin server.

รูปที่ 2.12 แสดง Header เมื่อ ICAP Client ส่ง Request ไปที่ ICAP Server

ตัวอย่าง Header ของ ICAP ใน mode Response Modification เมื่อ ICAP Server ส่ง response กลับไปที่ ICAP Client แสดงดังรูปที่ 2.13

```
ICAP/1.0 200 OK
Date: Mon, 10 Jan 2000 09:55:21 GMT
Server: ICAP-Server-Software/1.0
Connection: close
ISTag: "W3E4R7U9-L2E4-2"
Encapsulated: res-hdr=0, res-body=222

HTTP/1.1 200 OK
Date: Mon, 10 Jan 2000 09:55:21 GMT
Via: 1.0 icap.example.org (ICAP Example RespMod Service 1.1)
Server: Apache/1.3.6 (Unix)
ETag: "63840-1ab7-378d415b"
Content-Type: text/html
Content-Length: 92
```

This is data that was returned by an origin server, but with value added by an ICAP server.

รูปที่ 2.13 แสดง Header เมื่อ ICAP Server ส่ง Response ไปที่ ICAP Client

2.2.6 การ Implementation Protocol ICAP

Protocol ICAP เป็น Protocol ที่มี การ encapsulation แบบ lightweight คือ packaging number ทุกๆ packaging ของ HTTP Message จะ encapsulate เข้าไปใน ICAP body ตามลำดับ การอนุญาตของ vectoring ของ Request และ response ที่เป็นคู่กันของ ICAP Server และสำหรับการ Implement ที่ ICAP Client (ในการกำหนดค่าของ vectoring point สามารถกำหนดค่าได้ที่ ICAP Client) ในการทำ vectoring point หมายถึง การแบ่งส่วนการ adaptation เว็บไซต์ที่เข้ามาที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Cache โดยจะแบ่งเป็น 2 ส่วน คือ ส่วนที่เป็น Request ของ Client กับส่วนที่เป็น response ที่ตอบมาจาก Origin Server จะมีตำแหน่งของการ Adaptation 4 แบบ ที่แตกต่างกัน คือ

- Adaptation of Client Request
- Adaptation of Requests on their way to an Origin Server
- Adaptation of response coming from an Origin Server
- Adaptation of response coming from ICAP Client, heading back to the Client

โดยจะมีรายละเอียดของ Adaptation ทั้ง 4 แบบ ดังนี้

1. Adaptation of Client Request

ลักษณะการทำงาน คือ Adaptation แบบนี้จะทำทุกๆ ครั้งที่มี Request ที่ออกมาจาก Client ซึ่งมีการ Adaptation ซึ่งจะทำการเมื่อ Request อยู่ในทิศทางที่ไปสู่ Cache

ตัวอย่างการทำงาน: ใช้สำหรับ ตรวจสอบ URL ที่เป็น List Banded

การปรับแต่งค่าที่ ICAP Client: จะใช้เป็น reqmod_preCache

2. Adaptation of Requests on their way to an Origin Server

ลักษณะการทำงาน คือ Adaptation แบบนี้จะเหมือนกับแบบที่ 1 แต่แตกต่างกันตรงที่ทิศทางจะเป็นไปในทิศทางที่ออกจาก Cache ไป Origin Server

ตัวอย่างการทำงาน: การดึง cookie ออกจาก Request ของ Client

ในการปรับแต่งค่าที่ ICAP Client: จะใช้เป็น reqmod_postCache

3. Adaptation of response coming from an Origin Server

ลักษณะการทำงาน คือ Adaptation แบบนี้จะทำงานในตอนที่ response กลับมาจาก Origin Server โดยที่ทิศทางของ response จะอยู่ในทิศทางที่เข้าสู่ Cache ซึ่ง Adaptation แบบนี้จะทำการตรวจสอบทุกๆ เว็บเพจ ที่ Origin Server reply กลับมาก่อนที่จะเก็บเว็บเพจนั้นลง Cache

ตัวอย่างการทำงาน: ใช้สำหรับการ Scanning Virus ที่หน้าของ เว็บเพจ โดยจะทำการดึง Virus ออกก่อนที่จะเก็บลง Cache

ในการปรับแต่งค่าที่ ICAP Client: จะใช้เป็น respmod_preCache

4. Adaptation of response coming from ICAP Client, heading back to the Client

ลักษณะการทำงาน คือ Adaptation จะทำงานเหมือนกับแบบที่ 3 เป็น adaptation ของ response ที่กลับมาจาก Origin Server แต่จะมีการ adaptation หลังจากที่เก็บเว็บเพจนั้นลง Cache ไปแล้ว

ตัวอย่างการทำงาน: การกรองเนื้อหาของเว็บเพจที่ไม่เหมาะสมของ เว็บเพจ

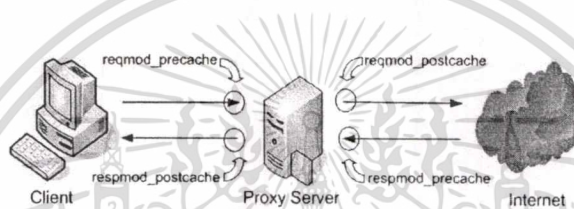
ตัวอย่างการทำงาน: การกรองเนื้อหาของเว็บเพจที่ไม่เหมาะสมของ เว็บเพจ

ในการปรับแต่งค่าที่ **ICAP Client**: จะใช้เป็น `respmod_postcache`

จากที่กล่าวมาในเรื่อง Vectoring Point ของ Protocol ICAP สามารถสรุปการปรับแต่งค่าที่ ต้อง set ที่ ICAP Client ในการ Implementation ของ Protocol ICAP โดยดูที่ลักษณะการทำงานที่ นั้น จะมีด้วยกัน 4 แบบ คือ

- reqmod_precache
- reqmod_postcache
- respmod_precache
- respmod_postcache

ตำแหน่งของ Vectoring point เมื่อ set ค่าในแบบต่างๆ แสดงดังรูปที่ 2.14

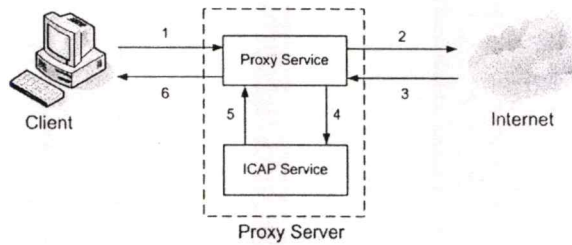


รูปที่ 2.14 แสดงตำแหน่งของ Vectoring point เมื่อ set ในแบบต่างๆ

2.3 การทำงานของ Proxy Server ร่วมกับการทำงานของ Protocol ICAP

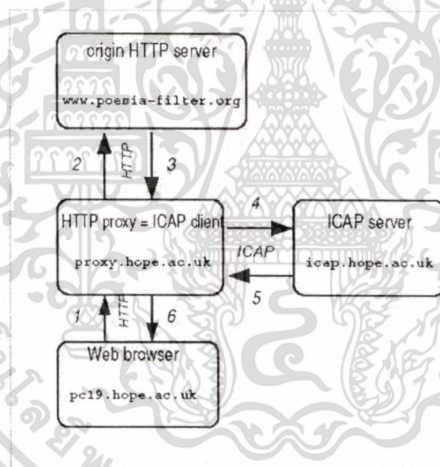
จากหลักการการทำงานของ Proxy Server และการทำงานของ Protocol ICAP ที่กล่าวมาข้างต้นนั้น เราจะนำ Protocol ICAP มาประยุกต์ใช้งานในการกรองเนื้อหาของเว็บเพจร่วมกับทำงานของ proxy Server ดังนี้

ในตอน that proxy Server ตรวจสอบเว็บเพจ และทำการ update version ของเว็บเพจนั้นเรียบร้อยแล้ว ก่อนที่จะส่งเว็บเพจนั้น มาให้กับ Client ตัว proxy Service (เป็น cache engine ที่จะไปนำเว็บเพจจากอินเทอร์เน็ตที่ Client ได้เคยร้องขอไว้มาเก็บไว้เพื่อใช้ในกรณีที่ Client มีการร้องขอเว็บเพจนี้ซ้ำอีก ก็จะส่งเว็บเพจจาก cache ไปให้ Client ได้เลย) จะส่งเว็บเพจนั้น ไปให้ Service ของ Protocol ICAP (โดยที่ Service ของ Protocol ICAP จะทำงานอยู่บน proxy Server) ICAP Service จะทำการประมวลผลเพื่อกรองเนื้อหาของเว็บเพจแล้ว ส่งกลับไปให้กับ proxy Service แล้ว proxy Service จึงจะส่ง เว็บเพจ ที่ถูกกรองเนื้อหาแล้วให้ Client ต่อไป แสดงดังรูปที่ 2.15



รูปที่ 2.15 แสดงการทำงานของ Proxy Service ร่วมกับการทำงานของ ICAP Service

ตัวอย่างขั้นตอนการทำงานของ Proxy Server ร่วมกับ Protocol ICAP แสดงตามตัวอย่าง
ผังรูปที่ 2.16



รูปที่ 2.16 แสดงการทำงานของ Protocol ICAP ใน mode respmod

มีขั้นตอนในการทำงานดังต่อไปนี้

1. เครื่อง Client ชื่อ pc19 ส่ง Request ด้วย Protocol HTTP ไปที่ Proxy Server
2. Proxy Server ชื่อ Proxy.hope.ac.uk จะ fetch เนื้อหา Web จาก Original Server www.poesia-Filter .org (โดยการส่ง Request ต่อไป)
3. Origin HTTP Server ส่ง response ด้วย Protocol HTTP กลับไปที่ Proxy Server
4. Proxy Server (ICAP Client) จะส่ง HTTP reply (ประกอบด้วยเนื้อหาของ HTML page) ไปที่ ICAP Server ชื่อ icap.hope.ac.uk โดยจะเป็นแบบ encapsulate ICAP ใน mode

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

respmod โดยใช้ ICAP Service จะมี URI เป็น icap:\icap.hope.ac.uk/filter/frence¶m=1 โดย URI ที่เป็นชนิดของ filter (response modification) แบบ Service Request ดังนั้น ICAP URI จะเป็น parameter ที่อนุญาตในการประมวลผลของ Protocol ICAP ในการกรองเนื้อหาของเว็บเพจ

5. ที่ icap.hope.ac.uk จะเป็น ICAP Server มีหน้าที่ในการวิเคราะห์เนื้อหาของเว็บเพจ (ถ้า accept content) จะตอบกลับด้วย ICAP OK ซึ่งจะเป็น original HTTP reply หรือ ในกรณีที่มีการ modified (HTML) content หรือ (ถ้ามีการ reject content) ตอบกลับด้วยข้อความ encapsulation ICAP OK HTTP 403 Forbidden error reply

6. ที่ proxy Server ชื่อ proxy.hope.ac.uk จะส่ง Response กลับไปที่ pc19 (ด้วย Protocol HTTP โดยจะแสดงผลที่ browser ของ pc19.hope.ac.uk)

2.4 ทฤษฎีที่ใช้ในการสร้างโปรแกรมกรองเว็บเพจที่เป็นแบบ Web Application

2.4.1 Web Server Apache

สำหรับ Apache นั้นเป็น web Server สำหรับระบบปฏิบัติการยูนิกซ์ หรือลินุกซ์ ที่นิยมใช้กันอย่างแพร่หลายเป็นอย่างมากด้วยประสิทธิภาพของ apache ทั้งความสามารถในการรองรับเทคนิคในการสร้างเว็บต่างๆ มากมายทำให้ apache เป็น web Server ที่ได้รับความนิยมมาก

ในการปรับแต่งคุณสมบัติต่างๆของ apache web Server เพื่อให้สามารถเขียนสคริปต์ Perl ได้ เริ่มจากเราจะเปิด file httpd.conf ขึ้นมาโดยถ้าเป็นระบบปฏิบัติการลินุกซ์ จะอยู่ที่ path : /usr/local/apache2/conf/httpd.conf ภายใน file httpd.conf ที่เราเปิดนั้นจะมีค่าคุณสมบัติต่างๆมากมายแสดงในแต่ละบรรทัด โดยที่ในบางบรรทัดจะมีการขึ้นต้นด้วย #(sharp) ซึ่งถ้าต้องการให้ web Server มีคุณสมบัติตามที่ได้กำหนดไว้ที่บรรทัดนั้นก็ทำการลบเครื่องหมาย sharp ออก apache web Server ก็จะสามารถทำงานโดยรองรับความสามารถนั้นได้โดยในหัวข้อนี้จะกล่าวถึงการ set ค่าคุณสมบัติที่ใช้ในเขียนสคริปต์ Perl คือ ให้ลบเครื่องหมาย sharp ออกตรงที่เขียนว่า AddHandler cgi-script .cgi และเพิ่มบรรทัด โดยเขียนว่า AddHandler cgi-script .pl แสดงดังรูปที่ 2.17

```
# To use CGI scripts outside of ScriptAliased directories:
# (You will also need to add "ExecCGI" to the "Options" directive.)
#
AddHandler.cgi-script .cgi.
AddHandler.cgi-script.pl.
```

รูปที่ 2.17 แสดงการแก้ไข file httpd.conf

2.4.2 Perl Script

ประวัติความเป็นมาของภาษา Perl

ภาษา Perl ย่อมาจากคำว่า Practical Extraction and Report Language

ถูกพัฒนาขึ้นโดยนายแลร์รี วอลล์ (Larry Wall) ในปี 1986 เพื่อใช้งานกับระบบปฏิบัติการยูนิกซ์ (Unix) ซึ่งในขณะนั้นการพัฒนา Application บนระบบปฏิบัติการยูนิกซ์เป็นเรื่องที่ยุ่งยากและต้องมีความรู้ความเข้าใจในภาษาโปรแกรมหลายภาษา เพราะในขณะนั้นภาษาโปรแกรมแต่ละภาษาจะทำงานในเรื่องใดเรื่องหนึ่งโดยเฉพาะ การจะทำงานออกมาสักงานหนึ่ง จะต้องใช้ภาษาโปรแกรมแทบทุกภาษาที่มีในขณะนั้นอย่าง เช่น ในส่วนของการทำงานกับไฟล์ต้องใช้ภาษา awk หรือ grep และแก้ไขไฟล์ด้วยภาษา sed เป็นต้น

ภาษา Perl ถูกสร้างขึ้นมาเพื่อให้ทำงานทุกส่วนเสร็จสมบูรณ์ในตัวเองไม่ต้องไปเรียกใช้งานภาษาอื่นอีก โดยที่ภาษา Perl ได้รวบรวมเอาส่วนดีของภาษาต่างๆ ในขณะนั้นเข้ามาไว้ด้วยกัน อย่างเช่น ภาษา grep/awk, sh, sed และ c ทำให้การสร้าง Application บนระบบปฏิบัติการยูนิกซ์เป็นไปได้อย่างมีประสิทธิภาพมากขึ้น และที่สำคัญภาษา Perl ถูกออกแบบมาให้ใช้งานง่าย ดังนั้นโครงสร้างของภาษาจึงง่ายไม่ซับซ้อน โดยมีโครงสร้างของภาษาคคล้ายกับภาษา c มาก

ในปัจจุบันภาษา Perl สามารถใช้งานได้กับระบบปฏิบัติการยูนิกซ์, ลินุกซ์, MVS, MS/DOS, Macintosh, OS/2, Amiga และอื่นๆ นอกจากนี้ด้วยความสามารถที่สำคัญของ Perl ในการจัดการกับไฟล์ text ได้เป็นอย่างดี ทำให้ภาษา Perl กลายมาเป็นภาษาสำคัญที่มีการใช้งานกันบนระบบเครือข่ายอินเทอร์เน็ต

ข้อดีของภาษา Perl

สำหรับข้อดีที่สำคัญของการใช้งานภาษา Perl นั้นจะแยกอธิบายอธิบายเป็นข้อๆ ได้ดังต่อไปนี้

- ไม่ต้องเสียค่าใช้จ่ายใดๆ ทั้งสิ้น
- ตัวแปลภาษา Perl แจกจ่ายให้ใช้โดยไม่ต้องเสียค่าใช้จ่ายใดๆ ทั้งสิ้นภายใต้ลิขสิทธิ์ GNU คือ หมายความว่าถ้าเราได้ติดตั้งและใช้งานตัวแปลภาษา Perl ในเครื่องของเราได้แล้ว เราสามารถจะไปหา download โค้ดพวกฟังก์ชัน หรือ โมดูลต่างๆ ที่มีการเขียนเอาไว้แล้วมาใช้งานได้โดยไม่ต้องเสียค่าใช้จ่ายใดๆ ทั้งสิ้น
- ใช้ได้กับระบบปฏิบัติการต่างๆ มากมาย

ภาษา Perl ถูกออกแบบในตอนเริ่มต้นให้ใช้งานกับระบบปฏิบัติการอื่นๆ เพิ่มขึ้นอีกมากมาย ดังแสดงต่อไปนี้

- ยูนิกซ์, ลินุกซ์, MS-DOS, Windows 95/98/NT, OS/2, Macintosh

ภาษา Perl ถูกสร้างขึ้นมาโดยได้รวบรวมเอาส่วนดีของภาษาต่างๆ เข้าไว้มากมาย แสดงในตารางที่ 2.1

ตารางที่ 2.1 ส่วนดีของภาษา Perl

ภาษา	ส่วนที่ภาษา Perl รับเข้ามา
grep/awk	การจัดการสตริง การค้นหาแทนที่สตริงด้วย pattern matching ซึ่งเป็นคุณสมบัติที่เด่นมากของ Perl
C	โครงสร้างของภาษา
sh	ส่วนที่สามารถเรียกรัน โปรแกรมที่เขียนด้วยภาษาอื่นได้
sed	ส่วนของการจัดการรับ text จากคีย์บอร์ด (STDIN) และแสดงผล text ออกหน้าจอ (STDOUT)

คุณสมบัติของภาษา Perl ในเรื่องการใช้งานง่ายและมีความยืดหยุ่นสูง

ภาษา Perl เป็นภาษาที่ออกแบบมาโดยมีโครงสร้างของภาษาไม่ซับซ้อนเข้าใจได้ง่าย ซึ่งโครงสร้างของภาษาจะคล้ายกับภาษา c มาก เพราะภาษา Perl สร้างขึ้นมาโดยใช้ภาษา c ทำให้ผู้ที่คุ้นเคยกับภาษา c อยู่แล้วใช้งานภาษา Perl ได้ไม่ยาก นอกจากนี้โดยตัวภาษาเองนั้นมีความยืดหยุ่นสูงทำให้จัดการกับงานด้านข้อความ, ไฟล์ text ได้เป็นอย่างดี

ความเป็นภาษาสคริปต์

ภาษาระดับสูงโดยส่วนใหญ่อย่างเช่น ภาษา c ที่เราเขียนขึ้นให้เป็นออบเจกต์ไฟล์ (.obj) แล้วจึงแปลเป็นไบนารีไฟล์ (.exe) จึงจะทำงานได้ซึ่งไม่เหมือนกับการทำงานของตัวแปลภาษา Perl ภาษา Perl เป็นภาษาสคริปต์ คือ โค้ดคำสั่งที่เราเขียนขึ้นมาจะอยู่ในรูปไฟล์ text และจะรับในรูปแบบของไฟล์ text โดยใช้อินเตอร์พรีเตอร์ (Interpreter) ไม่ต้องแปลโค้ดคำสั่งเป็นไบนารีไฟล์ (.exe) ซึ่งก็จะทำให้กระบวนการแปลความหมายคำสั่งของภาษา Perl เร็วกว่าภาษาที่ใช้คอมไพเลอร์ (compiler) เป็นตัวแปลภาษา ในหนังสือเล่มนี้จะเรียกใช้คำว่า “สคริปต์” แทนโปรแกรมภาษา Perl ที่เขียนขึ้นมา

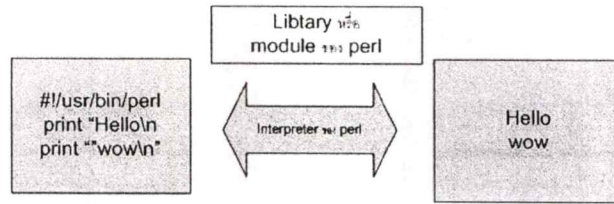
หลักการทำงานของตัวแปลภาษา

ตัวแปลภาษา คือ เมื่อเราเขียนโค้ดคำสั่งขึ้นมาตามโครงสร้างของภาษาโปรแกรมภาษาใดก็ตาม แล้วในการจะให้โค้ดคำสั่งเหล่านั้นทำงานได้ก็จะต้องมีตัวแปลภาษามาจัดการแปลโค้ดคำสั่งที่เราเขียนขึ้น เพื่อทำงานที่เราต้องการ

โดยลักษณะของตัวแปลภาษาจะแบ่งออกเป็น 2 ประเภทคือ ตัวแปลภาษาแบบคอมไพเลอร์ (compiler) และตัวแปลภาษาแบบอินเตอร์พรีเตอร์ (interpreter) ซึ่งตัวแปลภาษา Perl ก็เป็นแบบนี้ต่อไปเราจะมาดูหลักการทำงานของตัวแปลภาษาอินเตอร์พรีเตอร์กันก่อนเพื่อที่ว่า Perl ทำงานอย่างไร

อินเทอร์พรีเตอร์ (interpreter)

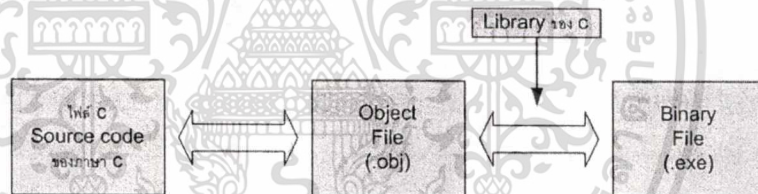
สำหรับการทำงานของตัวแปลภาษาแบบอินเทอร์พรีเตอร์นั้นจะทำงานอ่านบรรทัดต่อบรรทัดคือ อ่าน โค้ดคำสั่งแล้วก็ทำงานให้ผลออกมาเลยดังรูปที่ 2.18



รูปที่ 2.18 แสดงการทำงานของตัวแปลภาษาแบบอินเทอร์พรีเตอร์

คอมไพเลอร์ (compiler)

การทำงานของตัวแปลภาษาแบบคอมไพเลอร์ คือ มันจะทำการตรวจสอบหาข้อผิดพลาดของโค้ดคำสั่งของเราตั้งแต่ต้นจนจบก่อน หรือที่เรียกว่าการคอมไพล์ ถ้าไม่มีข้อผิดพลาดก็จะทำการแปลโค้ดคำสั่งของเราให้เป็นไฟล์นามสกุล .obj(Object file) จากนั้นก็จะทำการแปลจากไฟล์ .obj ให้เป็น ไบนารีไฟล์ .exe เพื่อทำงานต่อไป ดังตัวอย่างการทำงานของคอมไพเลอร์ภาษา c ดังรูปที่



2.19

รูปที่ 2.19 การทำงานของตัวแปลภาษาแบบคอมไพเลอร์

จากรูปตัวอย่างในกรณีที่มีการเรียกใช้ฟังก์ชันจากไลบรารี (library) ของภาษา c ฟังก์ชันเหล่านั้นจะถูกนำมาลิงค์รวมกับไฟล์ Object เป็นไบนารี .exe เพื่อทำงานต่อไป

ในส่วนของประสิทธิภาพการทำงานนั้นตัวแปลภาษาแบบคอมไพเลอร์จะทำงานได้เร็วกว่าตัวแปลภาษาแบบอินเทอร์พรีเตอร์ เพราะเมื่อโค้ดคำสั่งถูกคอมไพล์ผ่านแล้วได้เป็นไฟล์ .exe ออกมาจากนั้นก็จะเป็นขั้นตอนการทำงานอย่างเดียว

บทที่ 3

การออกแบบระบบงาน

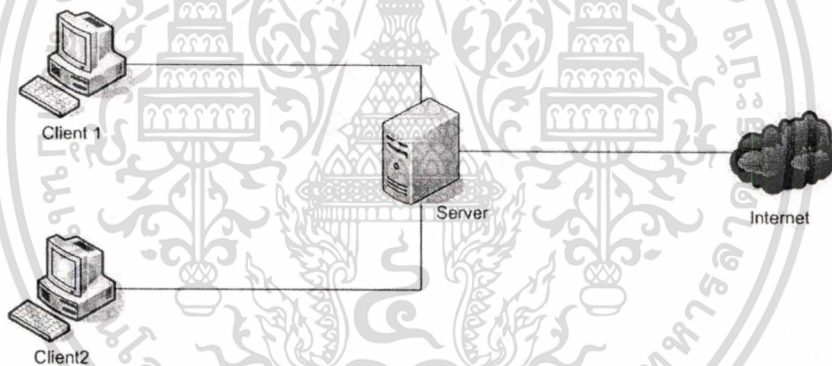
3.1 ความต้องการของระบบ

ความต้องการของโปรแกรมกรองเว็บเพจ มีดังต่อไปนี้

- โปรแกรมกรองเว็บเพจทำงาน Add On ไปบน Proxy Server Squid
- โปรแกรมกรองเว็บเพจสามารถทำการปรับแต่งค่าที่ต้องการกรองผ่านทางเว็บเพจได้
- โปรแกรมกรองเว็บเพจจะสามารถกรองคำที่เหมาะสมไม่ให้เห็นบนหน้าเว็บเพจได้
- โปรแกรมกรองเว็บเพจสามารถกำหนด URL, IP Address คำที่ใน Request URL ที่ไม่

อนุญาตให้ใช้เว็บเพจได้

3.2 โครงสร้างของระบบ



รูปที่ 3.1 แสดงโครงสร้างของระบบ

จากรูปที่ 3.1 จะสามารถแบ่งโครงสร้างของระบบเป็น 2 ส่วนคือ

Server

จะเป็นเครื่องที่ทำหน้าที่เป็น Proxy Server และ Web Server ที่ Server จะทำการติดตั้ง program ต่างๆ ดังนี้

- โปรแกรม Proxy Server Squid version ที่ support ICAP Client
- โปรแกรม ICAP Server
- โปรแกรม Web Server Apache
- โปรแกรม Perl Interpreter

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Client

จะแบ่ง Client ออกเป็นกลุ่มคือ

- Client 1 เป็นผู้ใช้อินเทอร์เน็ตจะทำปรับแต่งค่าที่ Browser ให้ทำการร้องขอเว็บเพจโดยผ่านทาง Server
- Client 2 เป็นส่วนที่ผู้ดูแลระบบจะใช้โปรแกรมในปรับแต่งค่าต่างๆ ผ่านทางเว็บ

3.3 การทำงานของระบบ

โปรแกรมกรองเว็บเพจจะเป็นโปรแกรมเขียนด้วยภาษา Perl จะเป็น Web Application ทำหน้าที่หลัก 2 ส่วน ในการปรับแต่งค่า คือ

1. ที่ Proxy Server Squid จะใช้ในการ

- ปรับแต่งค่า URL ที่ไม่อนุญาตในเข้าถึงเว็บเพจ
- ปรับแต่งค่า IP Address ที่ไม่อนุญาตในเข้าถึงเว็บเพจ
- ปรับแต่งค่าที่แสดงที่ Request URL ที่ไม่อนุญาตให้เข้าถึงเว็บเพจ

2. ICAP Server จะเป็นตัวที่ทำการวิเคราะห์ค่าที่ต้องการจะกรองออกจากเว็บเพจ จากทฤษฎีที่ได้กล่าวมาในบทที่ 2 ในการเขียน โปรแกรมที่เป็น Web application จะทำดังนี้

1. เข้าไป Scan ที่ file squid.conf ของ Proxy Server Squid แล้วทำการแสดงผล เพื่อที่จะทำการแก้ไข เพิ่ม หรือปรับแต่ง โดยจะเลือกเฉพาะค่า configuration ที่เกี่ยวข้องกับ

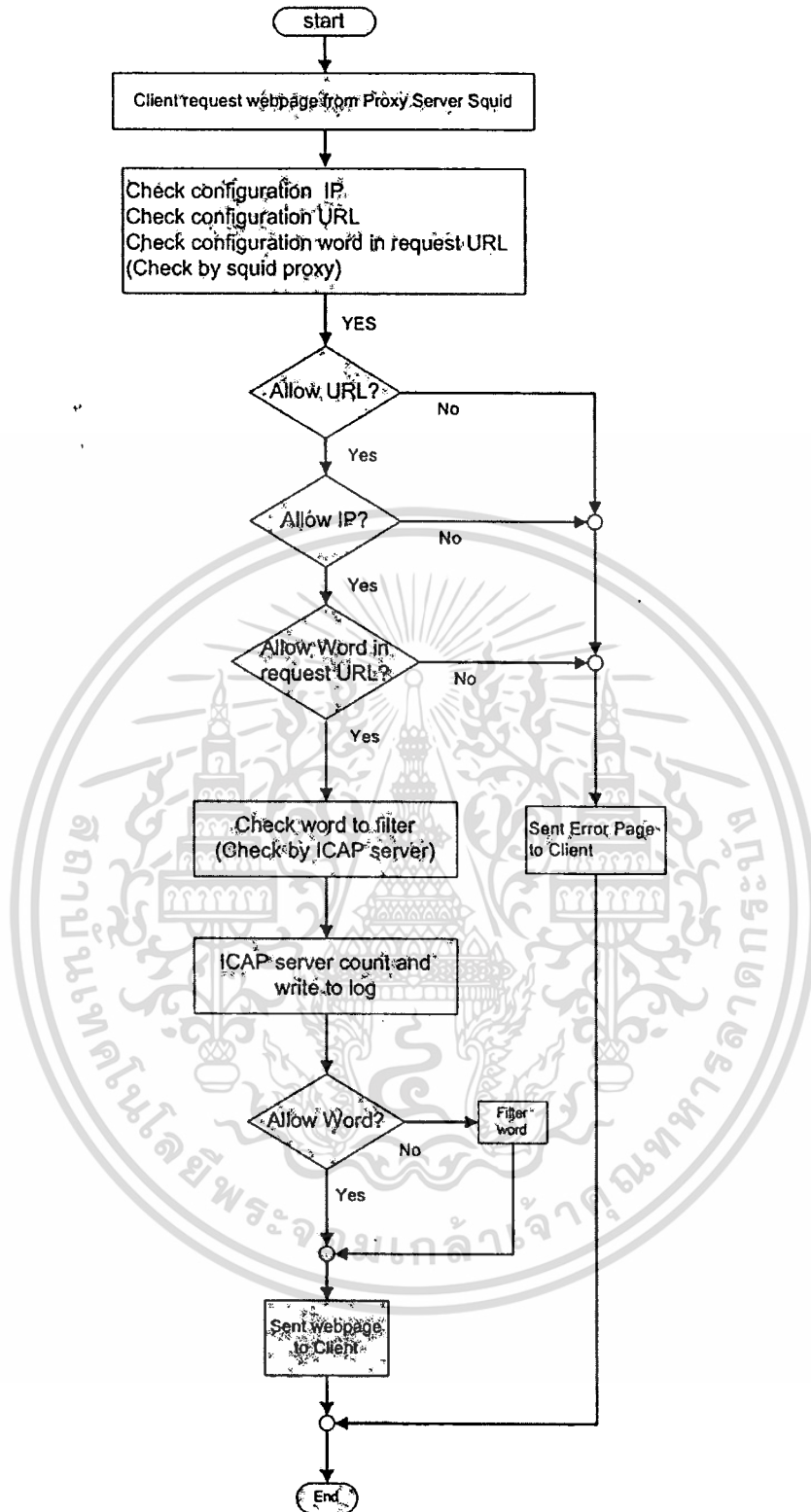
- การกำหนดคสิทธิของการเข้าถึงเว็บเพจของ URL
- การกำหนดคสิทธิของการเข้าถึงเว็บเพจของ IP Address
- การกำหนดเกี่ยวกับค่าที่อยู่ใน Request URL

เมื่อทำการปรับแต่งค่าต่างๆ เรียบร้อยแล้ว โปรแกรมก็จะทำการ Restart Service Squid เพื่อให้ Service ทำการเก็บค่าการปรับแต่งที่ได้ Update ไปแล้ว

2. เข้าไป Scan ที่ file valspeak.l ของ ICAP Server แล้วทำการแสดงผลเฉพาะที่ เกี่ยวข้องกับการกรองค่า โดยจะสามารถ แก้ไข เพิ่ม หรือปรับแต่งค่าที่ต้องการกรองและทำการบ เพิ่ม Codeในส่วนของการนับค่าที่กรองได้

เมื่อทำการปรับแต่งค่าต่างๆ เรียบร้อยแล้ว โปรแกรมก็จะทำการ Make file และ Restart Service ICAP Server เพื่อให้ Service ICAP Server ทำการเก็บค่าการปรับแต่งที่ได้ Update ไปแล้ว

หลังจากได้ทำทั้ง 2 ข้อเรียบร้อยแล้วจะกล่าวถึง flow chart แสดงการทำงานโดยรวมของ Proxy server Squid และ ICAP server



รูปที่ 3.2 แสดง flow chart การทำงานของ Proxy server Squid และ ICAP server

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

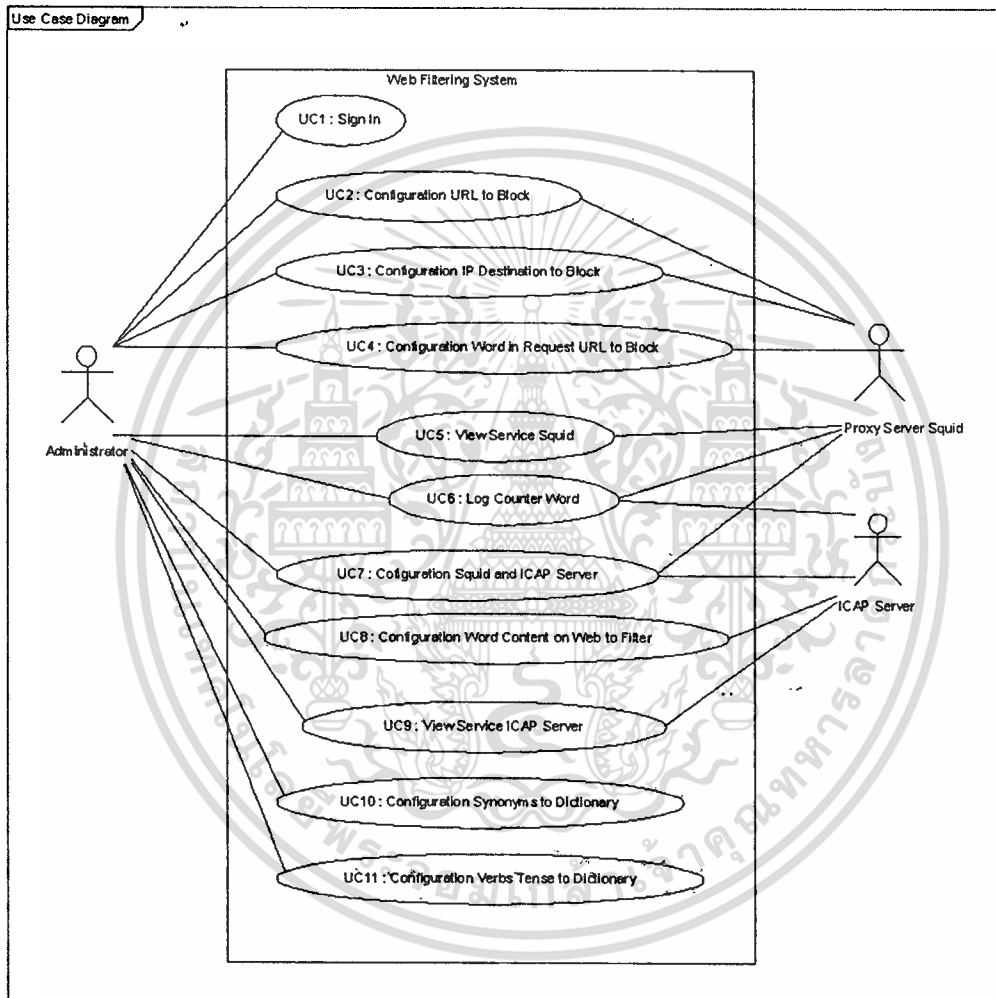
3.4 แผนภาพแสดง Use Case

ในหัวข้อนี้จะเป็นการสร้างแผนภาพที่อธิบายถึงการทำงานของระบบโดยรวม และกล่าวถึงบุคคลต่างๆ ที่มีส่วนเกี่ยวข้องในการทำงานของระบบทั้งหมด โดยจะมีแผนภาพแสดงดังรูปที่ 3.3

1. Actor แสดงถึงผู้มีส่วนเกี่ยวข้องกับระบบ ทั้ง บุคคลและ software
 - Administrator เป็นผู้ดูแลระบบและทำการปรับแต่งค่าต่างๆ
 - Proxy Server Squid เป็น Proxy Server ที่จะทำการปรับแต่งค่า
 - ICAP Server เป็น โปรแกรม ICAP Server
2. Process แสดงถึงกระบวนการทำงานหลักภายในระบบ โดยมีหลักการทำงานต่างๆ ดังนี้
 - UC1: Sign In เป็นกระบวนการตรวจสอบสิทธิในการเข้าไปใช้ระบบ
 - UC2: Configuration URL to Block เป็นการปรับแต่ง URL ที่ไม่อนุญาตให้เข้าถึง Web Site
 - UC3: Configuration IP Destination to Blockเป็นการปรับแต่งค่า IP Destination ที่ไม่อนุญาตให้เข้าถึง Web Site
 - UC4: Configuration Word in Request URL to Blockเป็นการปรับแต่งค่าคำที่อยู่ URL ที่ไม่อนุญาตให้เข้าถึง Web Site
 - UC5: View Service Squid เป็นการแสดง Status การทำงานของ Service Squid โดยจะสามารถ Start หรือ Stop Service Squid ได้
 - UC6: Log Counter Word เป็นการแสดง log ในการ access และมีการนับคำที่กรอง
 - UC7: Configuration Squid and ICAP Server เป็นการปรับแต่งค่า Port ที่ Proxy Squid ใช้งาน การปรับแต่งค่า IP Address ที่ ICAP Server ใช้งาน การปรับแต่งค่า Port ที่ ICAP Server ใช้งาน
 - UC8: Configuration Word Content to filter เป็นการปรับแต่งค่าคำที่ต้องการกรอง โดยจะสามารถปรับแต่งดังนี้
 - ลักษณะการกรองแบบ Prefix
 - ลักษณะการกรองแบบ Subfix
 - ลักษณะการกรองแบบ Case-Insensitive
 - ลักษณะการกรองแบบเป็นประโยค
 - ลักษณะการกรองแบบคำเหมือนโดยไปตรวจสอบคำที่ Dictionary
 - ลักษณะการกรองแบบ Verb Tense โดยไปตรวจสอบคำที่ Dictionary

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

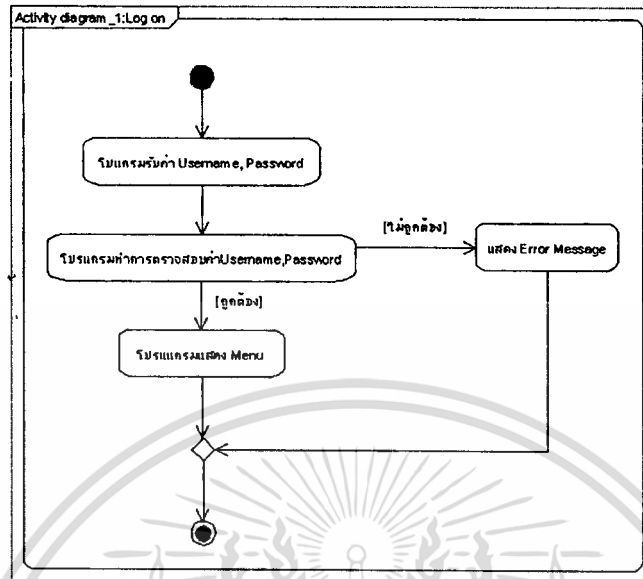
- UC9: Service ICAP Server เป็นการแสดง Status การทำงานของ Service ICAP Server โดยจะสามารถ Start หรือ Stop Service ICAP Server ได้
- U10: Configuration Synonym to Dictionary เป็นการปรับแต่งคำคำเหมือนที่ต้องการให้กรองที่ Dictionary
- UC11: Configuration Verb Tense to Dictionary เป็นการปรับแต่งคำคำที่เป็น Verb Tense ที่ต้องการให้กรอง ที่ Dictionary



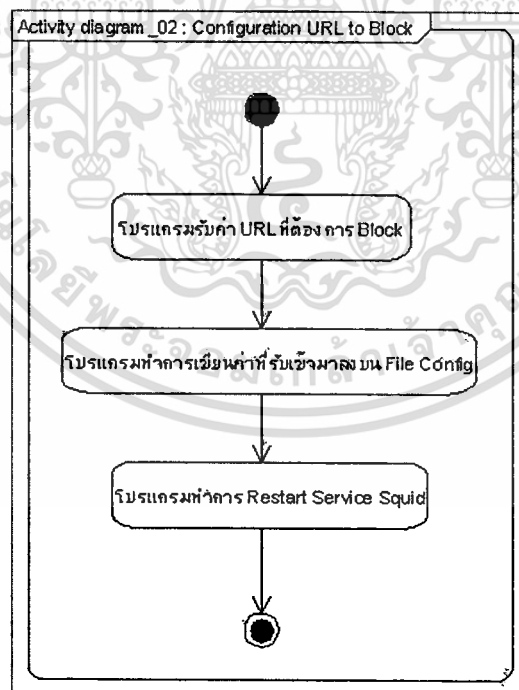
รูปที่ 3.3 แผนภาพแสดง Use Case ของระบบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 แผนภาพแสดง Activity

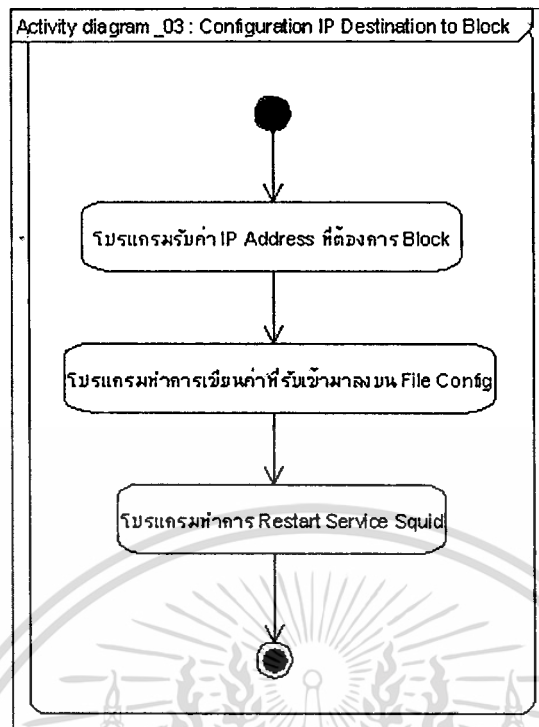


รูปที่ 3.4 แผนภาพแสดง Activity diagram Log on

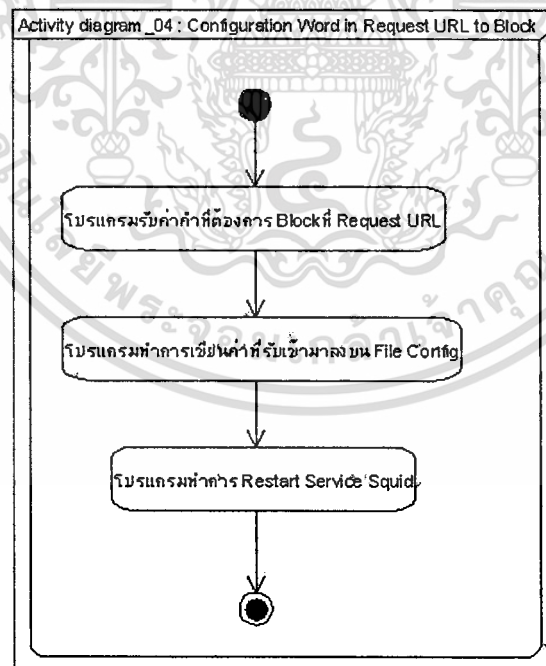


รูปที่ 3.5 แผนภาพแสดง Activity diagram Configuration URL to Block

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

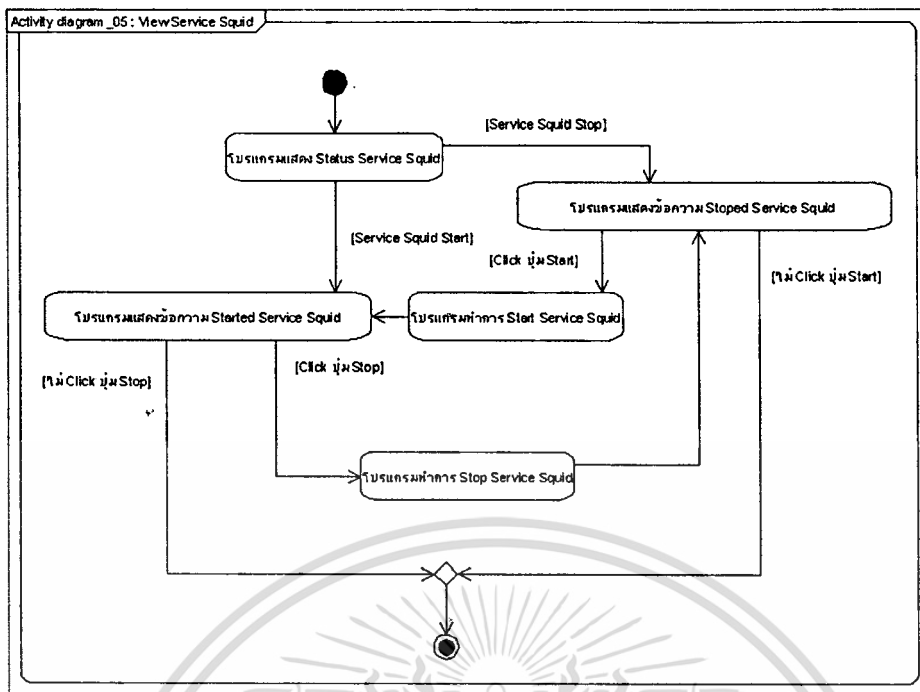


รูปที่ 3.6 แผนภาพแสดง Activity diagram Configuration IP Destination to Block



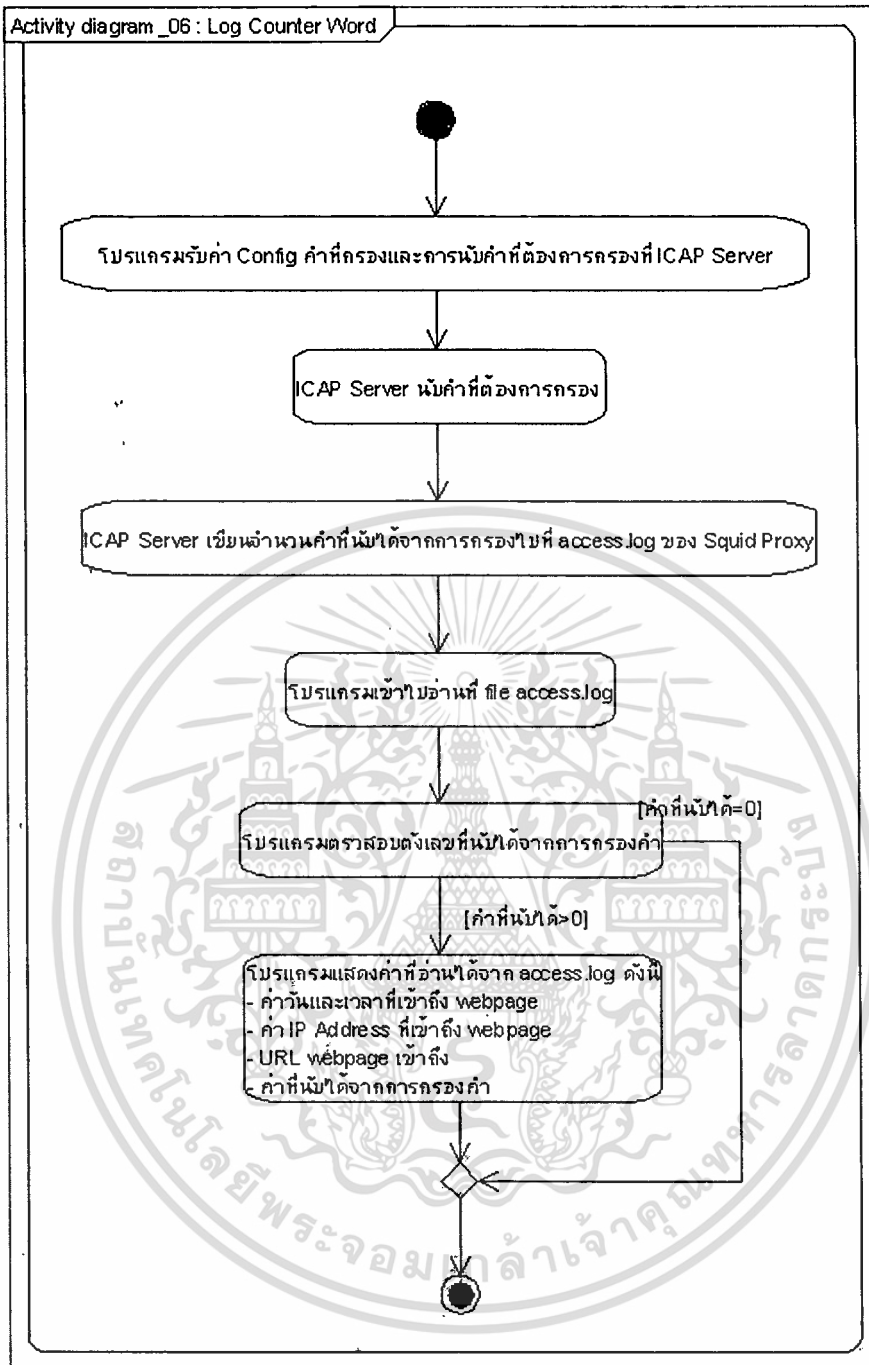
รูปที่ 3.7 แผนภาพแสดง Activity diagram Configuration Word in Request URL to Block

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



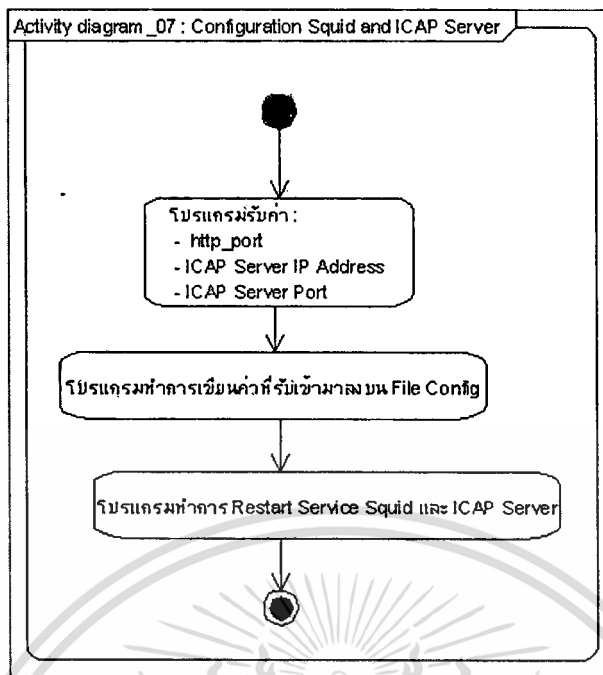
รูปที่ 3.8 แผนภาพแสดง Activity diagram View Service Squid

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

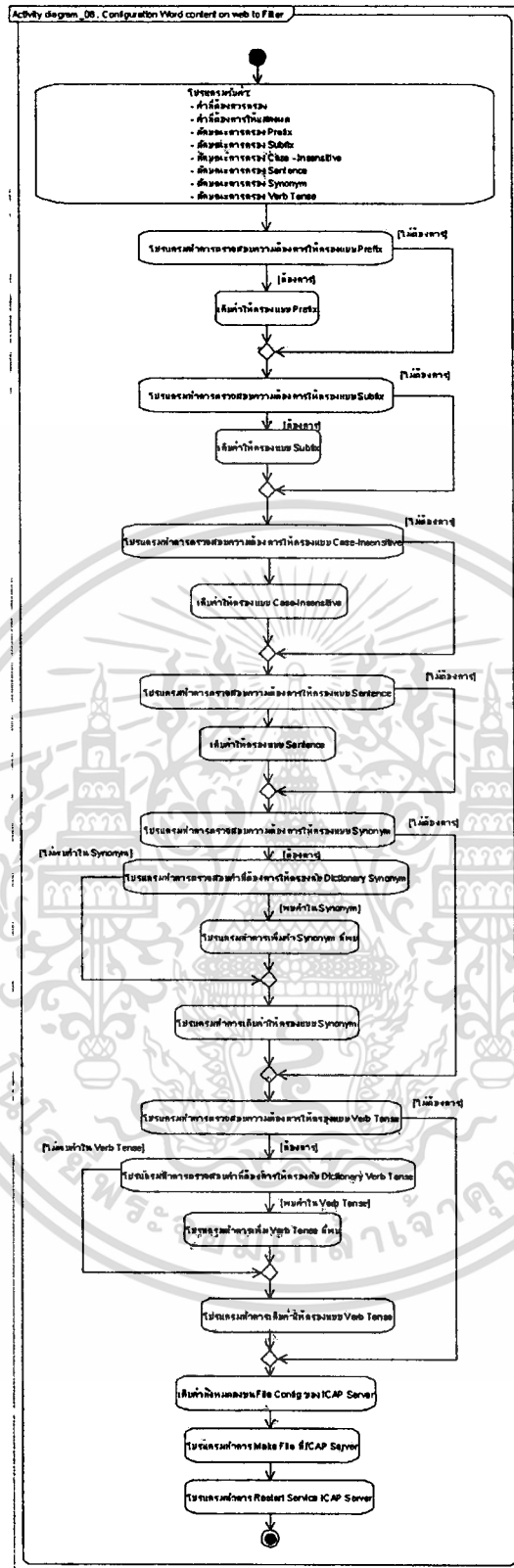


รูปที่ 3.9 แผนภาพแสดงActivity diagram Log Counter Word

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

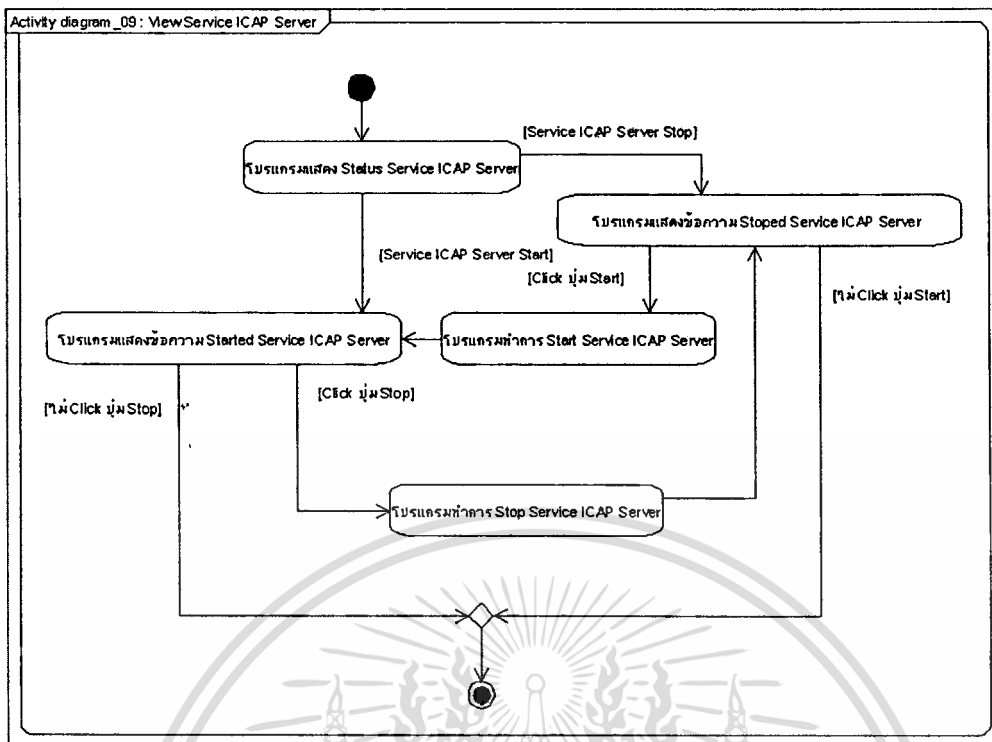


รูปที่ 3.10 แผนภาพแสดง Activity diagram Configuration squid and ICAP Server

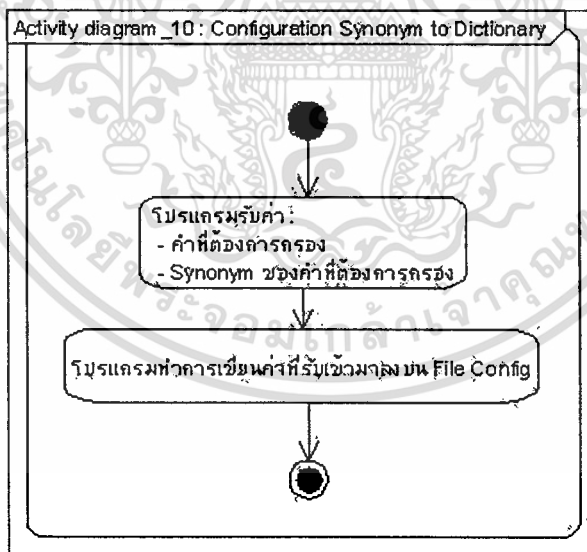


รูปที่ 3.11 แผนภาพแสดง Configuration Word Content on Web Filter

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

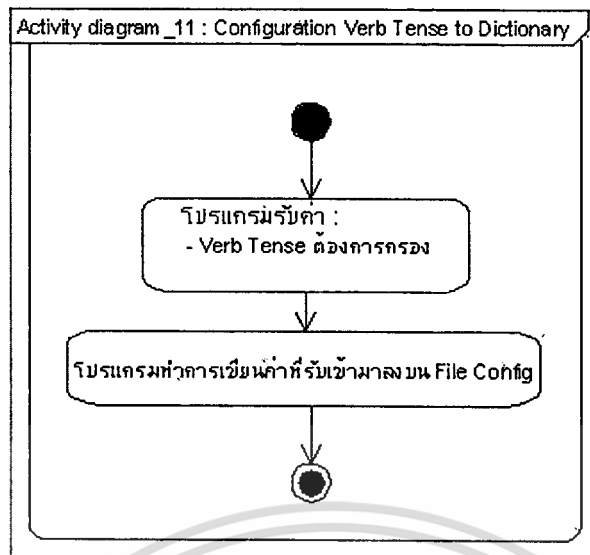


รูปที่ 3.12 แผนภาพแสดง Activity diagram View Service ICAP Server



รูปที่ 3.13 แผนภาพแสดง Activity diagram Configuration Synonym to Dictionary

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.14 แผนภาพแสดง Activity diagram Configuration Verb Tenses to Dictionary

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

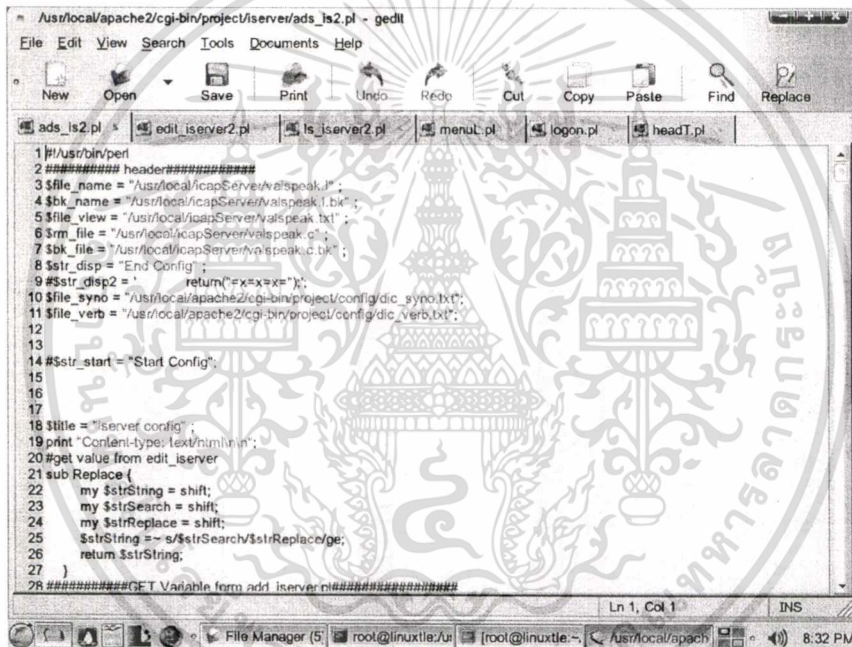
บทที่ 4

การพัฒนาระบบ

4.1 เครื่องมือที่ใช้ในการพัฒนา

1. โปรแกรม Proxy Server Squid version ที่ Support Protocol ICAP
2. โปรแกรม ICAP Server ที่ใช้ในการปรับแต่งค่าในการกรองคำ
3. โปรแกรม Apache Web Server เป็นโปรแกรมที่ใช้ทำ Web Application
4. Edit เป็นโปรแกรมที่ช่วยในการเขียนภาษา Perl และ HTML ในการพัฒนา Web

Application ของโปรแกรมกรองเว็บเพจแสดงดังรูปที่ 4.1



```
1 #!/usr/bin/perl
2 ##### header#####
3 $file_name = "/usr/local/icapServer/valspeak.l";
4 $bk_name = "/usr/local/icapServer/valspeak.l.bk";
5 $file_view = "/usr/local/icapServer/valspeak.txt";
6 $rm_file = "/usr/local/icapServer/valspeak.c";
7 $bk_file = "/usr/local/icapServer/valspeak.c.bk";
8 $str_disp = "End Config";
9 $$str_disp2 = "return( "x=x=x=");";
10 $file_syno = "/usr/local/apache2/cgi-bin/project/config/dic_syno.txt";
11 $file_verb = "/usr/local/apache2/cgi-bin/project/config/dic_verb.txt";
12
13
14 #$$str_start = "Start Config";
15
16
17
18 $title = "server config";
19 print "Content-type: text/html\n";
20 #get value from edit_issuer
21 sub Replace {
22     my $strString = shift;
23     my $strSearch = shift;
24     my $strReplace = shift;
25     $strString =~ s/$strSearch/$strReplace/g;
26     return $strString;
27 }
28 #####GET Variable form addl_issuer#####
```

รูปที่ 4.1 แสดงโปรแกรม GEdit

4.2 ขั้นตอนในการพัฒนาระบบ

เนื่องจากโปรแกรม ICAP client นั้นจะเป็นโปรแกรมที่ถูกพัฒนาให้ integrate ไปบน Squid proxy server ดังนั้นเมื่อเราต้องการที่จะใช้งาน protocol ICAP เราจะต้องเลือก version ของ squid ที่ support protocol ICAP และในการติดตั้งก็จะต้องทำการ configure ให้ enable icap support ด้วย

ส่วน ICAP server เป็นโปรแกรมที่แยกจาก Squid proxy server ดังนั้นในการใช้งานจะสามารถติดตั้งที่เครื่องเดียวกันกับที่ติดตั้ง Squid proxy server ก็ได้ หรือสามารถติดตั้งที่เครื่องอื่นก็ได้ เช่นเดียวกัน

1. ทำการติดตั้ง Proxy Server Squid version ที่ Support Protocol ICAP (Version ที่ใช้ในการพัฒนาคือ Squid Web Proxy as ICAP client version 1.2.1)
2. ทำการติดตั้ง ICAP Sever (Version ที่ใช้คือ ICAP Server 1.0)
3. ทำการปรับแต่งค่า Config ของ ICAP Client โดยจะทำการปรับแต่งที่ Proxy Serve Squid ที่ file Squid.conf แสดงดังรูปที่ 4.2

```
icap_enable on
icap_service s1 respmod_postcache 0 icap://10.49.0.8:1344/respmod
icap_class c1 s1

acl HTTP proto HTTP
acl GET method GET
icap_access c1 allow HTTP GET
```

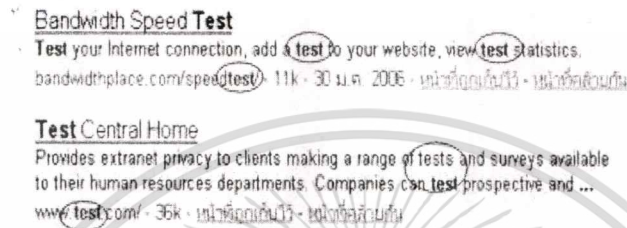
รูปที่ 4.2 แสดงการ set ค่าที่ squid ICAP client

- 3.1. set ค่าจะทำการ enable icap โดยเลือก icap_enable on
- 3.2. set ค่า service icap โดยคำสั่ง icap_service s1 respmod_postcache 0 icap://10.49.0.8:1344/respmod ซึ่ง vectoring point ที่ใช้ในการ adaptation คือ respmod_postcache ส่วน ICAP server ที่ใช้อยู่เป็น IP 10.49.0.8 (เป็น IP เดียวกับ Proxy server) port ที่ใช้ในการสื่อสารของ ICAP server และ ICAP client คือ 1344
- 3.3. icap_class c1 s1 เป็นการให้ icap_class ชื่อ c1 เป็น class ที่ใช้กับ icap_service s1
- 3.4. acl HTTP proto HTTP คือให้ HTTP เป็น access control list ของ protocol HTTP
- 3.5. acl GET method GET คือ ให้ GET เป็น access control list ของ method GET

3.6. icap_access c1 allow HTTP GET เป็นการอนุญาตให้ class c1 ใช้ protocol HTTP ใน method GET ได้

4. ทดลองปรับแต่งค่าคำที่ต้องการกรอง

ก่อนการทดลองจะทำการเก็บผลของการเข้าถึงหน้าเว็บเพจโดยให้เข้าเว็บเพจโดยตรง (ไม่ผ่านทาง proxy server ที่ config ให้มีการกรองเนื้อหาของเว็บเพจ) จะได้ผลลัพธ์แสดงดังรูปที่ 4.3



รูปที่ 4.3 แสดงการเข้าถึงเว็บเพจโดยตรง

ทดลองปรับแต่งค่าคำที่ต้องการกรองโดยการปรับแต่งค่าการกรองที่ file valspeak.l ที่ ICAP Server แสดงดังรูปที่ 4.4

```

1 /root/icapServer/valspeak.l (modified) - gedit
2
3
4
5 5<SCRIPT* (BEGINJAVA); return(yytext);
6 6<JAVA*</script> (BEGIN0); return(yytext);
7 7<JAVA*</SCRIPT*> (BEGIN0); return(yytext);
8 8<style* (BEGINSTYLE); return(yytext);
9 9<STYLE* (BEGINSTYLE); return(yytext);
10 10<STYLE*</style*> (BEGIN0); return(yytext);
11 11<STYLE*</style*> (BEGIN0); return(yytext);
12 12<* (BEGININ_TAG); return(yytext);
13 13<IN_TAG*> (BEGIN0); return(yytext);
14 14<JAVA* return(yytext);
15 15<STYLE* return(yytext);
16 16<IN_TAG* return(yytext);
17
18
19
20
21 21!test* return(##ทดสอบ##);
22
23
24
25
26
27
28
29
30
31
Ln 32, Col 1 DNS

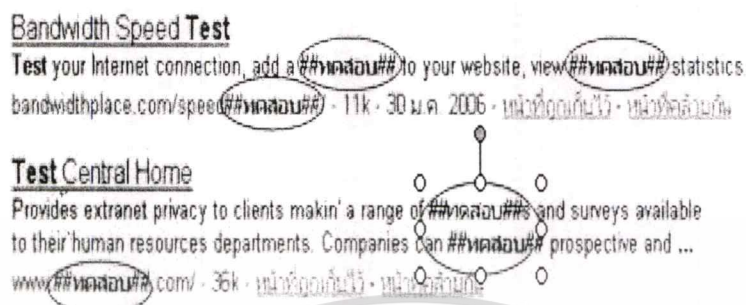
```

รูปที่ 4.4 แสดงการปรับแต่งค่าการกรองที่ file valspeak.l ที่ ICAP Server

หลังจากปรับแต่งค่าคำที่กรองแล้วก็ให้ทำการ Make file และ Restart Service ของ ICAP Sever จากนั้นให้ config คำที่ browser ให้ access เว็บเพจ ผ่านทาง proxy server (โดยที่มีการ config ให้มีการกรองเนื้อหาของเว็บเพจ)โดยให้มีการทดลอง set คำที่ file valspeak.l ของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรม ICAP server โดยทดลองเปลี่ยนจากคำว่า “test” เป็นคำว่า “##ทดสอบ##” จะได้ผลการทดลองดังนี้ ที่ browser ของ client ที่เรียกเว็บเพจเมื่อเป็นคำว่า “test” จะถูกเปลี่ยนเป็นคำว่า “##ทดสอบ##” แสดงดังรูปที่ 4.5

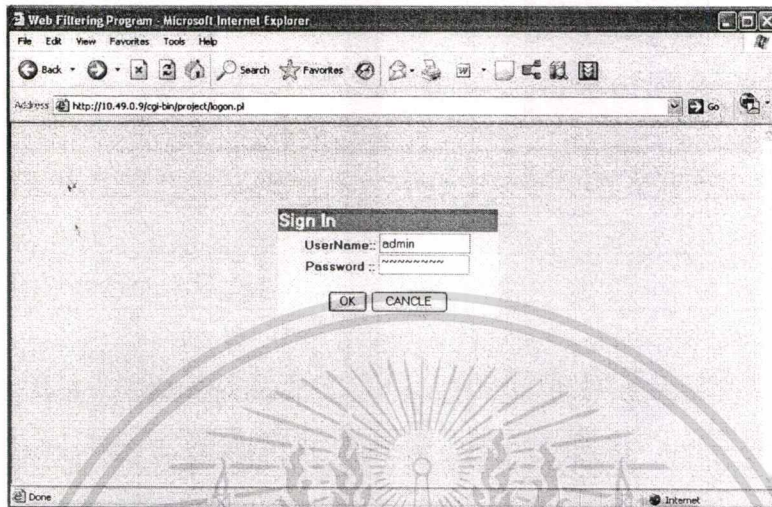


รูปที่ 4.5 แสดงเว็บเพจที่มีการกำหนดให้เข้าถึงเว็บเพจผ่านทาง proxy server

จากการทดลองพบว่า เมื่อเราทำการ config ค่าที่ file valspeak.1 ให้เปลี่ยนค่าจากคำว่า “test” เป็นคำว่า “##ทดสอบ##” แล้ว มีการประมวลผล โดยทำการเปลี่ยนค่าตามค่าที่ config ไว้ แสดงว่ามีการประมวลผลที่เนื้อหาของเว็บเพจทำให้มีการเปลี่ยนแปลงค่าที่จะแสดงผลที่ browser ของ client ดังนั้น เราสามารถนำวิธีการนี้มาประยุกต์ในการกรองเนื้อหาของเว็บเพจ โดยจะทำการเพิ่มค่าที่ต้องการกรองผ่านทาง Web Application และ Web Application ก็ทำการเพิ่มค่าใน file valspeak.1 และทำการ Make file และ Restart Service ของ ICAP Server

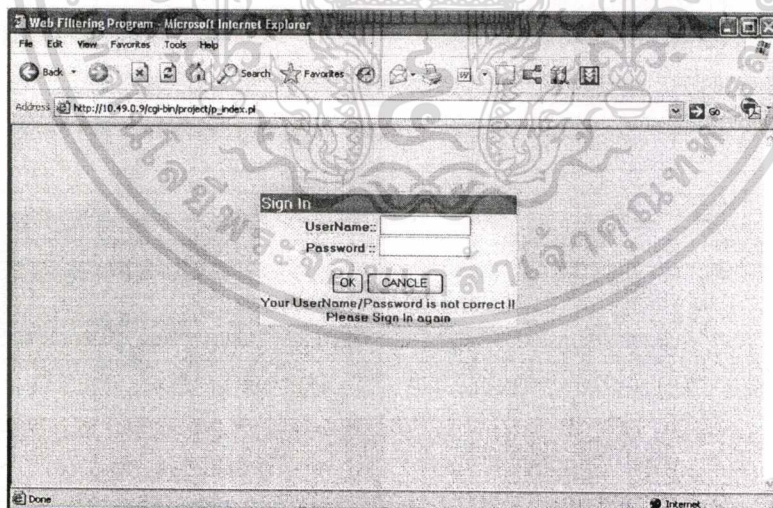
4.3 หน้าจอของระบบ

1. หน้าจอ Sign In เป็นหน้าจอแรกของระบบเพื่อตรวจสอบสิทธิการเป็น Admin ของระบบโดยให้ใส่ Username และ Password แสดงดังรูปที่ 4.6



รูปที่ 4.6 แสดงหน้าจอ Sign In

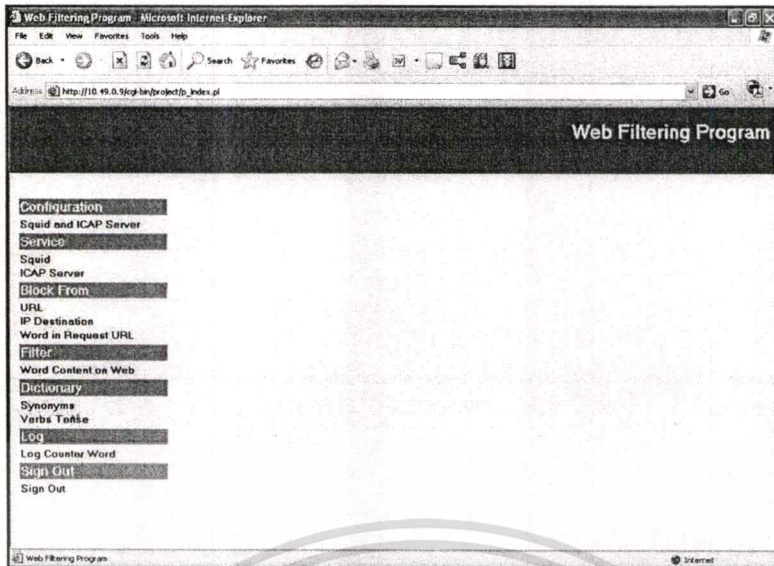
ถ้าใส่ Username Password ผิดจะแสดง error message ดังรูปที่ 4.7



รูปที่ 4.7 แสดง error message ถ้าใส่ Username Password ผิด

ถ้าใส่ Username และ Password ที่ถูกต้องจะแสดง Menu ของโปรแกรมดังรูปที่ 4.7 ที่ในการ Log Out ออกจากระบบให้ click ที่ Log Out

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

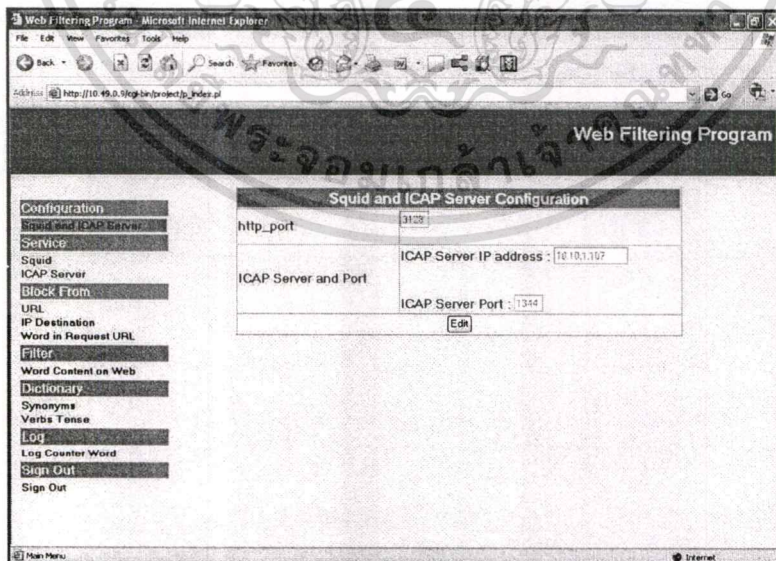


รูปที่ 4.8 แสดง Menu ของโปรแกรม

2. หน้าจอ Configuration Squid and ICAP Server หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งค่า ที่ Proxy Server Squid และ ICAP Server โดยจะสามารถปรับแต่งค่า

- http_port
- ICAP Server IP Address
- ICAP Server Port

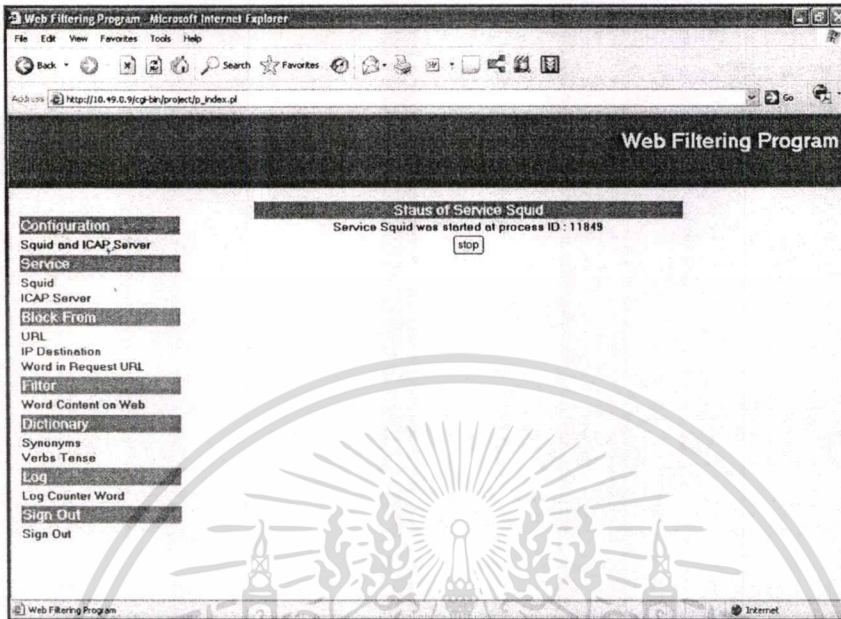
แสดงดังรูปที่ 4.9



รูปที่ 4.9 แสดงหน้าจอ Configuration Squid and ICAP Server

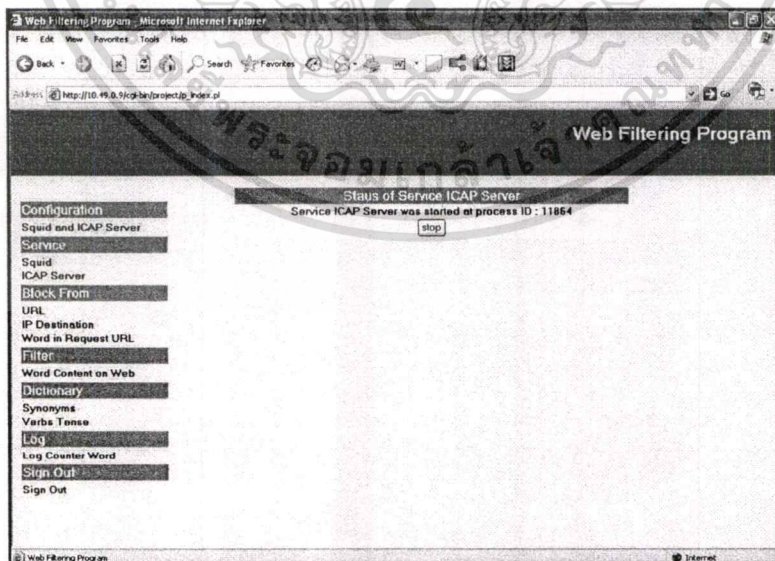
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. หน้าจอ Service Squid หน้าจอนี้จะเป็นหน้าจอแสดง Status ของ Service Squid หน้าจอจะใช้ในการ Stop และ Start Service ของ Squid ได้ด้วยรูปที่ 4.10



รูปที่ 4.10 แสดงหน้าจอ Service Squid

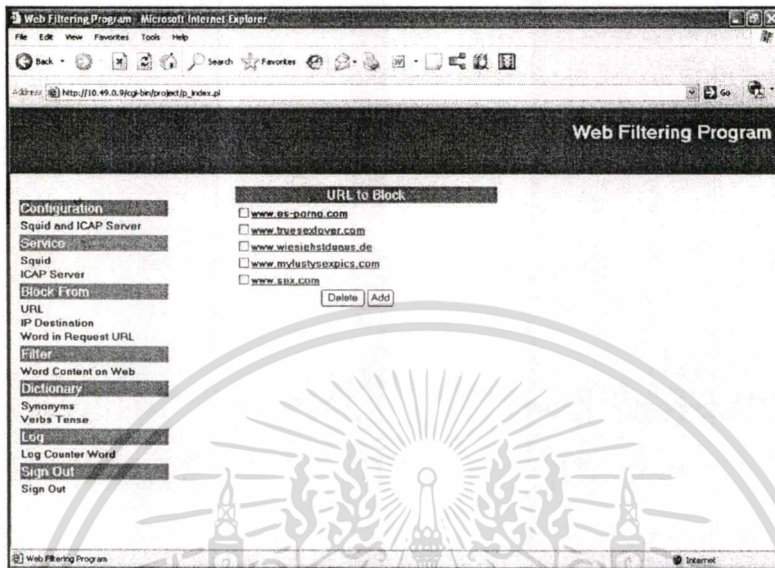
4. หน้าจอ Service ICAP Server หน้าจอนี้จะเป็นหน้าจอแสดง Status ของ Service ICAP Server หน้าจอจะใช้ในการ Stop และ Start Service ของ ICAP Server ได้ด้วยแสดงดังรูปที่ 4.11



รูปที่ 4.11 แสดงหน้าจอ Service ICAP Server

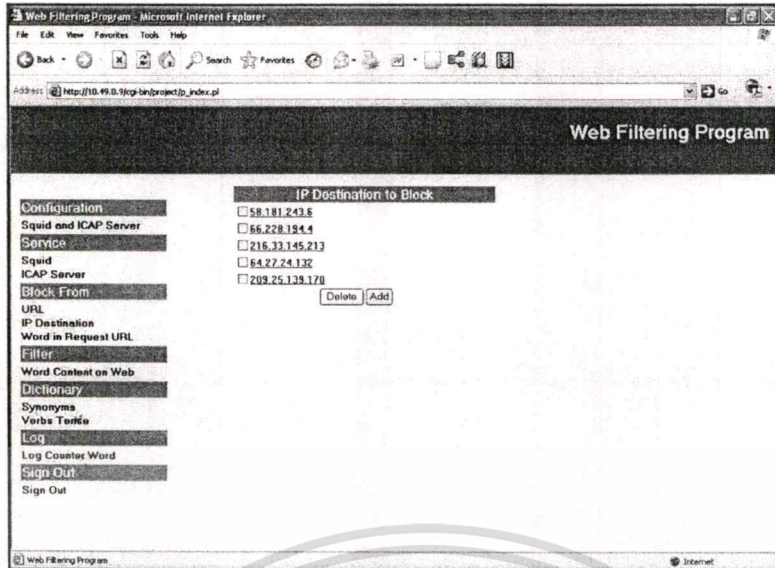
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. หน้าจอ URL to Block หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งในการ Block URL ถ้าทำการ Add URL ไปที่หน้าเว็บเพจนี้แล้วจะทำให้ ไม่สามารถเข้า เว็บดังกล่าวได้ แสดงดังรูปที่ 4.12



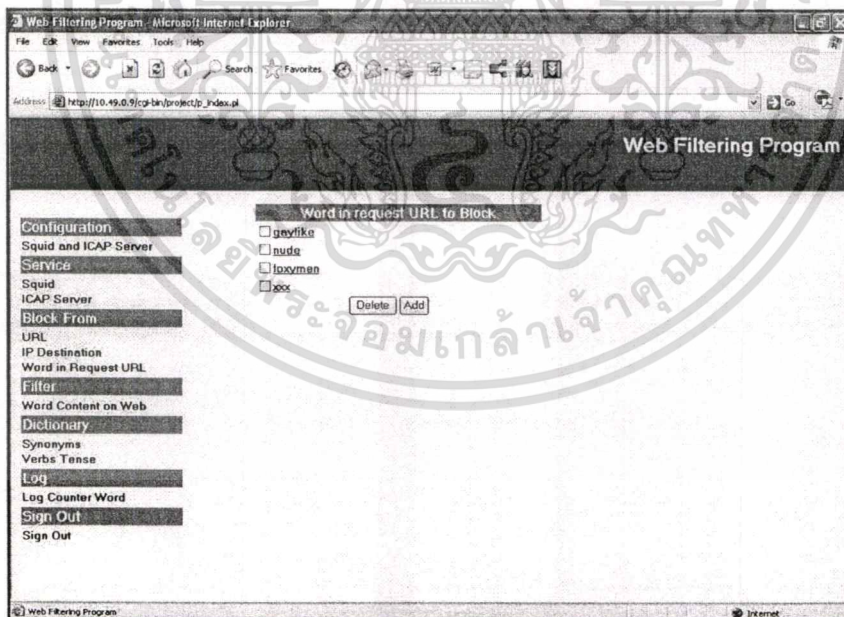
รูปที่ 4.12 แสดงหน้าจอ URL to Block

4. หน้าจอ IP Destination to Block หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งในการ Block IP ถ้าทำการเพิ่ม IP Address ไปที่หน้าเว็บเพจนี้แล้วจะทำให้ ไม่สามารถเข้า เว็บด้วย IP ดังกล่าวได้แสดงดังรูปที่ 4.13



รูปที่ 4.13 แสดงหน้าจอ IP Destination to Block

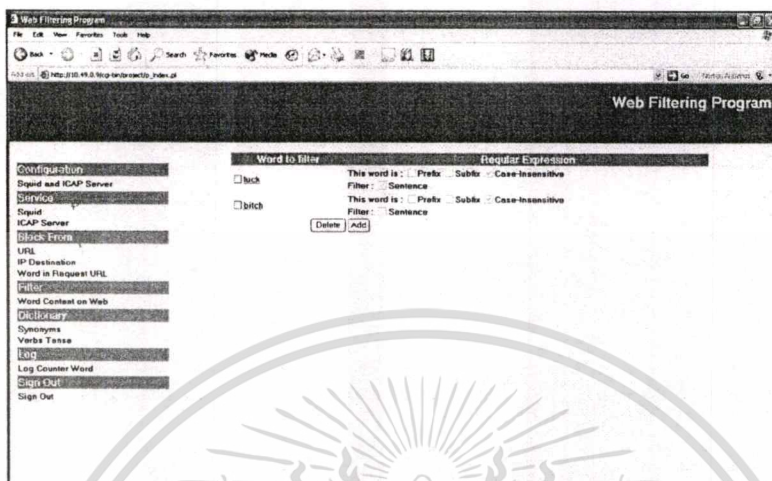
5. หน้าจอ Word in request to Block หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งในการ Block Word ที่ request ถ้าทำการ Add คำไปที่หน้าเว็บเพจนี้แล้วจะทำให้ไม่สามารถเข้าเว็บด้วยคำที่ Block ดังกล่าวได้แสดงดังรูปที่ 4.14



รูปที่ 4.14 แสดงหน้าจอ Word in request to Block

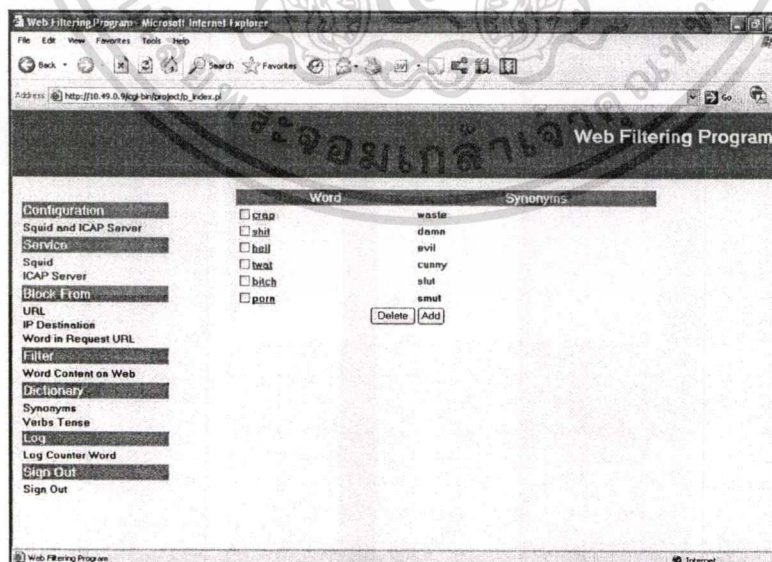
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

6. หน้าจอ Word Content on Web to Filter หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งในการกรองคำในเว็บเพจถ้าทำการเพิ่มคำไปที่หน้าเว็บเพจนี้แล้วจะทำให้กรองคำที่ใส่เข้าไปได้แสดงดังรูปที่ 4.15



รูปที่ 4.15 แสดงหน้าจอ Word Content on Web to Filter

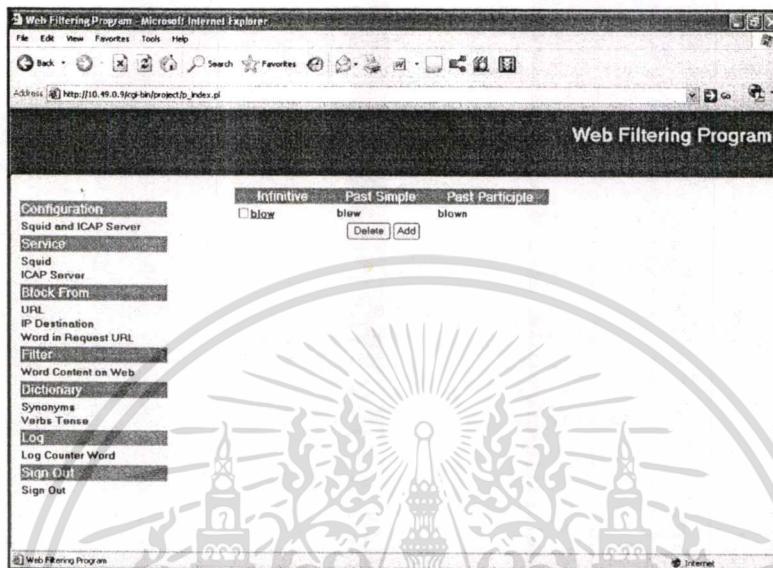
7. หน้าจอ Dictionary Synonyms หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งในการเพิ่มคำที่ต้องการกรองในแบบ Synonym ถ้า Add คำไปที่หน้าเว็บเพจนี้แล้วจะทำให้การกรองมีการกรองคำ Synonym เพิ่มขึ้นมาแสดงดังรูปที่ 4.16



รูปที่ 4.16 แสดงหน้าจอ Dictionary Synonyms

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

8. หน้าจอ Dictionary Verbs Tense หน้าจอนี้จะเป็นหน้าจอที่ใช้ในการปรับแต่งในการเพิ่มคำที่ต้องการกรองในแบบ Verb Tense ถ้าเพิ่มคำไปที่หน้าเว็บเพจนี้แล้วจะทำให้การกรองมีการกรองคำ Verb Tense เพิ่มขึ้นมาแสดงดังรูปที่ 4.17



รูปที่ 4.17 แสดงหน้าจอ Dictionary Verbs Tense

9. หน้าจอ Log Counter Word หน้าจอนี้จะเป็นหน้าจอที่ใช้แสดง access log ที่ user ได้เข้าถึงเว็บเพจโดยจะแสดง เวลา IP Address URL to Access และ Word Filter Count แสดงดังรูปที่ 4.18

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Web Filtering Program - Microsoft Internet Explorer

Address: http://10.49.0.9/cgi-bin/projectIp_index.pl

Web Filtering Program

Configuration	Date to Access	IP Address	URL to Access	Word Filter Count
Squid and ICAP Server	18 Oct 2006 11:13:02	10.49.0.13	http://www.jokes.in.th/no15411560.htm	13
Service	18 Oct 2006 11:14:26	10.49.0.13	http://www.google.co.th/search?	1
Squid	18 Oct 2006 11:16:09	10.49.0.13	http://search.pantip.com/cgi-bin/ss?	1
ICAP Server	18 Oct 2006 11:16:59	10.49.0.13	http://www.jokes.in.th/no15611580.htm	8
Block From	18 Oct 2006 11:17:04	10.49.0.13	http://www.jokes.in.th/no15811600.htm	2
URL	18 Oct 2006 11:26:20	10.49.0.13	http://www.jokes.in.th/No12211240.htm	1
IP Destination	18 Oct 2006 12:32:46	10.10.1.101	http://www.google.co.th/search?	40
Word in Request URL	18 Oct 2006 12:42:03	10.10.1.101	http://www.google.co.th/search?	17
Filter	18 Oct 2006 12:42:28	10.10.1.101	http://www.thebitchgirls.us/	1
Word Content on Web				
Dictionary				
Synonyms				
Verbs Tense				
Log				
Log Counter Word				
Sign Out				

รูปที่ 4.18 แสดงหน้าจอ Log Counter Word

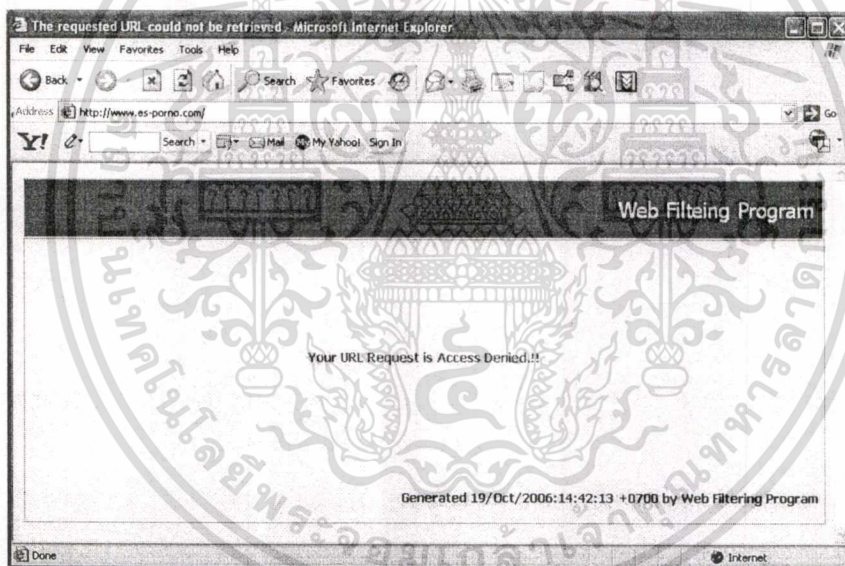
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 การทดสอบการทำงานของโปรแกรม

- URL Request to block ถ้ามีการ URL อยู่ใน list URL to Block แสดงดังรูปที่ 4.19 จะทำให้ไม่สามารถเข้าเว็บที่ Block URL นั้นได้และจะมี Error message แสดงดังรูปที่ 4.20



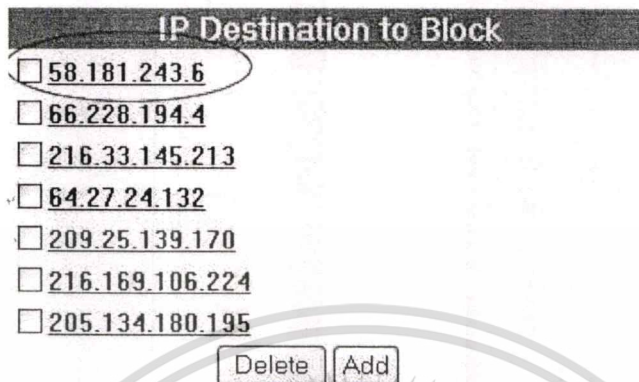
รูปที่ 4.19 แสดง list URL to Block



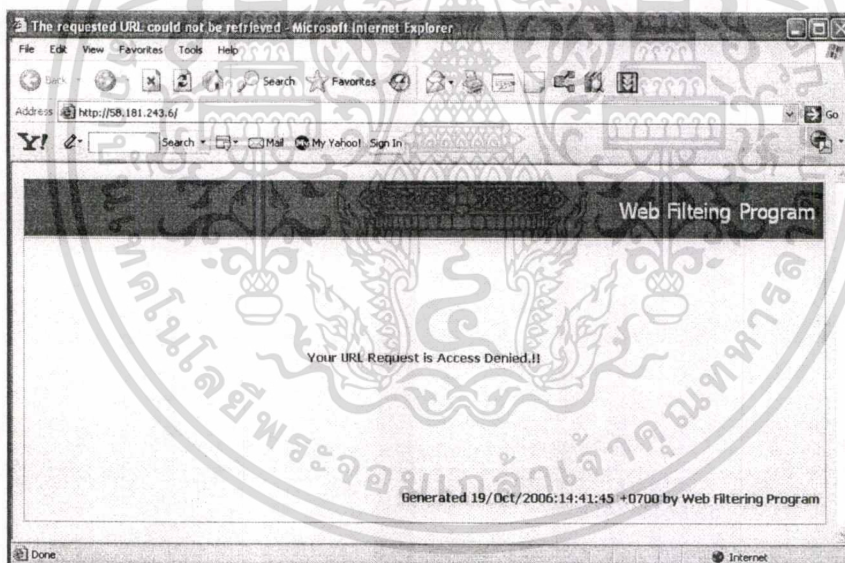
รูปที่ 4.20 แสดง Error message ที่ไม่สามารถเข้าเว็บที่ Block URL นั้นได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- IP Destination to Block ถ้ามี IP Address อยู่ใน List IP Destination to Block แสดงดังรูป 4.21 ที่ จะทำให้ไม่สามารถเข้าเว็บที่ IP Address นั้นได้และจะมี Error message แสดงดังรูปที่ 4.22



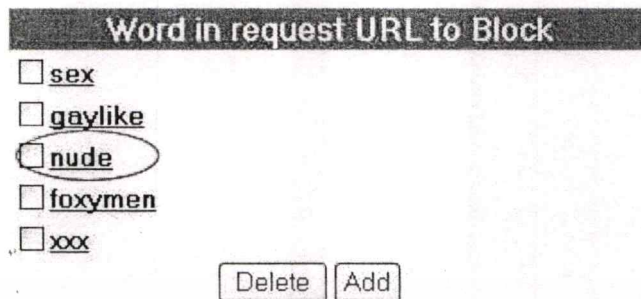
รูปที่ 4.21 แสดง List IP Destination to Block



รูปที่ 4.22 แสดง Error message ที่ไม่สามารถเข้า IP Address นั้นได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Word in Request URL to Block ถ้าคำมีที่ Block อยู่ใน Word in Request URL to Block แสดงดังรูปที่ 4.23 จะไม่สามารถเข้าเว็บนั้นได้และจะมี Error message แสดงดังรูปที่ 4.24



รูปที่ 4.23 แสดงคำมีที่ Block อยู่ใน Word in Request URL to Block



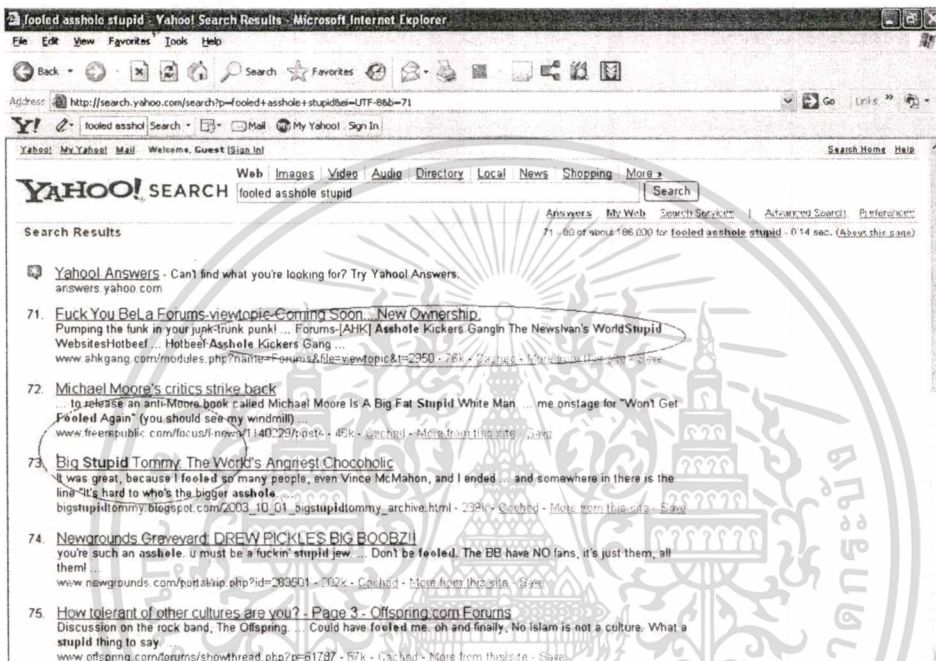
รูปที่ 4.24 แสดง Error message ที่ไม่สามารถเข้าเว็บนั้นได้เมื่อมีคำที่อยู่ใน Word in Request URL to Block

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรองคำที่อยู่บนเว็บเพจโดยทดลอง Feature ของโปรแกรมดังนี้

- Prefix, Subfix
- Case- Insensitive
- Dictionary Synonym

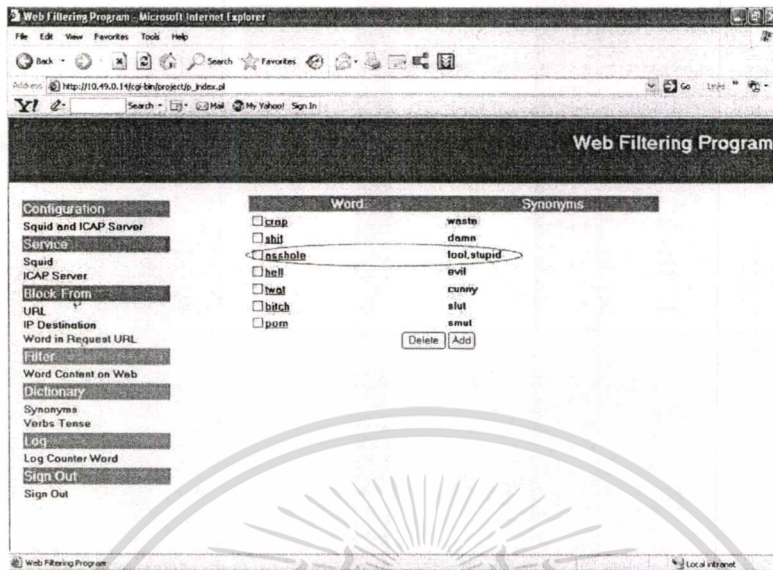
ก่อนทำการทดลองให้ทำการเก็บหน้าจอก่อนการเพิ่มคำที่ให้กรองจะแสดงผลดังรูปที่ 4.25 มีคำว่า “Stupid” “Asshole” “Fooled”



รูปที่ 4.25 แสดงการเก็บหน้าจอก่อนการเพิ่มคำที่ให้กรอง

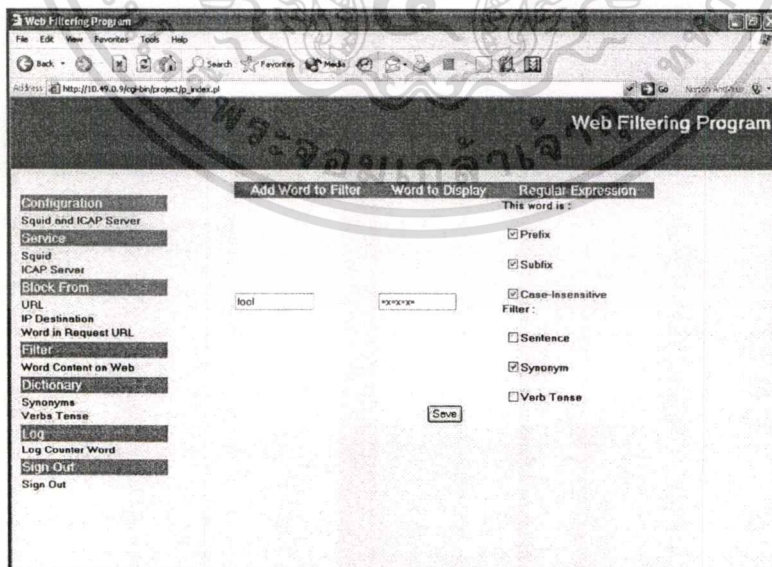
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Feature Dictionary Synonym แสดงดังรูปที่ 4.26



รูปที่ 4.26 แสดง Feature Dictionary Synonym

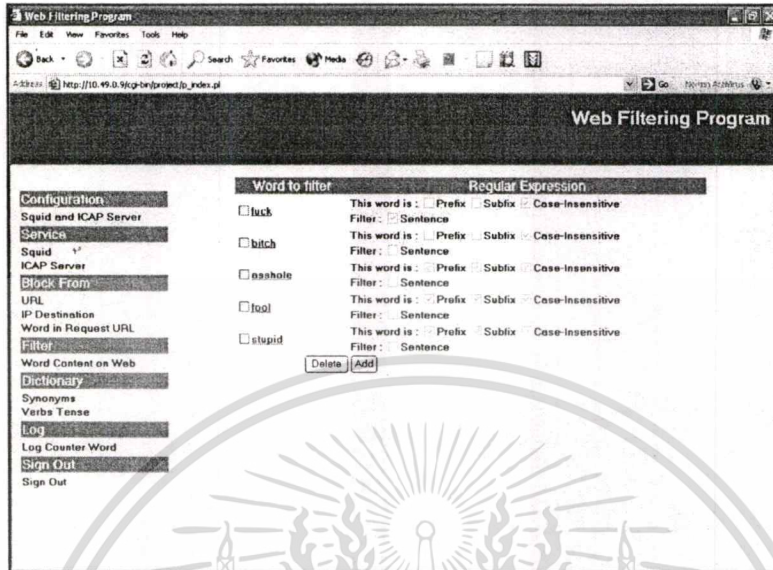
เราจะทดสอบเพิ่มคำว่า fool เข้าไปเพื่อทดสอบ Dictionary Synonym การทดสอบเพิ่มคำให้พิมพ์คำว่า fool ลงไปที่ Add Word to Filter แล้ว ใส่คำที่ต้องการให้แสดงผลที่ Word to Display จากนั้นให้เลือกลักษณะของการกรองคำเลือกที่ Prefix Subfix Case-Insensitive Synonym แล้ว click ที่ Save แสดงดังรูปที่ 4.27



รูปที่ 4.27 แสดงการเพิ่มคำเพื่อทดสอบการกรองคำ

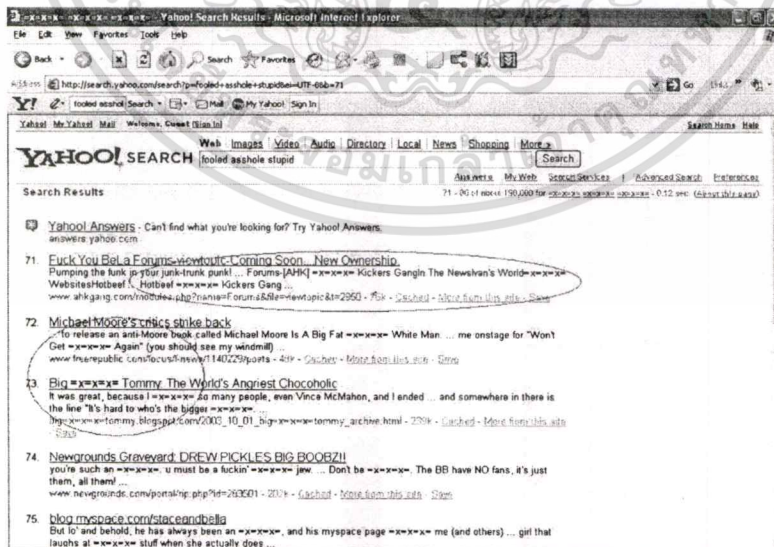
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรมจะทำการเข้าไปค้นคำว่า fool ใน Dictionary ของโปรแกรมและทำการเพิ่มคำที่
ต้องการกรอง Synonym ลงไปที่ Word to Filter แสดงดังรูปที่ 4.28



รูปที่ 4.28 แสดงคำที่เพิ่มจาก Dictionary Synonym

แล้วทดลองให้ Client ที่ Config คำที่ Browser ให้ใช้ Internet ผ่าน Proxy Server Squid
จะให้ผลลัพธ์แสดงดังรูปที่ 4.29 โปรแกรมจะทำให้คำว่า “Stupid” “Asshole” “Fooled”
แสดงผลเป็น “=X=X=X=” แสดงว่า Feature ที่ปรับแต่งค่าไปได้ทำงาน



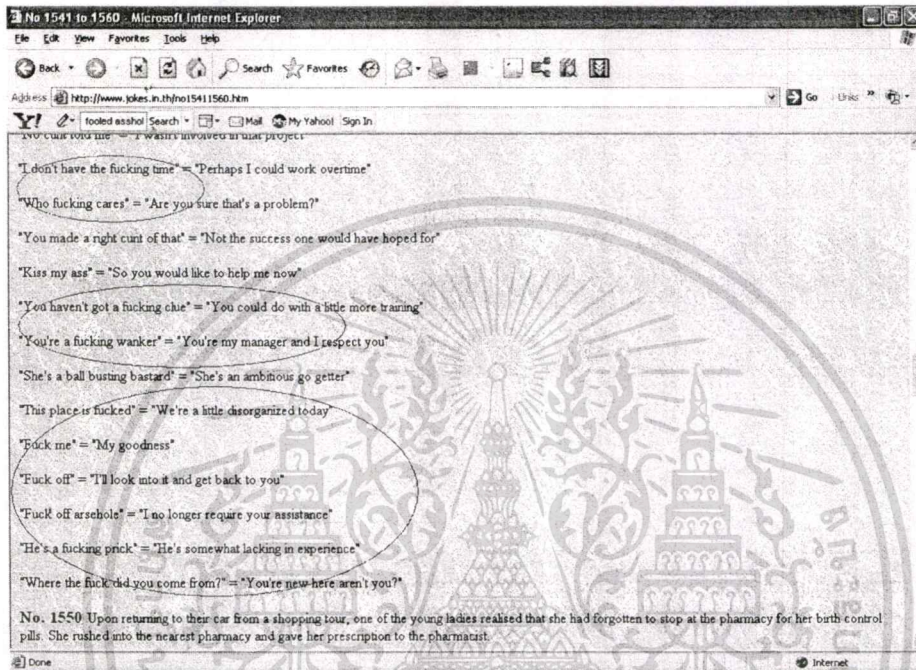
รูปที่ 4.29 แสดงผลลัพธ์ของการเพิ่มคำ โดยจะแสดงที่ Browser ที่แสดงผลที่ Client

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กรอกราคำที่อยู่บนเว็บเพจโดยทดลอง Feature ของโปรแกรมดังนี้

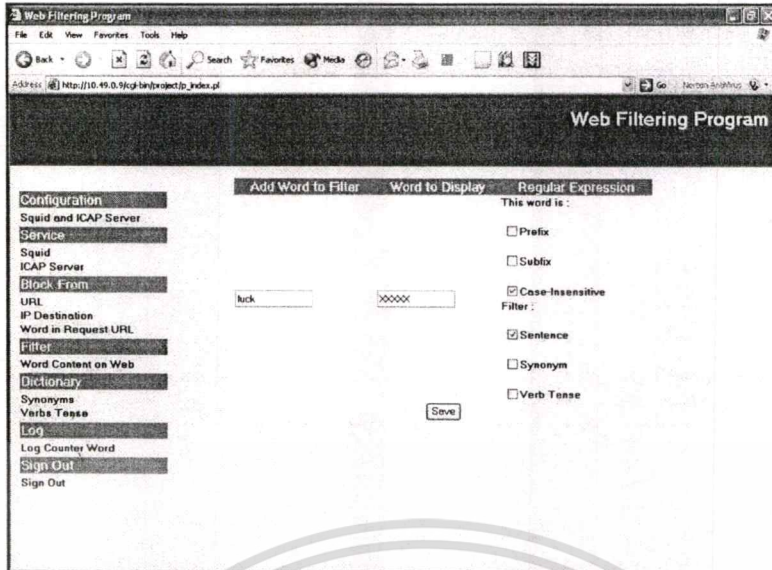
- Case- Insensitive
- Sentence

ก่อนทำการทดลองให้ทำการเก็บหน้าจอก่อนการเพิ่มคำที่ให้กรองประโยคที่มีคำว่า fuck อยู่ 10 ประโยค แสดงดังรูปที่ 4.30



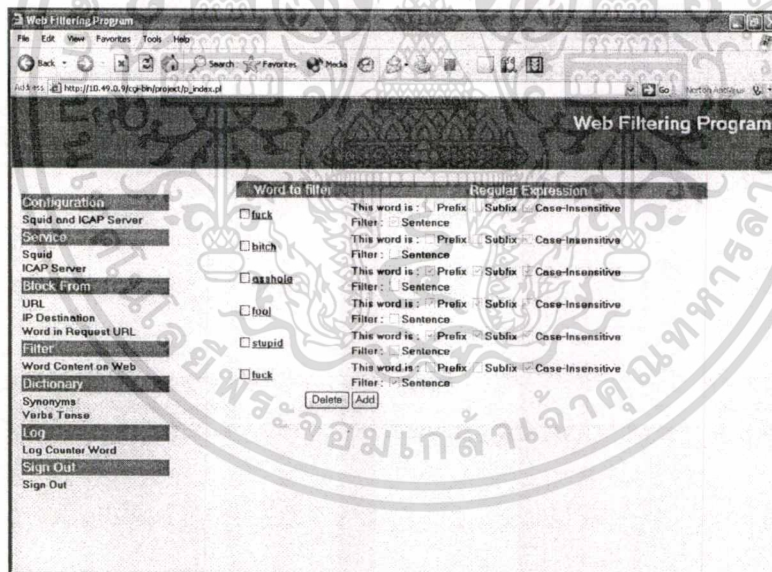
รูปที่ 4.30 แสดงการเก็บหน้าจอก่อนการเพิ่มคำที่ให้กรองประโยค

ทดลองเพิ่มคำว่า “fuck” เข้าไปในโปรแกรม โดยจะให้แสดงผลเป็น “XXXXXX” ลักษณะการกรองแบบ Case-Insensitive และกรองทั้งประโยคแสดงดังรูปที่ 4.31



รูปที่ 4.31 แสดงการเพิ่มคำเพื่อทดสอบการกรองคำ

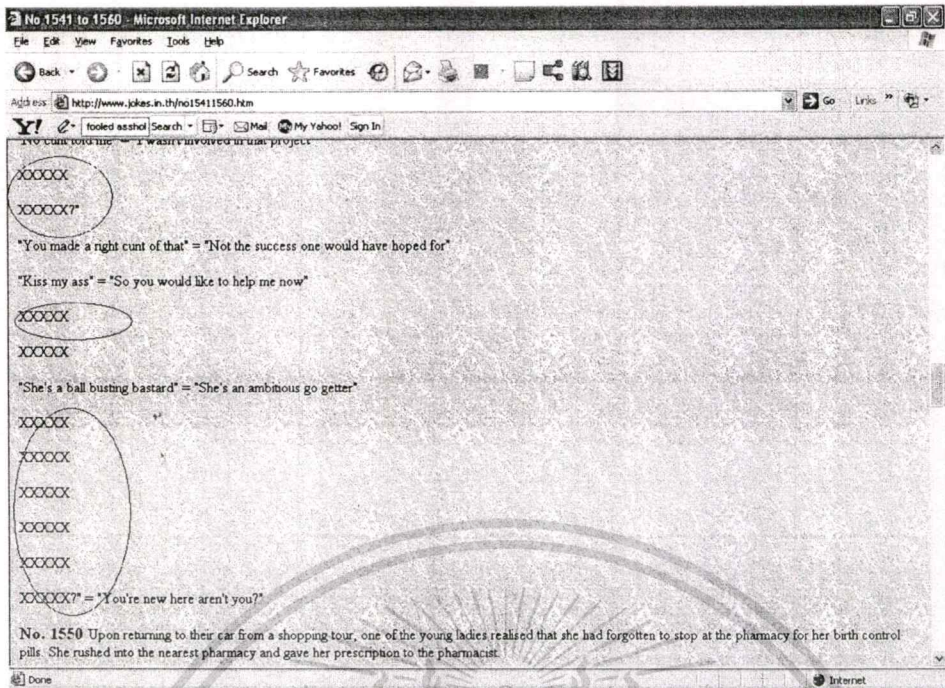
Click ที่ Save จะทำการเพิ่มคำว่า fuck เข้ามาที่โปรแกรมแสดงดังรูปที่ 4.32



รูปที่ 4.32 แสดงการเพิ่มคำเข้ามาที่โปรแกรม

แล้วทดลองให้ Client ที่ Config คำที่ Browser ให้ใช้ Internet ผ่าน Proxy Server Squid จะให้ผลลัพธ์แสดงดังรูปที่ 4.33 โปรแกรมจะทำให้ประโยคที่มีคำว่า “fuck” แสดงผลเป็น “XXXXX” แสดงว่า Feature ที่ปรับแต่งค่าไปได้ทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.33 แสดงผลลัพธ์ของการเพิ่มคำ โดยจะแสดงที่Browser ที่แสดงผลที่ Client



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

บทสรุป และข้อเสนอแนะ

5.1 บทสรุป

ในโครงการนี้ได้ดำเนินการศึกษาหาข้อมูลและวิธีในการกรองคำที่ไม่เหมาะสมที่แสดงบนหน้าเว็บเพจ โดยให้ทำการกรองคำโดยที่ทำงานผ่าน Proxy Server Squid และ Protocol ICAP โดยได้นำเทคโนโลยีของ Web Application เข้ามาช่วยในการจัดการ Block Web Site ที่ไม่เหมาะสม และกรองคำที่ไม่เหมาะสม ในการใช้ Web Application มาช่วยนั้นทำให้จัดการได้ง่ายยิ่งขึ้น โดยที่ผู้ดูแลระบบไม่จำเป็นต้องมีความรู้เรื่องการทำงานของ Proxy Server และ Protocol ICAP มากนัก

จากการที่ได้วิเคราะห์และออกแบบโปรแกรมนั้น ได้ทำการศึกษา การทำงานของ Proxy Server การทำงานของ Protocol ICAP ตลอดจนการปรับแต่งค่าต่างเพื่อทำให้กรองคำได้ โดยสามารถกรองคำได้ในลักษณะที่เป็น Prefix , Subfix, Case-Insensitive, ประโยค และมี Dictionary Synonym และ คำกริยา 3 ช่องใน Program เพื่อให้การกรองมีประสิทธิภาพมากยิ่งขึ้น

5.2 ข้อดีและข้อเสียของระบบ

5.2.1 ข้อดีของระบบ

- ในการทำการปรับแต่งค่าคำกรองของโปรแกรมใช้งานง่ายโดยให้ทำการเพิ่มค่าผ่านทางเว็บเพจได้
- มี Dictionary ที่ช่วยให้การกรองคำมีประสิทธิภาพมากขึ้น
- มีการประยุกต์ให้ใช้งานในการกรองลักษณะ Prefix, Subfix และ ประโยคได้
- มีการแสดงผลของ log access เพื่อใช้นับจำนวนคำที่กรองได้เพื่อช่วยในการตรวจสอบว่าผู้ใช้ได้เข้าเว็บ ไซด์อะไรไปและพบว่า มีคำที่ถูกกรองก็คำ

5.2.2 ข้อเสียของระบบ

- ในการปรับแต่งค่าคำกรองนั้นจะต้องไปทำการ Make file และ Restart Service ICAP Server ซึ่งจะส่งผลให้ User จะไม่สามารถเข้า Web Site ได้จนกว่าจะ Restart เสร็จ
- ยังไม่สามารถกรองคำที่เป็นภาษาไทยได้อย่างมีประสิทธิภาพเนื่องจากปัญหาเรื่องการตัดคำในภาษาไทย

5.3 ข้อเสนอแนะ

- ปรับปรุงโปรแกรมให้สามารถ ปรับตั้งค่าใหม่โดยไม่ต้องทำการ Make file และ Restart Service ICAP Server ใหม่
- ปรับปรุงให้โปรแกรมสามารถทำการกรองภาษาไทยได้อย่างมีประสิทธิภาพ



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- Ari Luotonen, **Web Proxy Servers**. United States of America : Prentice Hall PTR, 1997.
- Basile Starynkevitch, Mohamed Daoudi, Chridtophe Tombelle, Huicheng Zheng, Jose-Maria Gomez-Hidalgo, Paolo Allengrini, Nicoletta Calzolari, Simone Marchi, Simonetta Montemagni and Stafan Guerra, **POESIA A Software Architecture Definition Document Deliverable 3.1**. [Online]. Available: http://www.poesia-filter.org/pdf/Deliverable_3_1.pdf, 2002.
- Cisco ACNS Software Configuration Guide for Centrally Managed Deployments, Release 5.2. **Chapter 11 Configuring ICAP on Standalone Content Engines**. [Online]. http://www.cisco.com/en/US/products/sw/conntsw/ps491/products_configuration_guide_chapter09186a008023661f.html#anchor_to_feedback
- David Guerrero, **Caching The Web : Improve your users browsing and save your bandwidth by using proxy servers to cache web pages**. [online]. Available: <http://www.david-guerrero.com/papers/squid/squid.htm>, 1999.
- Geetha Manjunath and Hewlett Packard Labs ,**ICAP Client Support**. [Online]. Available: <http://devel.squid-cache.org/icap/>, 2004.
- J. Elson and A.Cerpa, **Internet Content Adaptation Protocol (ICAP)**. [Online]. Available: <http://www.i-cap.org/spec/rfc3507.txt>, 2003.
- Network Appliance, **Internet Content Adaptation Protocol (ICAP)**. [online]. Available: http://www.i-cap.org/docs/icap_whitepaper_v1-01.pdf, 2001.
- Riadh Elloumi andFares Triki, **Shweby User Manual**. [Online]. Available: http://shweby.sourceforge.net/doc/man-0_9_2.html
- Term squid, **SQUID Frequently Asked Questions**. [online]. Available: <http://www.squid-cache.org/Doc/FAQ/FAQ.html>, 2004



ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.

การติดตั้งโปรแกรมและการใช้งาน

ความต้องการของระบบ

สำหรับเครื่องที่สามารถติดตั้งได้นั้นมีรายละเอียดดังนี้
ระบบปฏิบัติการLinux พร้อมติดตั้ง โปรแกรมต่างดังนี้

- โปรแกรม Proxy Server Squid version ที่ Support ICAP Client
- โปรแกรม ICAP Server
- โปรแกรม Apache Web Server
- โปรแกรม Interpreter Perl
- โปรแกรม EGDIT

การติดตั้ง

ให้ทำการ Copy file project.tar ไปที่ path cgi-bin ของ Apache ก็จะสามารถใช้งานได้

การใช้งานระบบ

เข้าไปใช้งานระบบโดยผ่านทางเว็บเพจก็จะใช้งานได้ทันทีโดยใช้ Username = admin ,
Password = webfilter ก็จะใช้งานได้ทันที

ประวัติผู้เขียน

ชื่อ	นางสาวณัฐณี วิทยาปัญญา
วันเกิด	13 ธันวาคม 2521
สถานที่เกิด	กรุงเทพมหานคร
ประวัติการศึกษา	
ปริญญาตรี	วิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์ มหาวิทยาลัย ศรีนครินทรวิโรฒ
การทำงาน	Programmer บริษัท อัลฟาอินโฟ จำกัด System Engineer บริษัท ซีดีจี ซิสเต็มส์ จำกัด



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้