

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

โปรแกรมคัดกรองข้อมูลเว็บ
WEB CONTENT FILTERING



ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2549

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง

โปรแกรมคัดกรองข้อมูลเว็บ
WEB CONTENT FILTERING



เลขหมู่.....
เลขทะเบียน.....
วัน,เดือน,ปี.....

72976

26 ส.ย. 2550

b. 11 พ.ย. 2548
i.....

ปริญญาานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมคอมพิวเตอร์
คณะวิศวกรรมศาสตร์
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง
ปีการศึกษา 2549

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญาโทปีการศึกษา 2549

ภาควิชาวิศวกรรมคอมพิวเตอร์

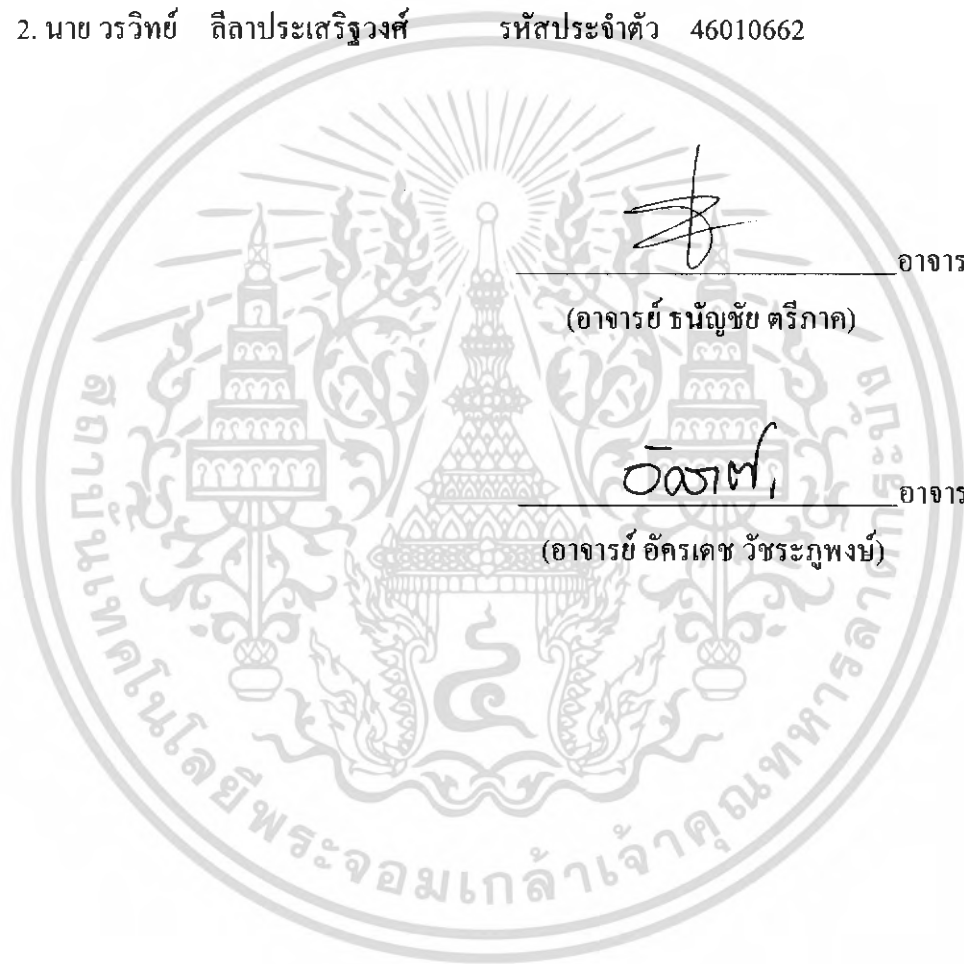
คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

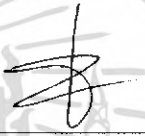
เรื่อง โปรแกรมคัดกรองข้อมูลเว็บ

WEB CONTENT FILTERING

ผู้จัดทำ

1. นาย พฤกษ์ รัตนานุสนธิ รหัสประจำตัว 46010499
2. นาย วรวิทย์ ลีลาประเสริฐวงศ์ รหัสประจำตัว 46010662




อาจารย์ที่ปรึกษา
(อาจารย์ ธานุชัย ตรีภาค)


อาจารย์ที่ปรึกษา
(อาจารย์ อัครเดช วัชรพงษ์)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โปรแกรมคัดกรองข้อมูลเว็บ

นาย พฤษัย รัตนสุนทร 46010499
นาย วรวิทย์ ถีลาประเสริฐวงศ์ 46010662
อาจารย์ ธนัญชัย ศรีภาค อาจารย์ที่ปรึกษา
อาจารย์ อัครเดช วัชรระภูพงษ์ อาจารย์ที่ปรึกษาร่วม
ปีการศึกษา 2549

บทคัดย่อ

เนื่องจากในปัจจุบันได้มีเว็บไซต์ต่างๆ เกิดขึ้นมากมาย โดยมีเว็บไซต์อยู่จำนวนหนึ่งที่เป็นเว็บไซต์ผิดกฎหมาย หรือมีเนื้อหาที่ไม่เหมาะสมต่อเด็ก และเยาวชน ซึ่งเว็บเหล่านี้ทำให้เกิดปัญหาทางสังคมต่างๆ ตามมาได้ ในภายหลัง เช่น ขายแผ่นซีดีลามกผ่านทางเว็บไซต์ เว็บลามกอนาจาร ทำให้วัยรุ่นเกิดความอยากรู้อยากลอง โดยที่พ่อแม่ของบุตรหลานไม่สามารถที่จะควบคุมได้ ถึงแม้ว่าในปัจจุบันทางภาครัฐได้จัดตั้งโครงการ Cyber Inspector เพื่อทำการตรวจสอบเว็บไซต์ต่างๆ ซึ่งดูแลได้ไม่ทั่วถึง จึงจัดทำโครงการ โปรแกรมคัดกรองข้อมูลเว็บ เพื่อให้การใช้งานเว็บไซต์บนเครือข่ายนั้นมีความปลอดภัยมากยิ่งขึ้น สามารถติดตั้งและใช้งานร่วมกับ เว็บเบราว์เซอร์ ต่างๆ ได้ โดยทำงานเหมือนกับ ฟร็อกซี่ส่วนตัว ซึ่งมีการพัฒนาส่วนหลักๆ คือ การพัฒนาโปรแกรมให้เป็นฟร็อกซี่ส่วนตัว สามารถทำงานร่วมกับเว็บเบราว์เซอร์ บนระบบปฏิบัติการวินโดวส์ , การพัฒนาเพื่อให้สามารถคัดกรองในส่วนของ ข้อความ ที่ไม่เหมาะสมบนเว็บไซต์ , การพัฒนาเพื่อให้สามารถคัดกรองในส่วนของรูปภาพ อัลกอริทึมการประมวลผลทางด้านรูปภาพ , การพัฒนาเพื่อให้สามารถคัดกรองในส่วนของสคริปต์ ที่เป็นปัญหาต่างๆ , สามารถให้ผู้ปกครองควบคุมปรับแต่งได้ รวมไปถึงสามารถใช้งานได้ทั้ง ไฟร์ฟอกซ์ และ เว็บเบราว์เซอร์ อื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

WEB CONTENT FILTERING

Mr. Prueak Thananusol

Mr. Worawit Leelaprasertwong

Mr. Thananchai Treepak Advisor

Mr. Akkradach Watcharapupong Co-Advisor

Academic Year 2006

ABSTRACT

The increasing prevalence of new websites, some of which are illegal or contain inappropriate content for youth such as online pornographic CD sales, pornographic websites, contributes to social issues, which are hard to be controlled. Although the government has already launched the campaign “Cyber Inspector” so as to evaluate those websites, still it does not cover throughout. Hence, our project is aimed to filter the content of websites for more security in network system. It can be installed and operate together with Web Browsers by working as Personal Web Proxy, which has been developed throughout our project mainly in 6 parts. To begin with, Personal Web Proxy has been developed so that it can be cooperated with Web Browser and Windows OS. In addition, text filtering ability (Text Detection) particularly for improper text content on websites as well as graphic filtering ability (Image Processing algorithms) and script filtering ability for those which are prone to problem has been implemented. For parental concerns, this program can be configured in their preferable ways. Last but not least, this program can operate with any Web Browsers such as Firefox.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

กิตติกรรมประกาศ

การทำปฏิญานิพนธ์ฉบับนี้อาจไม่สามารถสำเร็จลุล่วงไปด้วยดีได้ หากไม่ได้การสนับสนุน การคอยให้คำแนะนำ และการให้ความช่วยเหลืออย่างสม่ำเสมอจาก อาจารย์ ธานีชัย ตรีภาค และอาจารย์ อัครเดช วัชรระภูพงษ์ รวมไปถึงคณาจารย์ประจำภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังทุกท่าน ที่เป็นผู้อบรมสั่งสอนให้วิชาความรู้ต่างๆ ทั้งทางด้านวิชาความรู้และคุณธรรม แก่พวกเรามาโดยตลอด

ขอกราบขอบพระคุณ บิดา มารดา ผู้เป็นที่เคารพรักรักยิ่งของคณะผู้จัดทำ โดยท่านได้อุปการะเลี้ยงดู ให้โอกาสในด้านการศึกษาอย่างถึงที่สุด และคอยห่วงใย ทางคณะผู้จัดทำจึงขอกราบขอบพระคุณเป็นอย่างสูง

สุดท้ายนี้ขอขอบคุณเพื่อนๆทุกคนที่ให้คำแนะนำและขอขอบพระคุณทางภาควิชาวิศวกรรมคอมพิวเตอร์ และสถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบังที่ได้อำนวยความสะดวกในด้านต่างๆ อาทิ เช่น สถานที่ที่ใช้ศึกษาค้นคว้า และห้องวิจัยที่ใช้สำหรับทำปฏิญานิพนธ์ฉบับนี้

คุณค่าและประโยชน์อันพึงมาจากปฏิญานิพนธ์ฉบับนี้ ข้าพเจ้าขอมอบแด่ผู้มีพระคุณทุกท่าน

คณะผู้จัดทำ

นาย พลฤกษ์ ธานานุสนธิ์

นาย วรวิทย์ ธิลาประเสริฐวงศ์

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VIII
สารบัญรูป.....	IX
บทที่ 1 บทนำ.....	1
1.1 ความสำคัญและที่มาของโครงการ.....	1
1.2 วัตถุประสงค์ของโครงการ.....	1
1.3 ขอบเขตของโครงการ.....	2
1.4 วิธีการดำเนินการ.....	2
1.5 ประโยชน์ที่คาดว่าจะได้รับ.....	3
1.6 ส่วนประกอบของปฏิญานิพนธ์.....	3
บทที่ 2 โครงสร้างของซอฟต์แวร์ และการคัดกรองข้อมูล.....	5
2.1 โครงสร้างซอฟต์แวร์.....	5
2.2 ส่วนประกอบที่สำคัญที่ได้พัฒนาในโปรแกรมคัดกรองข้อมูลเว็บ.....	8
2.3 อัลกอริทึมพื้นฐานที่ใช้ในโปรแกรม.....	9
2.3.1 สถิติ.....	10
2.3.2 เมตริกซ์อัลจิบรา (Matrix Algebra).....	12
2.3.3 หลักการวิเคราะห์แยกแยะส่วนประกอบ (Principal Components Analysis: PCA).....	13
2.3.4 K-Means Clustering and K-nearest neighbor algorithm.....	15
2.3.5 K-Nearest Neighbor (KNN).....	16
2.4 การคัดกรองข้อความ (Text Detection).....	17
2.5 การคัดกรองรูปภาพ (Image Detection).....	18
2.5.1 พื้นฐานและระบบ โครงสร้างสี.....	18
2.5.1.1 ระบบโครงสร้างสีอาร์จีบี (RGB Color Model).....	18
2.5.1.2 ระบบ โครงสร้างสีวายซีบีซีอาร์ (YCbCr Color Model).....	19
2.5.1.3 การแปลงรูปแบบสี.....	19
2.5.2 กระบวนการค้นหาพื้นที่สีผิว.....	20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้า
2.5.3 กระบวนการปรับความคมชัดของภาพ.....	24
2.5.3.1 First derivative.....	24
2.5.3.2 Second derivative.....	24
2.5.3.3 การใช้ First Derivative ในการเพิ่มคุณภาพของภาพ.....	25
2.5.3.4 การใช้ Second Derivative ในการเพิ่มคุณภาพของภาพ.....	26
2.5.3.5 การตรวจหาเส้นขอบด้วย กระบวนการโซเบล (Sobel operators).....	26
2.5.4 กระบวนการทดสอบ.....	28
2.5.4.1การกำจัดภาพขนาดเล็กกว่ากำหนด (Size and palette Analysis).....	29
2.5.4.2 การตรวจหาสีผิวมนุษย์ (Skin detection module).....	29
2.5.4.3 การตรวจสอบคุณสมบัติของภาพ (Feature Extraction Module).....	30
2.5.4.4 การจัดประเภทของภาพด้วยคุณสมบัติ (Decision Classifier).....	32
บทที่3 หลักการพื้นฐานบนเครือข่าย โปรแกรม.....	34
3.1 โพรโตคอลเอชทีทีพี (HTTP: HyperText Transfer Protocol).....	34
3.1.1 วิธีการติดต่อของโพรโตคอล HTTP.....	34
3.1.2 โครงสร้างของโพรโตคอล HTTP.....	35
3.2 พร็อกซี่ส่วนตัว (Personal Web Proxy).....	37
3.2.1 พร็อกซี่ (Proxy) คืออะไร.....	37
3.2.2 พร็อกซี่ส่วนตัว (Personal Web Proxy) คืออะไร.....	38
3.3 การติดต่อระหว่างโปรเซส (Inter Process Communication).....	38
3.3.1 คลาส (Classes) ของการติดต่อระหว่างโปรเซสที่ใช้.....	38
3.3.2 Named Pipes UML static diagram.....	40
บทที่ 4 เทคนิคที่ใช้ในการเขียนโปรแกรม.....	41
4.1 การเขียนโปรแกรมแบบไม่ประสานเวลา (Asynchronous Programming).....	41
4.1.1 Asynchronous Client Application.....	43
4.1.2 Asynchronous Server Application.....	45
4.2 การสร้าง และการติดตั้งวินโดวส์เซอร์วิสแอปพลิเคชัน (Windows Service Application) โดย Visual Studio.NET.....	47
4.2.1 เขียนโปรแกรมวินโดวส์เซอร์วิส (Coding Windows Service).....	47
4.2.2 การเพิ่มคลาส Installer ในโปรแกรม.....	47

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้า
4.2.3 การใช้เครื่องมือรันเพื่อช่วยในการติดตั้งวินโดวส์เซอร์วิส	48
4.2.4 การรวมส่วนประกอบของตัวจับเวลา (Timer Component) กับวินโดวส์เซอร์วิส	48
4.2.5 การทดลองใช้โปรแกรม Personal Web Proxy เมื่อเป็นวินโดวส์เซอร์วิส	49
4.3 นิพจน์ปรกติ (Regular Expression)	54
4.3.1 หลักการพื้นฐาน	54
4.3.2 สัญลักษณ์ที่ใช้ในนิพจน์ปรกติ	55
4.3.3 ตัวอย่างการใช้งาน	56
4.4 การปรับแต่งค่าในรีจิสตรี (Registry Configuration)	57
4.4.1 หลักการที่เกี่ยวข้อง	57
4.4.2 การเลือก Subkey	58
4.4.3 การเปิดค่า Subkey ที่มีอยู่แล้ว	59
4.4.4 การเรียกค่าที่เกี่ยวข้องมาใช้	60
4.4.5 การปรับแต่งค่าของข้อมูล	60
4.4.6 การปิด Subkey	60
4.4.7 การลบ Subkey	61
4.4.8 การลบ Entire tree	61
4.4.9 การลบ Value	61
4.4.10 การรับค่า List of subkeys	61
4.4.11 การรับค่า List of value names	61
4.5 ส่วนต่อขยายบนไฟร์ฟอกซ์ (Firefox Extension)	62
4.5.1 โครงสร้างของไฟล์ Extension	62
4.5.2 ไฟล์ extension.rdf	62
4.5.3 ส่วนต่างๆที่ไฟล์ extension.rdf เป็นคุณสมบัติที่ต้องมี	63
4.5.4 ส่วนต่างๆที่ไฟล์ extension.rdf เป็นคุณสมบัติเพิ่มเติม	64
4.5.5 ตัวอย่างในการสร้างไฟล์ Extension	64
4.6 คลาสไดอะแกรมของโปรแกรมคัดกรองข้อมูล (Class Diagram)	69
4.6.1 คลาสที่เกี่ยวข้องกับWindowsService1	69
4.6.2 คลาสที่เกี่ยวข้องกับไฟล์ WindowsControl	72
บทที่ 5 การทดลอง และผลการทดลอง	76

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญ(ต่อ)

	หน้า
5.1 การทดสอบส่วนที่เป็นเนื้อหาของเว็บเพจ.....	76
5.1.1 ทำการศึกษาคำอธิบายที่จะต้องรู้ (Learning).....	76
5.1.2 นำกลุ่มทดลองมาทำการทดสอบ (Testing).....	77
5.2 การทดสอบภาพอนาจาร.....	80
5.2.1 การตรวจหาสีผิวมนุษย์.....	80
5.2.2 การตรวจสอบคุณสมบัติของภาพ.....	81
5.2.3 การจัดประเภทของภาพด้วยคุณสมบัติ.....	82
5.3 ประสิทธิภาพการทำงาน.....	87
บทที่ 6 บทวิจารณ์และสรุป.....	90
6.1 บทสรุป.....	90
6.2 ปัญหาอุปสรรคและแนวทางในการแก้ไข.....	91
6.3 แนวทางการพัฒนาต่อ.....	91
บรรณานุกรม.....	93
ภาคผนวก.....	94
ภาคผนวก ก. สถิติของเว็บไซต์ทั้งหมดที่นำมาทดสอบ.....	95
ภาคผนวก ข. คู่มือการติดตั้งซอฟต์แวร์.....	107
ภาคผนวก ค. คู่มือการใช้งานซอฟต์แวร์.....	112

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
2.1 คุณสมบัติ 6 ประการที่ได้จากภาพ.....	30
4.1 รายละเอียดของ subkey root.....	59
5.1 จำนวนเว็บไซต์ที่ลามกอนาจาร รวมทั้งต่างประเทศและในประเทศ.....	76
5.2 จำนวนเว็บไซต์ที่ไม่ลามกอนาจาร.....	77
5.3 จำนวนเว็บไซต์ที่เป็นเว็บไซต์ลามกอนาจารรวม.....	78
5.4 ผลการทดลองของกลุ่มตัวอย่าง.....	79
5.5 ผลการทดลองการตรวจสอบคุณสมบัติของภาพ.....	82
5.6 ประสิทธิภาพของการจัดกลุ่มข้อมูลของตัวอย่างภาพ.....	84
5.7 ประสิทธิภาพของการจัดกลุ่มข้อมูลของภาพนอกกลุ่มตัวอย่าง.....	86
5.8 เวลาที่ใช้ ในการทำงานกับเว็บไซต์ไม่อนาจารที่คัดกรองได้ถูกต้อง.....	87
5.9 เวลาที่ใช้ ในการทำงานกับเว็บไซต์ไม่อนาจารที่คัดกรองผิดพลาด.....	87
5.10 จำนวน Tag เหลือที่ได้จากเว็บไซต์ไม่อนาจาร 100 เว็บไซต์.....	88
5.11 เวลาที่ใช้ ในการทำงานกับเว็บไซต์อนาจารที่คัดกรองได้ถูกต้อง.....	88
5.12 เวลาที่ใช้ ในการทำงานกับเว็บไซต์อนาจารที่คัดกรองผิดพลาด.....	88
5.13 จำนวน Tag เหลือที่ได้จากเว็บไซต์อนาจาร 100 เว็บไซต์.....	89

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ

รูปที่	หน้า
2.1 โครงสร้างโดยรวมของโปรแกรมคัดกรองข้อมูลเว็บ.....	6
2.2 โครงสร้างอย่างละเอียดของโปรแกรมคัดกรองข้อมูลเว็บ.....	7
2.3 กลไกการทำงานของ K-means.....	16
2.4 โครงสร้างสี่อาร์จีบี เป็นลูกบาศก์.....	18
2.5 ความสัมพันธ์ระหว่างโครงสร้างสี YCbCr กับ RGB.....	19
2.6 กลุ่มตัวอย่างสีผิวในแกนของโครงสร้างสี YCbCr.....	20
2.7 กลุ่มตัวอย่างสีผิวในแกนของโครงสร้างสี HSV.....	21
2.8 กรอบพื้นที่สีผิว โครงสร้างสี YCbCr ที่ระนาบ $Y = 160$	22
2.9 กรอบพื้นที่สีผิว โครงสร้างสี HSV ที่ระนาบ $V = 70$	23
2.10(a) จุดของภาพที่มีขอบเขต 3×3 pixel (ค่า z เป็นค่าระดับสีเทา).....	26
2.10(b) (c) Roberts cross-gradient operators.....	26
2.10(d) (e) Sobel operators.....	26
2.11(a) ตัวอย่างภาพต้นฉบับที่ต้องการหาเส้นขอบ.....	27
2.11(b) ภาพที่ถูกตรวจหาเส้นขอบ (Sobel gradient).....	27
2.12 การทำงานของการตรวจสอบภาพอนาจาร.....	29
2.13 การทำงานของโปรแกรมการตรวจสอบภาพอนาจาร.....	31
2.14 การทำงานของการคำนวณหาไอแกนเวกเตอร์และ ไอแกนเวกซ์ของชุดข้อมูลตัวอย่างและ การหาจุดเซ็นทรอยด์ 3 จุด.....	32
2.15 การจำแนกภาพ.....	33
3.1 โครงสร้างของข้อมูลที่ส่งผ่าน โพรโตคอล HTTP.....	35
3.2 ตัวอย่างข้อความร้องขอด้วยเมธอด GET.....	36
3.3 ตัวอย่างข้อความร้องขอด้วยเมธอด POST.....	36
3.4 รูปแบบของ Personal Web Proxy.....	38
3.5 คลาสไคอะแกรมของ Inter Process Communication.....	40
4.1 แผนภาพขั้นตอนการทำงานของไคลเอ็นต์แอปพลิเคชัน.....	43
4.2 แผนภาพขั้นตอนการทำงานของเซิร์ฟเวอร์แอปพลิเคชัน.....	45
4.3 ส่วนของโค้ดของเมธอด เมื่อได้รับเหตุการณ์ครบรอบ interval ของตัวจับเวลา.....	48
4.4 ส่วนของโค้ดที่เพิ่มเพื่อสั่งให้โปรแกรม Personal Web Proxy ทำงาน.....	49
4.5 Command Prompt เมื่อเข้ามาที่ Visual Studio 2005 Command Prompt.....	50

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ(ต่อ)

รูปที่	หน้า
4.6 การใช้คำสั่งในการติดตั้งโปรแกรมวินโดวส์เซอร์วิส	51
4.7 ผลลัพธ์เมื่อทำการติดตั้งโปรแกรมวินโดวส์เซอร์วิสสำเร็จ	51
4.8 วินโดวส์เซอร์วิสที่ถูกติดตั้งเรียบร้อยแล้ว	52
4.9 ปุ่ม Start ใน Properties เพื่อเรียกให้วินโดวส์เซอร์วิสตัวนี้รัน	52
4.10 หน้าต่างควบคุมโปรแกรม เมื่อเริ่ม (Started) วินโดวส์เซอร์วิสชื่อ Service1	53
4.11 HTTP Message ที่ผ่านโปรแกรมเพื่อส่งคืนข้อมูลจากเว็บเซิร์ฟเวอร์ ให้กับเว็บเบราว์เซอร์ตามการร้องขอ	53
4.12 หน้าเว็บเบราว์เซอร์ที่ได้รับข้อมูลส่งกลับจากการร้องขอผ่านโปรแกรม	54
4.13 Registry Editor	58
4.14 การติดตั้ง Extension	68
4.15 Extension ที่ติดตั้งเสร็จเรียบร้อยแล้ว	68
4.16 Extension ที่ได้แสดงขึ้นมาในเมนู Add-ons และทำการกดAbout	68
4.17 คลาสชื่อ Class_image	69
4.18 คลาสชื่อ Classified	69
4.19 คลาสชื่อ PersonalWebProxy	70
4.20 สตักชื่อ PrivilegeDN	70
4.21 คลาสชื่อ Program	70
4.22 คลาสชื่อ ProjectInstaller	70
4.23 คลาสชื่อ Service1	70
4.24 คลาสชื่อ WebPage	71
4.25 สตักชื่อ WordPool	71
4.26 คลาสชื่อ WServiceNamePipe	71
4.27 คลาสชื่อ License_Agreement	72
4.28 คลาสชื่อ iniFileConfig	72
4.29 คลาสชื่อ CheckPassword	72
4.30 คลาสชื่อ WindowsControl	74
4.31 คลาสชื่อ WControlNamePipe	74
4.32 คลาสชื่อ ViewLog	75
5.1 กราฟของความสัมพันธ์ของคุณสมบัติของเว็บเพจ	78

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูปภาพ(ต่อ)

รูปที่	หน้า
5.2 สมการบนกราฟที่ใช้แบ่งแยกกลุ่มระหว่างเว็บอนาจาร และเว็บไม่อนาจาร	79
5.3(a) ภาพต้นฉบับ	81
5.3(b) ภาพเมื่อผ่านการทำกระบวนการตัดสีผิว	81
5.3(c) ภาพเมื่อผ่านการทำ Sobel Edge Oprator กับภาพต้นฉบับ	81
5.3(d) ภาพ เมื่อผ่านกระบวนการตัดสีผิวและผ่านการกรองด้วย Sobel Edge Oprator	81
5.3(e) ภาพเมื่อผ่านกระบวนการตัดเอาพื้นที่สีผิวที่มีขนาดเล็กออก	81
5.4 กราฟของความสัมพันธ์ของคุณสมบัติของภาพ	83
5.5 ผลการทดลองการจัดกลุ่มข้อมูลของตัวอย่างภาพ	84
5.6 ผลการแบ่งกลุ่มของภาพอนาจารนอกกลุ่มตัวอย่าง	85
5.7 ผลการแบ่งกลุ่มของภาพที่ไม่ใช่ภาพอนาจารนอกกลุ่มตัวอย่าง	86

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความสำคัญและที่มาของโครงการ

เนื่องจากในปัจจุบันได้มีเว็บไซต์ต่างๆ เกิดขึ้นมากมาย หลากหลายประเภทบนเครือข่ายอินเทอร์เน็ต เช่น เว็บไซต์ที่ให้ความรู้ ความบันเทิง และข่าวสารข้อมูล เหล่านี้ล้วนแต่เป็นสิ่งที่ดีทั้งสิ้น โดยจะมีเว็บไซต์อยู่จำนวนหนึ่งที่เป็นเว็บไซต์ผิดกฎหมาย หรือมีเนื้อหาที่ไม่เหมาะสมต่อเด็ก และเยาวชน ซึ่งเว็บเหล่านี้ทำให้เกิดปัญหาทางสังคมต่างๆ ตามมาได้ ในภายหลัง เช่น ขายแผ่นซีดีลามกผ่านทางเว็บไซต์ เว็บลามกอนาจาร ทำให้วัยรุ่นนอยากรู้อยากลอง นำไปปฏิบัติตาม หมกมุ่น หรืออาจผลเสียต่อสภาวะจิตใจ โดยที่พ่อแม่ของบุตรหลานไม่สามารถที่จะควบคุมได้

ถึงแม้ว่าปัจจุบันทางภาครัฐได้มีโครงการ Cyber Inspector เพื่อทำการตรวจสอบเว็บไซต์ ต่างๆ ที่มีอยู่ในปัจจุบันว่าเหมาะสมหรือไม่ โดยเว็บไซต์ไหนที่ไม่เหมาะสมก็จะถูก Block ไปซึ่งวิธีนี้เป็นวิธีการที่ค่อนข้างหนึ่ง แต่ก็อาจจะไม่ทั่วถึง เราจึงได้สร้าง โปรแกรมนี้ที่เรียกว่า “โปรแกรมคัดกรองข้อมูลเว็บ” (Web Content Filtering Application) นี้ขึ้นมาซึ่งเป็นโปรเจกต์ที่พัฒนาขึ้นมาใหม่ เพื่อให้บรรเทาภาระของภาครัฐและงบประมาณค่าใช้จ่ายที่จะต้องเสียไปอย่างมหาศาล และเพื่อให้ผู้ใช้ได้ท่องเว็บได้อย่างปลอดภัยในระดับหนึ่งด้วย

1.2 วัตถุประสงค์ของโครงการ

1. เพื่อสกัดกั้นเว็บไซต์ ที่มีเนื้อหาไม่เหมาะสม เช่น ข้อความ รูปภาพ เป็นอันตรายต่อเด็กและเยาวชน ไม่ว่าจะเป็นทางตรงหรือทางอ้อม
2. ผู้ใช้สามารถใช้งานเว็บไซต์ ได้อย่างปลอดภัยในระดับหนึ่งเพราะโปรแกรมสามารถดักจับสคริปต์ ของเว็บไซต์ ที่คาดว่าจะส่งผลอันตรายต่อข้อมูลของผู้ใช้และเครื่องคอมพิวเตอร์
3. ช่วยลดภาระของหน่วยงานรัฐบาล ในการป้องกันและปราบปรามปัญหาต่างๆ ที่เกี่ยวข้องกับเรื่องนี้ เช่น ลดภาระของ Cyber Inspector ลดปัญหาทางสังคม และอาชญากรรมต่างๆ
4. ลดภาระของผู้ปกครองในการดูแลบุตรหลานในการท่องเว็บไซต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.3 ขอบเขตของโครงการ

1. พัฒนาส่วนของการวิเคราะห์รูปภาพอนาจาร โดยใช้กระบวนการของการประมวลผลรูปภาพซึ่งสามารถวิเคราะห์ภาพที่มีช่วงสีในช่วงที่กำหนดเท่านั้นและภาพต้องอยู่ในสถานะแสงที่ค่อนข้างคงที่ ไม่มีการเปลี่ยนแปลงมากเกินไป
2. พัฒนาส่วนของการวิเคราะห์เนื้อหาบนเว็บไซต์ ซึ่งสามารถวิเคราะห์เนื้อหาได้บนเว็บไซต์ที่มีการใช้ภาษาอังกฤษเป็นหลัก
3. โปรแกรมสามารถคัดกรองหน้าเว็บเพจที่มีเนื้อหาไม่เหมาะสม หรือลามกอนาจาร (ในรูปแบบข้อความ รูปภาพ ลิงค์ เป็นต้น) ก่อนที่จะแสดงต่อผู้ใช้
4. โปรแกรมสามารถดักจับสคริปต์ของเว็บไซต์ ที่คาดว่าจะส่งผลกระทบต่อข้อมูลของผู้ใช้และเครื่องคอมพิวเตอร์
5. โปรแกรมสามารถทำงานร่วมกับ Fire Fox Web Browser นอกจากนี้ยังสามารถใช้งานร่วมกับ Web Browser อื่นได้ เช่น Internet Explorer เป็นต้น
6. ไม่สนับสนุนการใช้งานกับร่วมกับโปรโตคอล https และ ftp

1.4 วิธีการดำเนินการ

1. ศึกษาเนื้อหารายละเอียดต่างๆ ที่จำเป็นต้องใช้
 - ศึกษาหาข้อมูลและโครงสร้างของ Personal Web Proxy
 - ศึกษาวิธีการเขียน โปรแกรมเพื่อสร้าง Personal Web Proxy
 - ศึกษาโครงสร้างและการเขียนโปรแกรมแบบ Asynchronous Programming
 - ศึกษาโครงสร้างและการเขียนโปรแกรมเป็น Windows Service Application
 - ศึกษาปัจจัยที่จะแสดงให้เห็นว่าในหน้าเว็บ 1 หน้านั้นมีอะไรบ้างที่จะบ่งชี้ได้ว่าเว็บเพจนั้นเป็นเว็บไซต์ที่ไม่เหมาะสม
 - ศึกษาอัลกอริทึมที่จะนำมาพิจารณาในส่วนของ Text Detection ในที่นี้จะทำการศึกษาวิธีการทางสถิติที่เรียกว่า Eigen
 - ศึกษาการเขียนโปรแกรมเพื่อให้สามารถนำ link ในหน้านั้นมาได้จากสารบัญเว็บของ google และทำการเขียนโปรแกรมเพื่อให้ได้มาซึ่งโค้ดของหน้าเว็บไซต์นั้นๆ
 - ศึกษาการคัดกรองในส่วนของรูปภาพ (อัลกอริทึมทางด้าน Image Processing) เพื่อทำการพิจารณาในส่วนของรูปภาพ
 - ศึกษาและพัฒนากลไกการติดตั้งและบังคับใช้ ให้ Web Browser จำเป็นต้องเชื่อมต่อผ่าน Personal Web Proxy เสมอ
2. การทดลองเพื่อหาค่าทางสถิติ หรือ โคลิที่สามารรถนำมารวมเป็น โปรเจคได้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำสถิติต่างๆ โดยดำเนินการด้วย วิธีการทางสถิติที่เรียกว่า Eigen เพื่อหาค่า หรือ สมการที่จะนำไปใช้เป็นมาตรฐานสำหรับวิเคราะห์เนื้อหาบนเว็บต่อไป
 - ทำการศึกษาและทดลองเขียนโปรแกรมในส่วนของ Personal Web Proxy ให้ทำงานเป็น windows service
 - ทำการศึกษาและเขียน โปรแกรมแบบ Inter Process Communication
3. ออกแบบในส่วนของ User Interface
 4. นำแต่ละฟังก์ชันการทำงานมารวมกัน จากนั้นทำการทดสอบการใช้งาน และแก้ไขข้อผิดพลาดของโปรแกรมในส่วนต่างๆ

1.5 ประโยชน์ที่คาดว่าจะได้รับ

1. ได้รับความรู้ความเข้าใจเกี่ยวกับเอชทีทีพี โพรโตคอล (HTTP: HyperText Transfer Protocol)
2. ได้รับความรู้ความเข้าใจในการเขียน โปรแกรมเกี่ยวกับ Network Programming, Windows Service และ Inter Process Communication
3. ได้รับความรู้ความเข้าใจเกี่ยวกับกระบวนการของอัลกอริทึม ที่ใช้ ได้แก่ ไอแกนเวกเตอร์-ไอแกนแวลูส์ (Eigen Vector-Eigen Value) และหลักการวิเคราะห์แยกแยะส่วนประกอบ (Principal Components Analysis: PCA)

1.6 ส่วนประกอบของปริญญานิพนธ์

ปริญญานิพนธ์ฉบับนี้ได้แบ่งเนื้อหาออกเป็น 6 บท 3 ภาคผนวกด้วยกัน ดังนี้

บทที่ 1 กล่าวถึงความสำคัญและที่มาของโครงการ วัตถุประสงค์ของโครงการ ขอบเขตของโครงการ วิธีการดำเนินการ ประโยชน์ที่คาดว่าจะได้รับ และส่วนประกอบของปริญญานิพนธ์

บทที่ 2 กล่าวถึงโครงสร้างของโปรแกรมคัดกรองข้อมูลเว็บ โดยบอกถึงส่วนประกอบสำคัญที่ได้พัฒนาขึ้นเอง อัลกอริทึมที่ใช้ในการคัดกรองข้อมูลทั้งรูปภาพ และข้อความ

บทที่ 3 แสดงหลักการพื้นฐานบนเครือข่ายที่เกี่ยวข้องกับการทำงานของโปรแกรม โดยประกอบไปด้วย โพรโตคอลเอชทีทีพี, พร็อกซีส่วนตัว และการติดต่อส่งข้อมูลกันระหว่างโปรเซส

บทที่ 4 บอกถึงเทคนิคที่ใช้ในการเขียน โปรแกรมคัดกรองข้อมูลเว็บ โดยมีเนื้อหา รายละเอียดที่บอกถึงความหมาย เหตุผลที่ใช้ รวมไปถึงขั้นตอนการเขียน พร้อมติดตั้ง ซึ่งเทคนิคที่ใช้ประกอบไปด้วย การเขียน โปรแกรมแบบไม่ประสานเวลา, การสร้าง และการติดตั้งวินโดวส์เซอร์วิสแอปพลิเคชัน โดยใช้ Visual Studio .NET, รูปแบบในการเขียนนิพจน์ปรกติ, การปรับแต่งค่าในรีจิสตรี และส่วนที่ได้พัฒนาเพิ่มบนไฟร์ฟอกซ์ หรือส่วนต่อขยายบนไฟร์ฟอกซ์นั่นเอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5 กล่าวถึงการทดลองและผลการทดลอง โดยได้เขียนการทดลอง และผลการทดลอง เกี่ยวกับส่วนของเนื้อหาบนเว็บเพจ ว่าสามารถนำมาใช้จริงได้หรือไม่ รวมไปถึงการทดสอบภาพ อนุกรมตามสถานะของสีต่างๆ และหาค่าประสิทธิภาพการทำงานของโปรแกรม ในการทำงานจริง ร่วมกับเว็บเบราว์เซอร์ เพื่อเยี่ยมชมเว็บเพจต่างๆบนเครือข่ายอินเทอร์เน็ต

บทที่ 6 เป็นบทวิจารณ์และสรุป ซึ่งกล่าวถึงบทสรุปของโครงการ วิจารณ์สิ่งที่ได้รับจากโครงการ และข้อเสนอแนะสำหรับเป็นแนวทางในการพัฒนาต่อ

ภาคผนวก ก. กล่าวถึงสถิติของการเยี่ยมชมเว็บไซต์ต่างๆ โดยผ่านโปรแกรมคัดกรองข้อมูล ที่ได้จากการทดลองอย่างละเอียด ได้แก่ เวลาที่ใช้ จำนวน Tag ที่พบบนเว็บเพจนั้นๆ ทั้งเว็บไซต์ ลามกอนาจาร และไม่ลามกอนาจาร

ภาคผนวก ข. เป็นคู่มือการติดตั้งโปรแกรมคัดกรองข้อมูลเว็บบนเครื่องผู้ใช่ว่ามีขั้นตอนการติดตั้งอย่างไรบ้างอย่างละเอียด

ภาคผนวก ค. เป็นคู่มือของการใช้งาน โปรแกรมคัดกรองข้อมูลเว็บ อธิบายขั้นตอนการปรับแต่งค่าต่างๆ บนโปรแกรม ขั้นตอนการใช้งาน พร้อมรูปภาพแสดงประกอบ

บทที่ 2

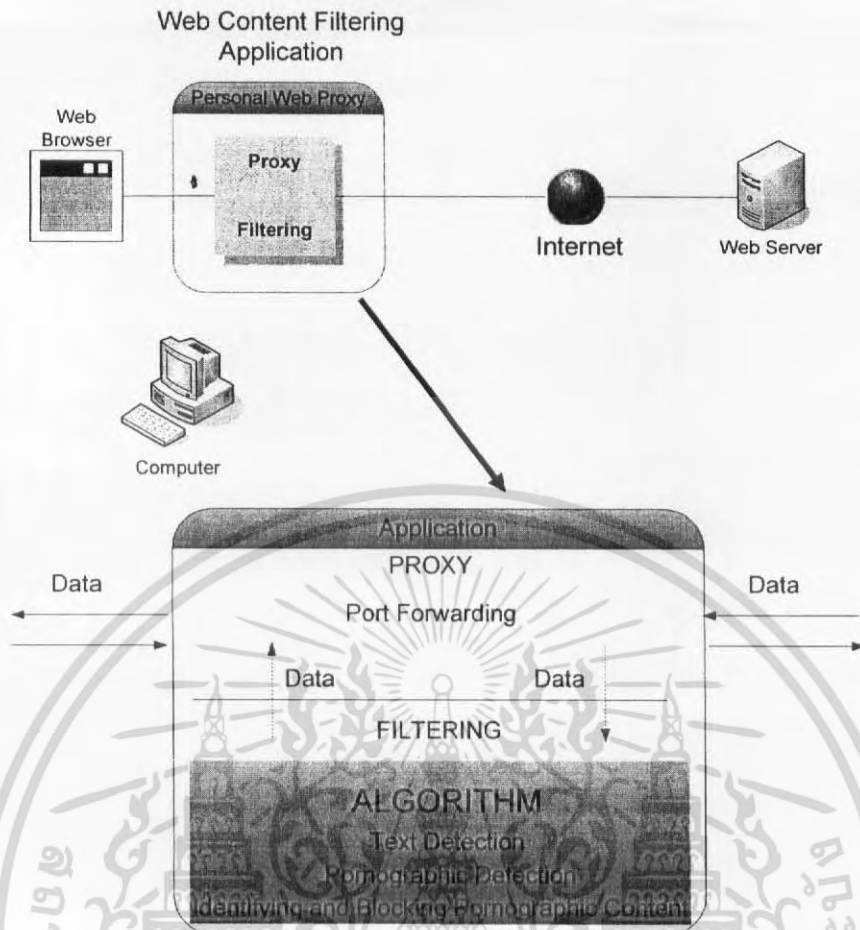
โครงสร้างของซอฟต์แวร์ และการคัดกรองข้อมูล

(Structural design and Detection)

2.1 โครงสร้างซอฟต์แวร์

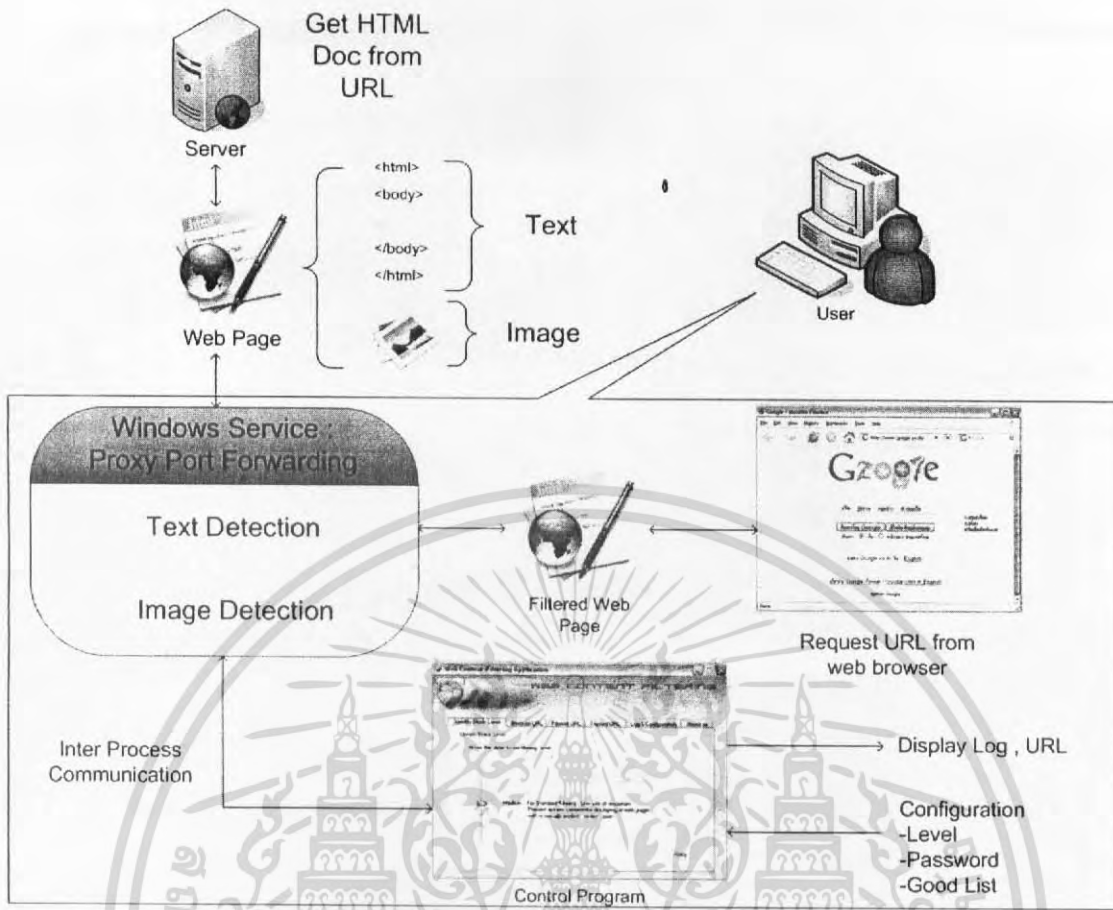
โปรแกรมคัดกรองข้อมูลเว็บ (Web Content Filtering Application) ทำหน้าที่คัดกรองข้อมูลที่ได้รับการร้องขอข้อมูล บนเว็บเบราว์เซอร์ (Web Browser) ทางฝั่งผู้ใช้ไคลเอนต์ (Client) โดยใช้เทคนิคที่เรียกว่า Personal Web Proxy คือการทำหน้าที่เป็นพร็อกซี่ บนเครื่องคอมพิวเตอร์ทางฝั่งผู้ใช้เว็บเบราว์เซอร์ เพื่อกรองเนื้อหาข้อมูลบนเว็บที่ได้จากเว็บเซิร์ฟเวอร์ (Web Server) ก่อนที่จะแสดงต่อผู้ใช้นบนเว็บเบราว์เซอร์ ซึ่งเมื่อผ่านการกรองนี้ที่โปรแกรมแล้วทำให้ข้อมูลที่ผ่านไปแสดงผลมีความปลอดภัยมากขึ้นระดับหนึ่ง และมีเนื้อหาที่เหมาะสมกับของผู้ใช้ โดยเฉพาะอย่างยิ่งเด็กและเยาวชน

โดยในการทำงาน โปรแกรมคัดกรองข้อมูลจะทำตัวเป็นวินโดว์เซอร์วิส (Windows Service) ซึ่งต่างจากการทำตัวเป็นวินโดว์แอปพลิเคชัน (Windows Application) คือ วินโดว์เซอร์วิสสามารถเริ่มการทำงานของโปรแกรมก่อนที่ผู้ใช้ (User) คนใดคนหนึ่งจะเข้าสู่ระบบ (หากมีการปรับตั้งค่าให้เริ่มโปรแกรมเมื่อวินโดว์ทำการ Boot Up) และประโยชน์อีกอย่างหนึ่ง คือ สามารถรันโปรแกรมได้ตลอดเวลา ครอบคลุมที่ระบบปฏิบัติการยังทำงานอยู่ โดยผู้ใช้ทั่วไปที่ไม่ใช่ admin จะไม่สามารถทำการปิดเซอร์วิส หรือโปรเซสของโปรแกรมนี้ได้ โดยมีโครงสร้างโดยรวมดังรูปที่ 2.1 และโครงสร้างโดยละเอียดดังรูปที่ 2.2



รูปที่ 2.1 โครงสร้างโดยรวมของโปรแกรมคัดกรองข้อมูลเว็บ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



Text Detection

Detect parameter from tag
 ex. `` `<meta>` `<a href>` → Use Eigen for Dimension Reduction

Pornographic content?
 Yes – Blocked URL ← Consider by Algorithm
 No – Passed URL

Image Detection

Get pixel
 (255,0,3)(150,43,200)(49,169,32)...

→ Skin Color Model (YCbCr)(RGB)

↓ Edge Detection & Filter

↓ Cal Weight

Pornographic image?
 Yes – No Display
 No – Display

รูปที่ 2.2 โครงสร้างอย่างละเอียดของ โปรแกรมคัดกรองข้อมูลเว็บ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- Personal Web Proxy Application ประกอบด้วยส่วนของ Port Forwarding เพื่อทำหน้าที่เป็นตัวกลางในการส่งผ่านข้อมูลระหว่างทางฝั่งผู้ใช้ กับทางฝั่งเว็บเซิร์ฟเวอร์ และส่วนของการทำ Filtering ทำหน้าที่ตรวจสอบข้อมูลเนื้อหาบนหน้าเว็บไซต์ก่อนที่จะส่งต่อไปให้ผู้ใช้จากการร้องขอ โดยใช้ Text Detection , Pomographic Detection และ Identifying and Blocking Pomographic Content ในการพิจารณาว่าเว็บไซต์นั้นๆ มีเนื้อหาที่เหมาะสมหรือไม่ ก่อนที่จะส่งผ่านต่อไปที่ทางฝั่งผู้ใช้
- สร้างพรีอ็อกซี่ส่วนตัว เพื่อรับการร้องขอข้อมูลเว็บจากทางฝั่งผู้ใช้ จากนั้นทำการสร้างการร้องขอข้อมูลไปทางฝั่งเว็บเซิร์ฟเวอร์ พร้อมทั้งสร้างกระบวนการในการรอรับข้อมูล เพื่อส่งกลับทางฝั่งผู้ใช้
- นำอัลกอริทึมตรวจจับภาพที่ไม่เหมาะสมมาคำนวณ เพื่อทำการคัดกรองส่วนรูปที่ไม่เหมาะสม และนำฟังก์ชันการตรวจหาเนื้อหาเช่น การดักจับข้อมูล เนื้อหา หรือสคริปต์ที่ไม่เหมาะสม มาคำนวณหาเปอร์เซ็นต์ความไม่เหมาะสม เพื่อจะกรองข้อมูลในเว็บไซต์นั้นๆ
- จะมีส่วนของอินเตอร์เฟซซึ่งจะมีการเรียกแก้ไข และเพิ่มรีจิสตรี สามารถเรียก Log File ของโปรแกรมได้ สามารถปรับแต่งระดับการกรองข้อมูลได้
- มีส่วนของ Inter Process Communication เพื่อทำการติดต่อส่งข้อมูลกันระหว่างอินเตอร์เฟซของโปรแกรมและ วินโดว์เซอร์วิส
- วินโดว์เซอร์วิส จะเป็นทำหน้าที่ในส่วนของพรีอ็อกซี่ส่วนตัว
- สามารถพัฒนาต่อ ให้เพิ่มการดักจับเนื้อหา หรือปรับเปลี่ยนอัลกอริทึมที่ใช้ในการตรวจสอบความเหมาะสมของรูปภาพ เพื่อให้เพิ่มประสิทธิภาพมากยิ่งขึ้น

2.2 ส่วนประกอบที่สำคัญที่ได้พัฒนาในโปรแกรมคัดกรองข้อมูลเว็บ

ส่วนที่ได้พัฒนาขึ้นเอง

1. วินโดว์เซอร์วิส จะเป็นแกนหลักของการทำงานในส่วนของพรีอ็อกซี่ส่วนตัว
2. โปรแกรมพรีอ็อกซี่ส่วนตัว ประกอบไปด้วยการทำงานโดยเทคนิคต่างๆ ดังนี้
 - Port Forwarding เพื่อทำหน้าที่เป็นตัวกลางในการส่งผ่านข้อมูลระหว่างทางฝั่งผู้ใช้ กับทางฝั่งเว็บเซิร์ฟเวอร์
 - Asynchronous Programming ทำให้การเชื่อมต่อจะไม่ถูกปิดกั้นโดยตัวแอปพลิเคชันระหว่าง ที่รอคอยกระบวนการทำงานของระบบเครือข่ายเสร็จสมบูรณ์
 - Socket Programming เป็นการเขียนโปรแกรมเพื่อให้สามารถติดต่อสื่อสารกันบนเน็ตเวิร์คได้
3. ส่วนของ Text Detection ประกอบไปด้วยฟังก์ชันต่างๆ ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ฟังก์ชันที่ทำการค้นหาคุณสมบัติต่างๆ บนหน้าเว็บเพจ
 - ฟังก์ชันการคำนวณหาค่าน้ำหนักของหน้าเว็บเพจนั้นๆ
4. ส่วนของ Image Processing: Pornographic Detection ประกอบไปด้วยฟังก์ชันต่างๆ ดังนี้
- ฟังก์ชันที่ทำการค้นหาพิกเซลและพิจารณาว่าพิกเซลนั้นเป็นพิกเซลของสีผิวหรือไม่
 - ฟังก์ชันที่ทำการค้นหาขอบของภาพ
 - ฟังก์ชันที่ทำการคำนวณหาค่าน้ำหนักของรูปภาพ
5. ส่วนของ Principal component analysis (PCA) จะประกอบไปด้วยฟังก์ชันต่างๆ ดังนี้
- ฟังก์ชันวิเคราะห์ถึงคุณสมบัติและทำการเลือกคุณสมบัติที่สำคัญที่สุด
 - ฟังก์ชันที่ใช้ค้นหา หรือระบุถึงรูปแบบเฉพาะตัวบางส่วนของข้อมูลต่างๆ ที่ประกอบไปด้วยหลายๆ ปัจจัย ว่ามีความเหมือนหรือต่างกันอย่างไร
6. ส่วนของอินเตอร์เฟซ (User Interface) ประกอบด้วย
- การปรับแต่งค่าพรีอ็อกซ์ริจิสตรี และ รีจิสตรี ของโปรแกรมคัดกรองข้อมูลเว็บ
 - การปรับเปลี่ยนค่า prefs.js ที่อยู่ใน path ของแต่ละUser เพื่อให้ ไฟร์ฟอกซ์สามารถปรับเปลี่ยนค่าของ พรีอ็อกซ์ได้
 - การเขียนและเรียก Log File เพื่อทำการบันทึกการใช้งานเว็บเพจ ประกอบไปด้วย URL เวลาที่ใช้งาน
 - ควบคุมการทำงานของ วินโดว์เซอร์วิส

ส่วนที่ได้นำ Source Code มาพัฒนาต่อ

1. Inter Process Communication ได้นำ คลาสพื้นฐานที่อยู่ใน <http://www.codeproject.com> และได้นำมาประยุกต์ใช้กับโปรแกรมเพื่อติดต่อระหว่างในส่วนของ อินเตอร์เฟซ (User Interface) กับ Windows Service

2.3 อัลกอริทึมพื้นฐานที่ใช้ในโปรแกรม

ในการทดลองจะต้องทำการแบ่งกลุ่มชุดข้อมูลตัวอย่างเพื่อใช้เป็นมาตรฐานในการพิจารณาข้อมูลใหม่กว่าเป็นข้อมูลในกลุ่มใด โดยจะทำการแบ่งกลุ่มข้อมูลเป็นสองกลุ่ม คือ กลุ่มของภาพที่ไม่ใช่ภาพอนาจารและกลุ่มของข้อมูลที่เป็นภาพอนาจาร โดยจะใช้หลักการของการวิเคราะห์ส่วนประกอบ, K-Means และ K-Nearest Neighbor (KNN)

หลักการการวิเคราะห์ส่วนประกอบเป็นเทคนิคทางสถิติที่มีประโยชน์ในการทำงานประเภทต่างๆ อาทิเช่น การจดจำใบหน้า(Face recognition) และการบีบอัดภาพ (Image Compression) และเป็นเทคนิคพื้นฐานในการค้นหารูปแบบในข้อมูลที่มีหลายมิติ ก่อนที่จะอธิบายเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ถึงรายละเอียดของหลักการการวิเคราะห์ส่วนประกอบ จะกล่าวถึงทฤษฎีทางคณิตศาสตร์ที่ใช้ในหลักการการวิเคราะห์ส่วนประกอบ ประกอบด้วย สถิติ(statistics) ซึ่งพิจารณาการกระจายตัวของข้อมูล และเมทริกซ์อัลจิบรา(Matrix Algebra) พิจารณาไอแกนเวกเตอร์(Eigenvectors) และไอแกนแวลูส์(Eigenvalues) ซึ่งเป็นคุณสมบัติสำคัญของ เมทริกซ์ ที่ซึ่งเป็นรากฐานของหลักการการวิเคราะห์ส่วนประกอบ

2.3.1 สถิติ

สถิติสามารถนำมาใช้เมื่อมีชุดข้อมูลขนาดใหญ่และต้องการวิเคราะห์ความสัมพันธ์ระหว่างแต่ละจุดของชุดข้อมูลนั้น ในที่นี้จะกล่าวถึงการวัดที่กระทำกับกลุ่มของข้อมูล เช่น ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation), วาเรียนซ์ (Variance), โควาเรียน (Covariance) และ โควาเรียนซ์เมทริกซ์ (Covariance Matrix)

- ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation)

ค่าเบี่ยงเบนมาตรฐานของชุดข้อมูลเป็นการวัดการกระจายตัวของข้อมูล คำนิยามของค่าเบี่ยงเบนมาตรฐานคือระยะทางเฉลี่ยจากค่ากลางของชุดข้อมูลถึงจุดข้อมูลหนึ่งๆ การคำนวณค่าเบี่ยงเบนมาตรฐานมีสมการดังนี้

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}} \quad (2.1)$$

การคำนวณหาค่าเบี่ยงเบนมาตรฐานสามารถหาได้จากกำลังสองของระยะทางจากแต่ละจุดของข้อมูล ไปยังค่าเฉลี่ยของกลุ่มข้อมูล ทำการรวมค่าที่คำนวณได้ทั้งหมดหารด้วย จำนวนสมาชิกของกลุ่มข้อมูล (n) ลบด้วย 1 แล้วทำการหาค่ารากที่สอง

- วาเรียนซ์ (Variance)

วาเรียนซ์เป็นตัววัดการกระจายตัวของข้อมูลอีกแบบหนึ่ง มีลักษณะแบบเดียวกับ ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) ซึ่งมีสมการดังนี้

$$s^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{(n-1)} \quad (2.2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งอาจกล่าวได้ว่าวาเรียนซ์เป็นกำลังสองของค่าเบี่ยงเบนมาตรฐาน (s^2) โดยทั้งค่าเบี่ยงเบนมาตรฐานและวาเรียนซ์ต่างก็เป็นการวัดการกระจายตัวของข้อมูล แต่ค่าเบี่ยงเบนมาตรฐานจะนิยมใช้มากกว่า

- โควาเรียนซ์ (Covariance)

จากที่กล่าวมาค่าเบี่ยงเบนมาตรฐานและวาเรียนซ์เป็นการวัดการกระจายตัวของข้อมูล 1 มิติ อย่างไรก็ตามมีชุดข้อมูลจำนวนมากที่มีมากกว่า 1 มิติ และจุดประสงค์ของการวิเคราะห์ข้อมูลทางสถิติเพื่อหาว่าข้อมูลมีความสัมพันธ์ระหว่างมิติหรือไม่ โดยปกติแล้ว โควาเรียนจะใช้วัดข้อมูล 2 มิติ ถ้าทำการคำนวณ โควาเรียนซ์ระหว่างข้อมูล 1 มิติและตัวมันเองจะได้ค่าวาเรียนซ์ แต่ถ้าข้อมูลมี 3 มิติ (x, y, z) สามารถคำนวณค่า โควาเรียนซ์ได้โดยคำนวณ โควาเรียนซ์ระหว่าง x กับ y , x กับ z และ y กับ z การคำนวณค่า โควาเรียนซ์ระหว่าง x กับ x , y กับ y และ z กับ z จะได้ค่าวาเรียนซ์ของ x , y และ z

สมการของ โควาเรียนซ์คล้ายกับสมการของวาเรียนซ์ ซึ่งสมการของวาเรียนซ์เป็นดังนี้

$$\text{var}(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})(X_i - \bar{X})}{(n-1)} \quad (2.3)$$

และสมการของ โควาเรียนซ์เป็นดังนี้

$$\text{cov}(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(n-1)} \quad (2.4)$$

ถ้าค่า โควาเรียนซ์ที่คำนวณออกมามีค่าเป็นบวกแสดงว่าค่าข้อมูลมีการเพิ่มขึ้นทั้งสองมิติ แต่ถ้าค่า โควาเรียนซ์เป็นลบแสดงว่าข้อมูลค่าเพิ่มขึ้นหนึ่งมิติ ส่วนอีกมิติมีค่าลดลง แต่ถ้าค่าเป็นศูนย์แสดงว่าข้อมูลทั้งสองมิติไม่ได้ขึ้นต่อกัน โดยค่าของ $\text{cov}(X, Y)$ จะมีค่าเท่ากับ $\text{cov}(Y, X)$ พิจารณาจากสมการ ที่ 2.4

- โควาเรียนซ์เมตริกซ์ (Covariance Matrix)

จากที่ได้กล่าวมาข้างต้นโดย โควาเรียนซ์จะคำนวณข้อมูล 2 มิติ ถ้ามีข้อมูลมากกว่า 2 มิติ แสดงว่ามีการคำนวณ โควาเรียนซ์มากกว่าหนึ่งตัว เช่น ชุดข้อมูลที่มี 3 มิติ (x, y, z) สามารถคำนวณ $\text{cov}(x, y)$, $\text{cov}(x, z)$ และ $\text{cov}(y, z)$ ถ้ามีข้อมูล n มิติ จะสามารถคำนวณค่า โควาเรียนซ์ได้แตกต่างกัน

$$\frac{n!}{(n-2)! \cdot 2} \text{ ค่า}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ทางที่ง่ายสำหรับการคำนวณค่าโควาเรียนซ์ระหว่างมิติหลายๆมิติ ทำได้โดยการนำค่าโควาเรียนซ์ทั้งหมดใส่ลงในเมตริกซ์ ซึ่งนิยามสำหรับโควาเรียนซ์เมตริกซ์สำหรับชุดข้อมูลที่มี n มิติเป็นดังนี้

$$C^{n \times n} = (c_{i,j}, c_{i,j} = \text{cov}(Dim_i, Dim_j)) \quad (2.5)$$

ซึ่ง $C^{n \times n}$ คือเมตริกซ์ที่มี n แถว n คอลัมน์ และ Dim_x เป็นมิติที่ x ถ้ามีชุดข้อมูล n มิติ ดังนั้นเมตริกซ์จะมี n แถว n คอลัมน์ และข้อมูลในเมตริกซ์คือผลลัพธ์ของการคำนวณโควาเรียนซ์ระหว่างมิติ 2 มิติที่ต่างกัน เช่น ข้อมูลที่อยู่ในแถว 2 คอลัมน์ 3 คือการคำนวณโควาเรียนซ์ระหว่างมิติที่ 2 และมิติที่ 3

ตัวอย่างเช่น ข้อมูลมี 3 มิติ (x,y,z) ดังนั้นโควาเรียนซ์เมตริกซ์จะมี 3 แถวและ 3 คอลัมน์ดังนี้

$$C = \begin{pmatrix} \text{cov}(x,x) & \text{cov}(x,y) & \text{cov}(x,z) \\ \text{cov}(y,x) & \text{cov}(y,y) & \text{cov}(y,z) \\ \text{cov}(z,x) & \text{cov}(z,y) & \text{cov}(z,z) \end{pmatrix}$$

พิจารณาจากโควาเรียนซ์เมตริกซ์ จะเห็นว่าตามแนวเส้นทะแยงมุมบนซ้ายไปยังล่างขวา เป็นค่าโควาเรียนซ์ระหว่างมิตินั้นและตัวมันเอง หรือก็คือวาเรียนซ์ของมิตินั้น และจาก $\text{cov}(a,b)$ มีค่าเท่ากับ $\text{cov}(b,a)$ จะเห็นได้ว่าเมตริกซ์นี้ สมมาตรบนเส้นทะแยงมุมนี้

2.3.2 เมตริกซ์อัลจิบรา (Matrix Algebra)

ในหัวข้อนี้จะเป็นพื้นฐานของเมตริกซ์อัลจิบราที่ถูกใช้ในหลักการการวิเคราะห์ส่วนประกอบ ซึ่งไอแกนเวกเตอร์และไอแกนแวลูส์ที่จะกล่าวต่อไปมาจากเมตริกซ์อัลจิบรา ตัวอย่างที่ 1 เป็นเมตริกซ์อัลจิบราซึ่งมีนอนไอแกนเวกเตอร์ (non-eigenvector)

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} x \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 11 \\ 5 \end{pmatrix}$$

ตัวอย่างที่ 2 เป็นเมตริกซ์อัลจิบราซึ่งมีไอแกนเวกเตอร์

$$\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} x \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 8 \end{pmatrix} = 4x \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

- ไอแกนเวกเตอร์ (Eigenvectors)

จากคุณสมบัติของเมตริกซ์ซึ่งสามารถคูณกันได้ระหว่าง 2 เมตริกซ์ ไอแกนเวกเตอร์นั้นเป็นกรณีพิเศษรูปแบบหนึ่ง โดยพิจารณาจากการคูณกันระหว่างเมตริกซ์และเวกเตอร์ในตัวอย่างของหัวข้อเมตริกซ์อัลจิบรา ในตัวอย่างแรกเวกเตอร์ผลลัพธ์ที่ไม่ได้เป็นตัวเลขจำนวนเต็มคูณกับเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เวกเตอร์ตั้งต้น ต่างกับตัวอย่างที่สองซึ่งผลลัพธ์เป็นสี่เท่าของเวกเตอร์ตั้งต้น จากตัวอย่างจะพบว่าเวกเตอร์ต้นฉบับประกอบด้วย 2 มิติ โดยเวกเตอร์ $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ จากตัวอย่างที่ 2 แทนเวกเตอร์ที่มีจุดกำเนิดที่ (0,0) ไปยังจุด (3,2)

ไอแกนเวกเตอร์สามารถหาได้จากเมทริกซ์ที่จำนวนคอลัมน์และจำนวนแถวเท่ากันเท่านั้น แต่ไม่ใช่ทุกเมทริกซ์เช่นนั้นจะมีไอแกนเวกเตอร์ ถ้ากำหนดให้เมทริกซ์ $n \times n$ มีไอแกนเวกเตอร์แล้ว จะมี n ไอแกนเวกเตอร์

การหาไอแกนเวกเตอร์ โดยจะทราบว่าคุณค่าของเวกเตอร์ไม่มีผลต่อการดูว่าเวกเตอร์นั้นเป็นไอแกนเวกเตอร์หรือไม่ แต่ทิศทางมีผล ดังนั้นการหาไอแกนเวกเตอร์มักจะมีกรทำให้ความยาวของเวกเตอร์เป็นหนึ่ง ซึ่งจะทำให้ทุกไอแกนเวกเตอร์มีความยาวเท่ากัน จากตัวอย่างที่ 2 ไอแกนเวกเตอร์คือ $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ และมีความยาวของเวกเตอร์เท่ากับ $\sqrt{3^2 + 2^2} = \sqrt{13}$ เมื่อได้ความยาวของเวกเตอร์แล้วจะนำไปหารเวกเตอร์ตั้งต้น ซึ่งจะให้ได้เวกเตอร์ที่มีความยาวเท่ากับหนึ่ง

- ไอแกนแวลูส์ (Eigenvalues)

ไอแกนแวลูส์มีความสัมพันธ์ใกล้ชิดกับ ไอแกนเวกเตอร์จากตัวอย่างที่ 2 ค่าจำนวนเต็ม (4) เรียกว่า ไอแกนแวลูส์ซึ่งสัมพันธ์กับ ไอแกนเวกเตอร์ ดังนั้นสามารถพบ ไอแกนแวลูส์ และ ไอแกนเวกเตอร์มาเป็นคู่ นั่นหมายถึงเมื่อหาไอแกนเวกเตอร์ได้ก็จะได้ไอแกนแวลูส์ด้วย

2.3.3 หลักการวิเคราะห์แยกแยะส่วนประกอบ (Principal Components Analysis: PCA)

หลักการวิเคราะห์แยกแยะส่วนประกอบ (PCA) คือ วิธีการบ่งชี้ถึงรูปแบบของข้อมูล และช่วยแสดงข้อมูลให้เห็นถึงจุดเด่นของความเหมือนและความต่างของข้อมูล อย่างไรก็ตามรูปแบบของข้อมูลนั้นยากที่จะหาได้ เมื่อข้อมูลนั้นมีหลายมิติ (Dimension) และยิ่งยากที่จะแสดงเป็นรูปภาพหรือกราฟได้ ดังนั้น PCA จึงเป็นเครื่องมือที่มีประสิทธิภาพอย่างมากในการแยกแยะหรือจำแนกข้อมูล

นอกจากนี้ PCA ยังมีประโยชน์ในการช่วยกำหนดรูปแบบของข้อมูล และบีบอัดข้อมูลได้ โดยการลดจำนวนมิติ ปราศจากการสูญเสียเนื้อหาของข้อมูล โดยการวิเคราะห์แยกแยะส่วนประกอบการใช้งานมีวิธีการทำดังนี้

ขั้นตอนที่ 1 รวบรวมข้อมูล

- รวบรวมข้อมูล โดยพิจารณาถึงมิติ ของปัจจัย

ขั้นตอนที่ 2 หักลบออกจากค่ากลาง (Mean)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- เราจะต้องห้กลับค่ากลางออกจากแต่ละข้อมูลของมิติ ค่ากลางที่จะทำการลบนี้เป็นค่าเฉลี่ยตามขวางของแต่ละมิติ ดังนั้น ค่า x จะต้องลบด้วย \bar{x} และค่า y จะต้องลบด้วย \bar{y} ทำให้ผลลัพธ์ของชุดข้อมูลที่มีค่าเท่ากับค่ากลาง จะเท่ากับ 0 (ข้อมูลที่ถูกลบออกจากค่ากลางแล้ว จะเรียกว่า DataAdjust)

ขั้นตอนที่ 3 การคำนวณค่าเมทริกซ์โควาเรียนซ์ (Covariance Matrix)

- สมมติว่าหากข้อมูลเป็น 2 มิติ ทำให้เมทริกซ์โควาเรียนซ์เป็น 2x2

ขั้นตอนที่ 4 การคำนวณไอแกนเวกเตอร์ (Eigenvectors) และไอแกนแวลูส์ (Eigenvalues) ของค่าเมทริกซ์โควาเรียนซ์

- เมื่อค่าเมทริกซ์โควาเรียนซ์เป็น สี่เหลี่ยมจัตุรัสแล้ว(จำนวน Row เท่ากับจำนวนของ Column)เราจะสามารถคำนวณค่าของไอแกนเวกเตอร์ และไอแกนแวลูส์ สำหรับเมทริกซ์ได้ โดยค่าไอแกนเวกเตอร์เป็นเวกเตอร์ 1 หน่วย ที่มีความยาวเท่ากับ 1 หากเราดูข้อมูลที่ลงไปในกราฟ จะเห็นว่าข้อมูลจะค่อนข้างมีรูปแบบที่เห็นได้ชัด เป็นจุดแนวเส้นทแยงมุม ซึ่งไอแกนเวกเตอร์จะแสดงให้เราเห็นว่าจำนวนข้อมูลใน 2 ชุดมีความสัมพันธ์กับเส้นตรงของกราฟอย่างไร ดังนั้นในกระบวนการทำงานของการสร้างไอแกนเวกเตอร์ของเมทริกซ์โควาเรียนซ์ เราจึงสามารถแยกเส้นที่อยู่บนกราฟ ซึ่งเป็นลักษณะเด่นของข้อมูลออกมาได้

ขั้นตอนที่ 5 การเลือกส่วนประกอบ และเปลี่ยนให้อยู่ในลักษณะของเวกเตอร์

- จะสังเกตได้ว่าไอแกนแวลูส์จะเป็นค่าที่ค่อนข้างแตกต่างกัน ในความเป็นจริงไอแกนเวกเตอร์กับค่าไอแกนแวลูส์ที่มีค่าสูงสุด จะเป็นหลักการส่วนประกอบของชุดข้อมูล เช่น หากไอแกนเวกเตอร์ กับไอแกนแวลูส์ที่มีค่าสูง จะเป็นจุดที่อยู่ตรงกลางของชุดข้อมูล ซึ่งเป็นความสัมพันธ์ที่เห็นได้อย่างชัดเจนระหว่างมิติของข้อมูล โดยไอแกนเวกเตอร์กับที่ถูกลบจากเมทริกซ์โควาเรียนซ์ ในขั้นตอนต่อไปจะเป็นลำดับโดยไอแกนแวลูส์ จากค่าสูงสุดไปต่ำสุด นี้จะให้เราสามารถเรียงลำดับความสำคัญของปัจจัยได้ ดังนั้นจึงสามารถตัดสินใจที่จะเพิกเฉยต่อ ปัจจัยที่มีความสำคัญน้อย หรือไอแกนแวลูส์มีค่าน้อย ซึ่งไม่ถือว่าเป็นการสูญเสียข้อมูลมากนัก หากคุณปล่อยบางปัจจัยออกไป ข้อมูลสุดท้ายจะมีมิติน้อยกว่าข้อมูลเริ่มต้น ถ้าข้อมูลเริ่มต้นมี n มิติ และทำการคำนวณ n ของไอแกนเวกเตอร์ และไอแกนแวลูส์ และดังนั้นเราจะเลือกเพียงไอแกนเวกเตอร์ p แรก และข้อมูลชุดสุดท้ายจะมีเพียง p มิติ จากนั้นต้องทำการแปลงให้อยู่ในรูปของเวกเตอร์ (Feature vector) โดยการนำค่าไอแกนเวกเตอร์ที่เลือกมาจาก list ของไอแกนเวกเตอร์ และเปลี่ยนเป็นเมตริกซ์ให้ไอแกนเวกเตอร์อยู่ในคอลัมภ์

$$\text{FeatureVector} = (\text{eig1 eig2 eig3 ... eign})$$

ขั้นตอนที่ 6 กลายมาเป็นข้อมูลชุดใหม่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- นี่จะเป็นขั้นตอนสุดท้ายของ PCA และเป็นขั้นตอนที่ง่ายที่สุด โดยเราจะเลือกส่วนประกอบหรือปัจจัย (eigenvector) ซึ่งเราอยากจะนำข้อมูลทั้งหมดของเรา และเปลี่ยนเป็นรูปแบบของเวกเตอร์ โดยทำการทรานส์โพส (Transpose) ของเวกเตอร์ และคูณส่วนนี้ด้วยชุดข้อมูลเริ่มต้น

$$\text{FinalData} = \text{RowFeatureVector} \times \text{RowDataAdjust},$$

ซึ่ง RowFeatureVector เป็นเมทริกซ์ที่มีไอแกนเวกเตอร์อยู่ในคอลัมน์ที่ถูกทรานส์โพส ดังนั้นไอแกนเวกเตอร์ในตอนนี้จะอยู่ในแนวนอน, ไอแกนเวกเตอร์ที่มีค่าสูงจะอยู่ด้านบนสุด และ RowDataAdjust เป็นข้อมูลที่หักออกจากค่ากลาง และทรานส์โพส ส่วน FinalData เป็นชุดข้อมูลสุดท้าย และประกอบข้อมูล (Data Items) ในคอลัมน์ และมีดัชนีทานไปกับแถว

2.3.4 K-Means Clustering and K-nearest neighbor algorithm

K-Means เป็นอัลกอริทึมที่ใช้แยกข้อมูลออกเป็นกลุ่มๆ โดยอาศัยคุณลักษณะต่างๆ เพื่อแยกข้อมูลออกเป็น K กลุ่ม การจัดกลุ่มแบบ K-Means ทำได้โดย

1. หาจุดเซ็นทรอยด์ (Centroid)
2. หาระยะทางของแต่ละข้อมูลกับจุดเซ็นทรอยด์
3. จัดกลุ่มข้อมูลไปยังกลุ่มที่มีระยะทางจากจุดเซ็นทรอยด์ของกลุ่มนั้นสั้นที่สุด

ในตอนแรกต้องกำหนดจุดเซ็นทรอยด์ตั้งต้นไว้ก่อน โดยการเลือกจุดข้อมูลใดๆ ในชุดข้อมูลนั้น ในการทดลองจะทำการกำหนดจุดเซ็นทรอยด์ 3 จุด จากนั้นทำการคำนวณระยะทางจากข้อมูลแต่ละตัวถึงจุดเซ็นทรอยด์ทั้งสามจุด โดยใช้การคำนวณระยะทางแบบยูคลิดีน (Euclidean distance) โดยมีสมการดังนี้

$$\text{Distance} = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \quad (2.6)$$

เมื่อคำนวณหาระยะทางจากข้อมูลแต่ละตัวกับจุดเซ็นทรอยด์ทั้งสามจุดแล้ว จะทำการแยกข้อมูลแต่ละตัวไปยังกลุ่มที่มีระยะทางจากจุดเซ็นทรอยด์ของกลุ่มนั้นน้อยที่สุด เมื่อได้ชุดข้อมูลของกลุ่มที่ 1, 2, 3 มาแล้ว ก็จะต้องทำการคำนวณหาจุดเซ็นทรอยด์ของกลุ่มนั้นใหม่โดยดูจากชุดข้อมูลในกลุ่มนั้นๆ คำนวณหาจุดเซ็นทรอยด์ โดยการหาค่าเฉลี่ยของค่า x และ y ของข้อมูลแต่ละตัวในกลุ่มนั้น

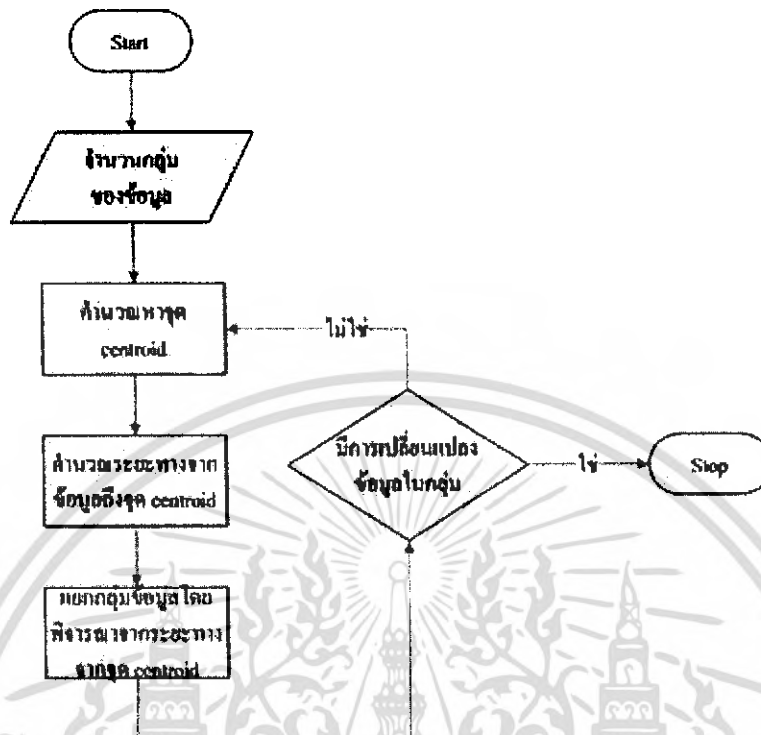
$$c = \left(\frac{x_1 + x_2 + \dots + x_n}{m} \right), \left(\frac{y_1 + y_2 + \dots + y_n}{m} \right) \quad (2.7)$$

n = จำนวนข้อมูลของกลุ่มนั้น

m = จำนวนจุด เซ็นทรอยด์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อได้จุดเซ็นทรอยด์ใหม่ของแต่ละกลุ่มแล้วก็ทำการคำนวณหาระยะทางแบบเดิมทำวนไปเรื่อยๆ จนกว่าค่าของจุดเซ็นทรอยด์จะคงที่



รูปที่ 2.3 กลไกการทำงานของ K-means

2.3.5 K-Nearest Neighbor (KNN)

วิธีการคำนวณอัลกอริทึมของ KNN มีขั้นตอนดังนี้

1. พิจารณาพารามิเตอร์ K = จำนวนของ nearest neighbor
2. คำนวณระยะทางระหว่างจุดเซ็นทรอยด์ที่มาจาก K-means กับข้อมูล
3. เรียงลำดับระยะทางและพิจารณา nearest neighbor โดยมีพื้นฐานบน K-th minimum distance
4. รวมกลุ่มของ nearest neighbor เข้าด้วยกัน
5. ใช้กลุ่มหลักของ nearest neighbor เป็นค่าที่ไว้คาดการณ์ความน่าจะเป็นของข้อมูลใหม่ๆว่า เป็นกลุ่มไหน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 การคัดกรองข้อความ (Text Detection)

การตรวจสอบเนื้อหาของหน้าเว็บเพจสามารถบ่งบอกว่าเว็บไซต์นั้นเป็นเว็บธนาคารหรือไม่ โดยพิจารณาคุณสมบัติต่างๆของโครงสร้าง และรูปแบบที่ได้รับมาจากการทดลองจากหน้าเว็บเพจจำนวนมาก ผลลัพธ์ของการตรวจสอบคุณสมบัติต่างๆ จะช่วยบ่งชี้ถึงโครงสร้างของเว็บไซต์ธนาคารแยกจากเว็บไซต์ที่ไม่ใช่ธนาคาร ซึ่งมีหลายรูปแบบที่ปรากฏเฉพาะในเว็บไซต์ธนาคาร ในการค้นหาข้อมูลที่ปรากฏภายในเอกสาร HTML จะต้องพิจารณาแต่ละส่วน มีดังนี้

- Title : จะอยู่ภายใน tag <TITLE> และ </TITLE> ซึ่งอยู่ในส่วนของ <HEAD>
- Meta : จะอยู่ภายใน tag <META> และ </META>
- Body : จะเป็นส่วนที่เราเห็นปรากฏอยู่ในเว็บเพจ

ในการทดลองนี้จะพิจารณาคุณสมบัติของหน้าเว็บเพจด้วยกันทั้งหมด 10 ปัจจัย ดังนี้

- 1) จำนวน IMG Tag ทั้งหมดในหน้าเว็บเพจ
- 2) จำนวน META Tag ทั้งหมดในหน้าเว็บเพจ
- 3) จำนวนคำหยาบคายทั้งหมดใน Meta และ Title
- 4) จำนวน SCRIPT Tag ทั้งหมดในหน้าเว็บเพจ
- 5) จำนวนลิงค์ทั้งหมดในหน้าเว็บเพจ
- 6) จำนวนลิงค์ที่เป็นรูปภาพทั้งหมดในหน้าเว็บเพจ
- 7) จำนวนลิงค์ที่เป็นตัวหนังสือทั้งหมดในหน้าเว็บเพจ
- 8) จำนวนลิงค์ที่ไม่ใช่ที่อยู่ภายในเว็บตัวเองในหน้าเว็บเพจ (Non-Relative Link)
- 9) จำนวนลิงค์ที่อยู่ภายในเว็บตัวเองในหน้าเว็บเพจ (Relative Link)
- 10) จำนวน PARAM Tag ทั้งหมดในหน้าเว็บเพจ (เพื่อนับจำนวนสื่อมัลติมีเดียทั้งหลาย)

เมื่อเราได้คุณสมบัติทั้ง 10 ปัจจัยของหน้าเว็บเพจมาแล้ว นำมาทำการ Normalized ข้อมูลก่อน จากนั้นนำค่าเหล่านี้มาทำลดมิติของข้อมูล (Dimension Reduction) จาก 10 ปัจจัย (10 มิติ) ให้เหลือ 2 มิติโดยไม่มีการสูญเสียความสำคัญของข้อมูลมากนัก ซึ่งใช้เทคนิคที่เรียกว่า Eigen (อธิบายอยู่ในเนื้อหา 2.3.3 หลักการวิเคราะห์แยกแยะส่วนประกอบ Principal Components Analysis: PCA)

สุดท้ายนำข้อมูลที่เหลือใน 2 มิติ มาทำการพล็อตกราฟแล้วจึงพิจารณาว่าข้อมูลที่ได้ควรเป็นกลุ่มของข้อมูลเว็บไซต์ธนาคาร หรืออยู่ในกลุ่มของเว็บไซต์ปกติ ที่ไม่มีสื่อลามกอนาจาร

72976

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

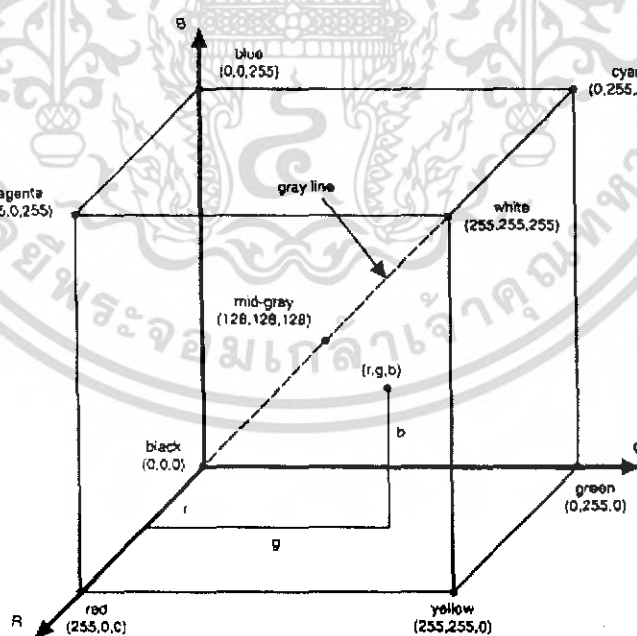
2.5 การคัดกรองรูปภาพ (Image Detection)

2.5.1 พื้นฐานและระบบโครงสร้างสี

2.5.1.1 ระบบโครงสร้างสีอาร์จีบี (RGB Color Model)

ในโครงสร้างนี้ สีแต่ละสีจะปรากฏในรูปแบบของแม่สี คือ สีแดง สีเขียว และ สีน้ำเงิน โครงสร้างสีนี้มีโครงสร้างเป็นลักษณะแกนคาร์ทีเซียนโคออดิเนต (Cartesian coordinate) โดยมีลักษณะเป็นทรงลูกบาศก์ มีค่าสีแดง สีเขียว และสีน้ำเงินอยู่ที่มุมทั้งสามที่เป็นแกน และมีสีฟ้าคราม สีม่วง สีเหลือง อยู่ที่มุมอีก 3 มุม สีดำจะอยู่ที่จุดกำเนิดคือถ้าทุกสีมีค่าเป็น 0 คือเป็น สีดำ สีขาวก็คือตรงข้ามกับสีดำซึ่งอยู่ที่มุมไกลสุดจากจุดกำเนิดในลักษณะทะแยงมุม ในรูปแบบนี้ค่าระดับสีเทา (Gray scale) จะอยู่บนเส้นระหว่างสีดำและสีขาว และสีอื่นๆ ก็จะมีตำแหน่งอยู่ภายในลูกบาศก์นี้

ค่าของสีคือจุดที่อยู่บนพื้นผิวหรือในลูกบาศก์ถูกกำหนดค่าโดยเวกเตอร์ที่ชี้ออกจากจุดกำเนิด ซึ่งช่องว่างแต่ละที่ในลูกบาศก์เรียกว่า ความลึกของพิกเซล (Pixel depth) โดยแม่สีแต่ละสีมีค่า 8 บิต มีได้ 256 ค่า เมื่อรวมสีทั้งหมดของความลึกของพิกเซลในลูกบาศก์ (24 บิตอาร์จีบี) จะได้ทั้งหมด $(2^8)^3 = 16,777,216$ สี ภาพในโครงสร้างสีอาร์จีบี ประกอบด้วยภาพสามระนาบที่เป็นอิสระจากกัน

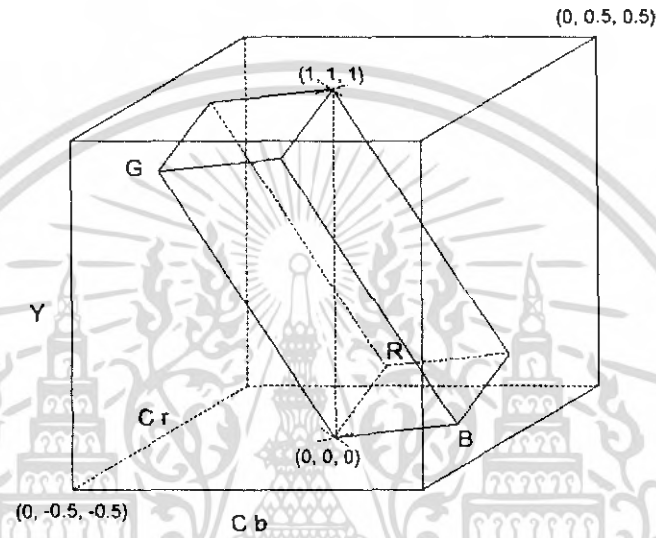


รูปที่ 1.4 โครงสร้างสีอาร์จีบี เป็นลูกบาศก์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.1.2 ระบบโครงสร้างสีวายซีบีซีอาร์ (YCbCr Color Model)

ในโครงสร้างนี้จะใช้เป็นที่แพร่หลายสำหรับคิวิตอลวีดีโอ ในรูปแบบของโครงสร้างนี้ ค่าปริมาณของแสงในการส่องสว่างจะเก็บข้อมูลไว้ในส่วนของ (Y) และในส่วนความแตกต่างของสีนั้นจะแบ่งเป็น 2 สีคือ Cb และ Cr โดย Cb จะแสดงให้เห็นถึงความแตกต่างของสีแดง และอ้างถึงค่าในหมวดสีแดงโดยโครงสร้างสีวายซีบีซีอาร์ มีความเที่ยงตรงและแม่นยำมาก ในส่วนของการส่องสว่าง และหมวดสีซึ่งเป็นโครงสร้างสีที่ใช้กันในการเข้ารหัสแบบเอ็มพีอีจี (MPEG & JPEG)



รูปที่ 2.5 ความสัมพันธ์ระหว่างโครงสร้างสี YCbCr กับ RGB

2.5.1.3 การแปลงรูปแบบสี

การแปลงรูปแบบสีจาก RGB เป็น YCbCr

$$Y = (0.299 * R) + (0.587 * G) + (0.114 * B) \quad (2.8)$$

$$Cb = (-0.168736 * R) - (0.331264 * G) + (0.5 * B) \quad (2.9)$$

$$Cr = (0.5 * R) - (0.418688 * G) - (0.081312 * B) \quad (2.10)$$

โดยกำหนดให้ค่าของ R, G และ B เป็นค่าของ สีแดง สีเขียว และ สีน้ำเงินตามลำดับ ซึ่งมีค่าตั้งแต่ 0 ถึง 255 แล้วค่าของ Y จะมีค่าในช่วงตั้งแต่ 0 ถึง 255 และ ค่า Cb, Cr อยู่ในช่วงตั้งแต่ 128- ถึง 128

การแปลงรูปแบบสีจาก YCbCr เป็น RGB

$$R = 1.164 * (Y - 16) + 1.596 * (Cr - 128) \quad (2.11)$$

$$G = 1.164 * (Y - 16) - 1.391 * (Cb - 128) - 0.813 * (Cr - 128) \quad (2.12)$$

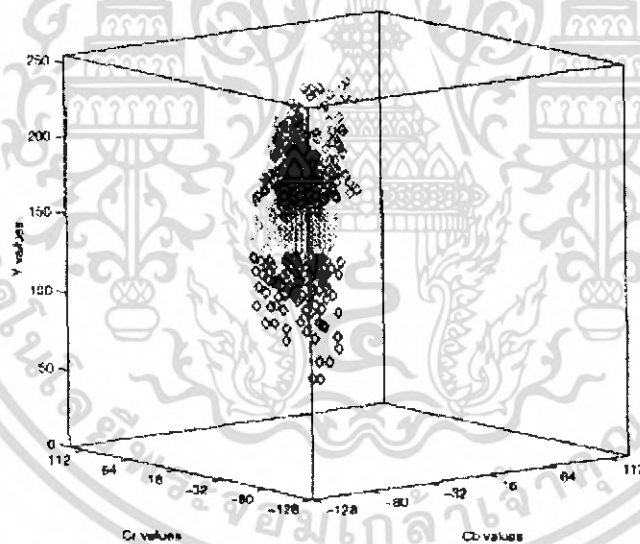
$$B = 1.164 * (Y - 16) + 2.018 * (Cb - 128) \quad (2.13)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.2 กระบวนการค้นหาพื้นที่สีผิว

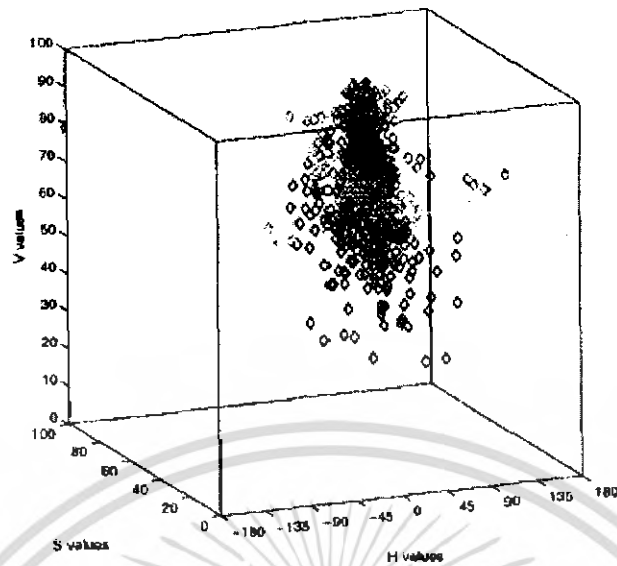
การตรวจหาภาพนารต้องอาศัยกระบวนการค้นหาพื้นที่สีผิว (Detection of Skin color regions) เพื่อแยกสีผิวมนุษย์ออกจากส่วนที่ไม่ใช่สีผิวในรูปภาพ จากงานวิจัยเรื่อง [Face Detection Using Quantized Skin Color Regions Marging and Wavelet Packet Analysis] ในหัวข้อเรื่อง “Skin Color segmentation” เสนอไว้ว่า ในงานวิจัยนั้นใช้โครงสร้างสีสองแบบ คือ YCbCr ซึ่งเป็นโครงสร้างสีที่ใช้กันในการเข้ารหัสแบบ MPEG และ JPEG และใช้โครงสร้างสีแบบ HSV (Hue ,Saturation ,Value)ที่ใช้ในงานคอมพิวเตอร์กราฟฟิก และใกล้เคียงกับการผสมสีของศิลปิน ข้อมูลที่ใช้วิเคราะห์เป็นการสุ่มตัวอย่างสีผิว 950 ตัวอย่าง โดยสุ่มจากหลากหลายเชื้อชาติและมีความสว่างแตกต่างกันไป โดยจะมีการแสดงกลุ่มตัวอย่างสีผิวในแกนของ YCbCr และ HSV ซึ่งแสดงเป็นกราฟข้อมูล 3 มิติ

จากรูปที่ 2.6 และ 2.7 เราสามารถสังเกตเห็นได้ว่า ตัวอย่างสีผิวทั้งใน YCbCr และ HSV ต่างก็มีการกระจายตัวที่มีลักษณะคล้ายกัน โดยข้อมูลค่อนข้างที่จะมีการรวมตัวกันอยู่เป็นกลุ่ม ในงานวิจัยนั้นมุ่งที่จะสร้างขอบเขต 3 มิติเพื่อให้ครอบคลุมกลุ่มของตัวอย่างสีผิวที่สุ่มได้มากที่สุดด้วยสมการเส้นตรง



รูปที่ 2.6 กลุ่มตัวอย่างสีผิวในแกนของโครงสร้างสี YCbCr

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

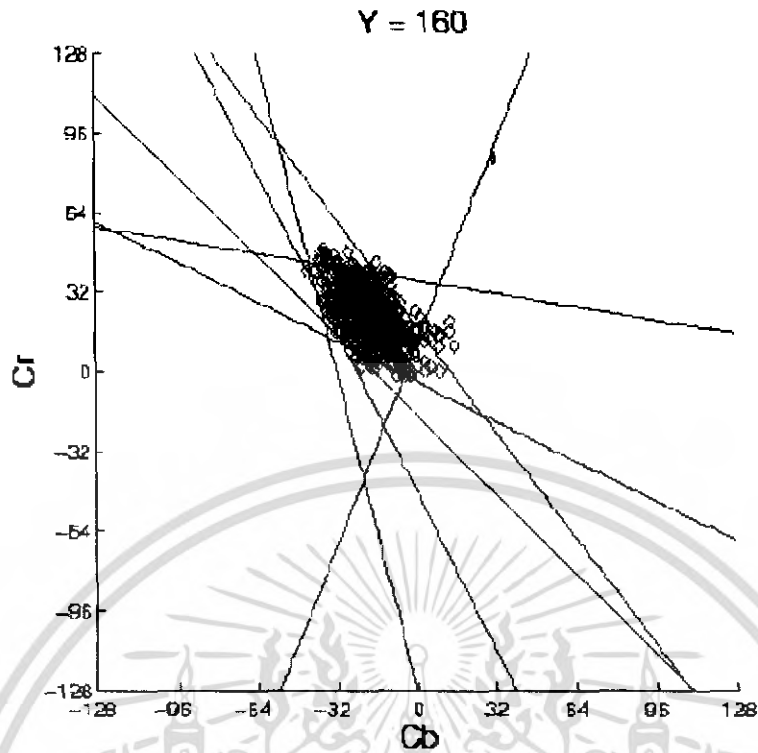


รูปที่ 2.7 กลุ่มตัวอย่างสีผิวในแกนของ โครงสร้างสี HSV

ในกราฟ YCbCr พบว่าการเปลี่ยนแปลงของค่า Y มีผลเพียงเล็กน้อยต่อการเปลี่ยนแปลงในระนาบ CbCr โดยทำการพิจารณาใน 2 กรณีคือ ภาพที่มีความมืด (ค่าของ Y มีค่าประมาณ 50) และภาพที่สว่าง (ค่าของ Y มีค่าประมาณ 240)

เมื่อกำหนดค่า $Y = 160$ เพื่อให้ได้กราฟในระนาบ CbCr และกำหนดสมการเส้นตรงให้ครอบคลุมกลุ่มตัวอย่างดังรูปที่ 2.8 การกำหนดค่า Y ค่าเดียวนี้ก่อให้เกิดความผิดพลาดในกระบวนการตรวจหาสีผิว จึงแบ่งช่วง Y ออกเป็นสองช่วงที่จุด $Y = 128$ ให้มีขอบเขตที่ต่างกันโดยกำหนดเป็นตัวแปรที่ต่างกันโดยที่จุด $Y = 128$ เป็นจุดแบ่งขอบเขตความมืดและสว่างและกำหนดสมการเส้นตรงแปดสมการบนพื้นที่ทั้งสองส่วน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.8 กรอบพื้นที่สีผิวโครงสร้างสี YCbCr ที่ระนาบ $Y = 160$

ทำการคำนวณหาค่าของ $\theta_1, \theta_2, \theta_3$ และ θ_4 ใน 2 กรณีคือ

เมื่อ $Y > 128$

$$\theta_1 = -2 + \frac{256 - Y}{16}$$

$$\theta_2 = 20 - \frac{256 - Y}{16}$$

$$\theta_3 = 6$$

$$\theta_4 = -8$$

เมื่อ $Y \leq 128$

$$\theta_1 = 6$$

$$\theta_2 = 12$$

$$\theta_3 = 2 + \frac{Y}{32}$$

$$\theta_4 = -16 + \frac{Y}{16}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากช่วงของค่า Y ทั้งสองช่วง สามารถกำหนดขอบเขตสีผิวได้ด้วยสมการต่อไปนี้

$$Cr \geq -2(Cb + 24)$$

$$Cr \geq -(Cb + 17)$$

$$Cr \geq -4(Cb + 32)$$

$$Cr \geq 2.5(Cb + \theta_1)$$

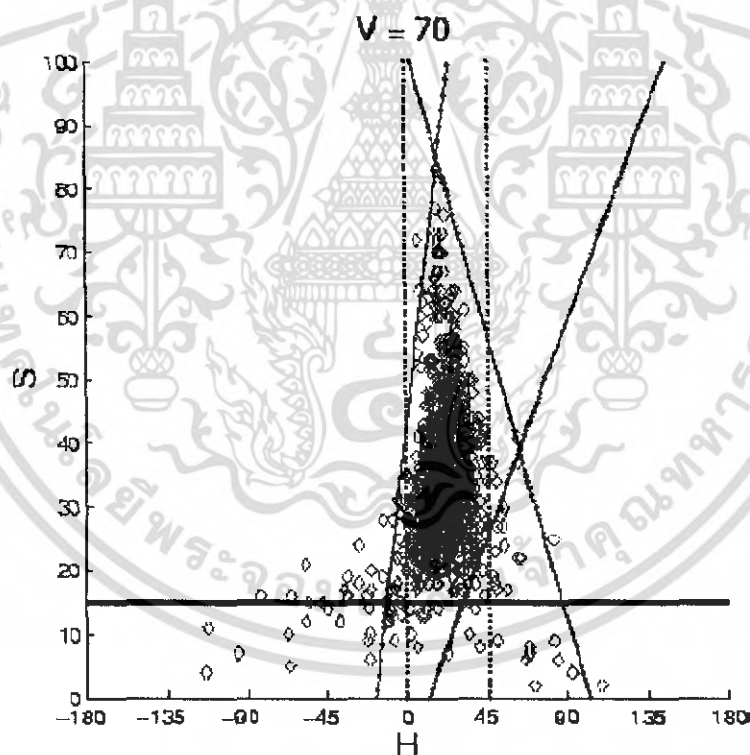
$$Cr \geq \theta_3$$

$$Cr \geq -0.5(\theta_4 - Cb)$$

$$Cr \leq \frac{220 - Cb}{6}$$

$$Cr \leq \frac{4}{3}(\theta_2 - Cb)$$

สมการเหล่านี้หมายความว่าพิกเซลเป็นสีผิวของมนุษย์ นอกเหนือจากนี้ไม่ใช่สีผิวมนุษย์ โครงสร้างสี YCbCr นั้นเป็นกรรมวิธีที่ดีในการตรวจสอบสีผิวมากกว่า โครงสร้างสี HSV อย่างไรก็ตามก็ยังมีผู้ใช้บางคนใช้การฉายภาพบนระนาบของ HS ซึ่งทำได้โดยการกำหนดค่า thresholds ให้กับ Hue และ Saturation ผลของการใช้ค่า thresholds นี้ เราพบว่าผลของการจำแนกสีผิวจะได้รับผลกระทบจากความแตกต่างของค่าความสว่าง



รูปที่ 2. 9 กรอบพื้นที่สีผิวโครงสร้างสี HSV ที่ระนาบ $V = 70$

ในทำนองเดียวกันกับกรณีของ YCbCr ทำการกำหนดค่า $V = 70$ และได้สมการความสัมพันธ์ของค่า Hue และ Saturation ดังนี้

$$S \geq 10$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$V \geq 40$$

$$S \leq -H - 0.1V + 110$$

$$\text{ถ้า } H \geq 0 \text{ แล้ว } S \leq 0.08(100-V)H + 0.5V$$

$$\text{ถ้า } H < 0 \text{ แล้ว } S \leq 0.5H + 35$$

2.5.3 กระบวนการปรับความคมชัดของภาพ

จุดประสงค์หลักของการปรับความคมชัดของภาพคือเพื่อทำการเน้นส่วนรายละเอียดในภาพ หรือ ปรับปรุงรายละเอียดของภาพที่ถูกทำให้พร่ามัว (Blur) อาทิเช่นความผิดพลาดหรือผลกระทบที่เกิดขึ้นในกระบวนการ ได้มาซึ่งภาพ (Image acquisition) การใช้กระบวนการของความแตกต่างของสีภายในภาพ (image differentiation) ที่แสดงถึงอัตราความไม่ต่อเนื่องของสีเพื่อตรวจหาเส้นขอบของภาพ และส่วนที่ไม่ต่อเนื่องกันอื่นๆ เช่น สิ่งรบกวน (noise)

กระบวนการปรับความคมชัดของภาพมีพื้นฐานบน first และ second-order derivatives ซึ่งแสดงจุดสนใจบนพื้นที่ที่มีค่าระดับสีเทาคงที่ (flat segments), จุดเริ่มและจุดสิ้นสุดของความไม่ต่อเนื่อง (step and ramp discontinuities) และ พื้นที่ที่มีระดับสีไล่เลี่ยกัน (ramps) โดยมีคุณสมบัติดังนี้

2.5.3.1 First derivative

Derivative ในฟังก์ชันดิจิทัลอนันต์ได้ถูกอธิบายในรูปของความแตกต่าง (Difference) โดย First derivative ต้องมีคุณสมบัติดังนี้

- ต้องมีค่าเท่ากับ ศูนย์ บนพื้นที่ที่มีค่าระดับสีเทาคงที่ (flat segments)
- ต้องมีค่าไม่เท่ากับศูนย์ที่จุดเริ่มต้นของพื้นที่ที่มีระดับสีเทาต่างกัน (gray-level step or ramp)
- ต้องมีค่าไม่เท่ากับศูนย์บนพื้นที่ที่มีระดับสีไล่เลี่ยกัน (ramps)

โดย First Order derivative ของฟังก์ชันหนึ่งมิติแนวแกน เอ็กซ์ ($f(x)$) คือ

$$\frac{\partial f}{\partial x} = f(x+1) - f(x) = f'(x) \quad (2.14)$$

2.5.3.2 Second derivative

Derivative ในฟังก์ชันดิจิทัลอนันต์ได้ถูกอธิบายในรูปของความแตกต่าง (Difference) โดย Second derivative ต้องมีคุณสมบัติดังนี้

- ต้องมีค่าเท่ากับ ศูนย์ บนพื้นที่ที่มีค่าระดับสีเทาคงที่ (flat segments)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ต้องมีค่าไม่เท่ากับศูนย์ที่จุดเริ่มต้นของพื้นที่ที่มีระดับสีเทาต่างกัน(gray-level step or ramp)
- ต้องมีค่าเท่ากับศูนย์บนพื้นที่ที่มีระดับสีไล่เฉียดกัน (ramps)

โดย Second Order derivative ของฟังก์ชันหนึ่งมิติแนวแกน เอ็กซ์($f(x)$) คือ

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2} &= f'(x) - f'(x-1) \\ &= (f(x+1) - f(x)) - (f(x) - f(x-1)) \quad (2.15) \\ &= f(x+1) + f(x-1) - 2f(x)\end{aligned}$$

2.5.3.3 การใช้ First Derivative ในการเพิ่มคุณภาพของภาพ

FirstDerivativeในการประมวลผลภาพนั้นเป็นเครื่องมือในการหาค่าความลาดชัน (gradient) สำหรับฟังก์ชัน $f(x,y)$ ค่าความชันของ f ที่ระยะพิกัด (x,y) ถูกแสดงในรูปของเวกเตอร์ที่มีสองคอลัมน์ก็ดังนี้

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (2.16)$$

ในทางปฏิบัติเราสามารถประมาณค่าขนาดความชันได้ด้วยการใช้ค่าสมบูรณ์ ได้ดังนี้

$$\nabla f \approx |G_x| + |G_y| \quad (2.17)$$

พิจารณารูปที่ 2.10(a) จุดของภาพที่มีขอบเขต 3×3 พิกเซล (ค่า z เป็นค่าระดับสีเทา) Robert[1965] ได้เสนอคำนิยามในการประมวลผลภาพยุคแรกโดยการใช้ค่าความแตกต่างในแนวทะแยงเป็น $G_x = (z_9 - z_5)$ และ $G_y = (z_8 - z_6)$ และจากสมการที่ 2.9 สามารถคำนวณหาขนาดของความชันได้ดังนี้

$$\nabla f \approx |z_9 - z_5| + |z_8 - z_6| \quad (2.18)$$

สมการที่ 2.17 นี้สามารถใช้ในลักษณะของ masks ในรูปที่ 2.10(b) และ (c) โดยเรียกว่า Roberts cross-gradient operators

อย่างไรก็ตามการใช้ขนาดของ mask เป็นเลขคู่่นั้นทำได้ไม่สะดวก โดยขนาดของ mask ที่เล็กที่สุดที่เราสนใจคือ 3×3 พิกเซลซึ่งเมื่อทำการคำนวณหาขนาดของความชันที่จุดที่มีค่าระดับสีเทาเท่ากับ z_5 แต่ใช้ mask ขนาด 3×3 ได้ดังนี้

$$\begin{aligned}\nabla f \approx & |(z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)| \quad (2.19) \\ & + |(z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)|\end{aligned}$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

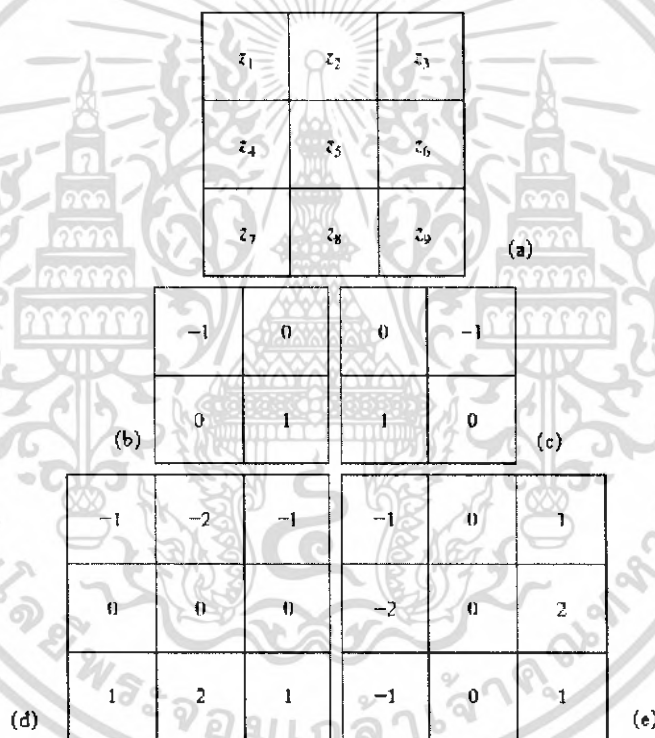
2.5.3.4 การใช้ Second Derivative ในการเพิ่มคุณภาพของภาพ

ลักษณะของการเพิ่มคุณภาพของภาพด้วย second-order derivatives เราสนใจในการทำตัวกรองแบบสมมาตร (isotropic filters) ซึ่งเป็นอิสระต่อทิศทางของความไม่ต่อเนื่องบนภาพ (rotation invariant) กระบวนการแบบสมมาตรที่ง่ายที่สุดได้แก่ Laplacian ซึ่งมีสมการดังนี้

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (2.20)$$

ซึ่งสามารถแสดงในรูปผลบวกของ 2 องค์ประกอบดังนี้

$$\nabla^2 f = [f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1)] - 4f(x, y) \quad (2.21)$$



รูปที่ 2.10 (a) จุดของภาพที่มีขอบเขต 3 x 3 pixel (ค่า z เป็นค่าระดับสีเทา)

(b)(c) Roberts cross-gradient operators

(d)(e) Sobel operators

2.5.3.5 การตรวจหาเส้นขอบด้วย กระบวนการโซเบล (Sobel operators)

กระบวนการโซเบล (Sobel Edge Operator) เป็นการตรวจหาขอบของภาพ ซึ่งข้อสันนิษฐานเบื้องต้น คือผิวของมนุษย์จะมีความราบเรียบต่อเนื่องกันไม่มีขอบเกิดขึ้นภายใน โดยการ

ตรวจหาขอบของ Sobel จะทำโดยการทำ คอนโวลูชัน (convolution) ซึ่งจะนำพิกเซลที่อยู่ล้อมรอบ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปเผยแพร่โดยไม่ได้รับอนุญาตจากเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

พิกเซลที่กำลังพิจารณาทำการคำนวณ โดยเราจะมีตัวเลขคงที่อยู่กลุ่มหนึ่ง เรียกว่า ค่าตัวร่วม (Mask Coefficient) เข้าไปคูณกับพิกเซลเหล่านั้น จากนั้นนำผลคูณแต่ละตัวมาทำการบวกเข้าด้วยกัน ผลลัพธ์สุดท้ายจะเก็บไว้ในตำแหน่งพิกเซลที่กำลังพิจารณาของภาพที่ผ่านการประมวลผลแล้ว

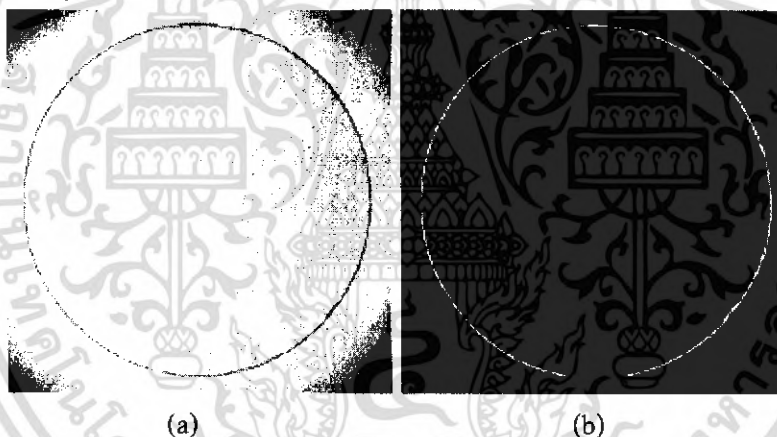
ในทิศทางแนวนอนใช้ Mask Coefficient ขนาด 3x3 ดังรูปที่ 2.10 (d) คือ

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

และในทิศทางแนวตั้งใช้ mask coefficient ขนาด 3x3 ดังรูปที่ 2.10 (e) คือ

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \text{ or } \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$$

อย่างไรก็ตาม Sobel operators สามารถทำได้ด้วยสมการที่ 2.19 เช่นกัน โดยการให้ค่าน้ำหนักด้วยสองนั้นเพื่อให้ความสำคัญกับพิกเซลที่กึ่งกลางมากที่สุดโดยค่าน้ำหนักทั้งหมดต้องรวมกันแล้วเท่ากับ ศูนย์



รูปที่ 2.11 (a) ตัวอย่างภาพต้นฉบับที่ต้องการหาเส้นขอบ

(b) ภาพที่ถูกตรวจหาเส้นขอบ (Sobel gradient)

2.5.4 กระบวนการทดสอบ

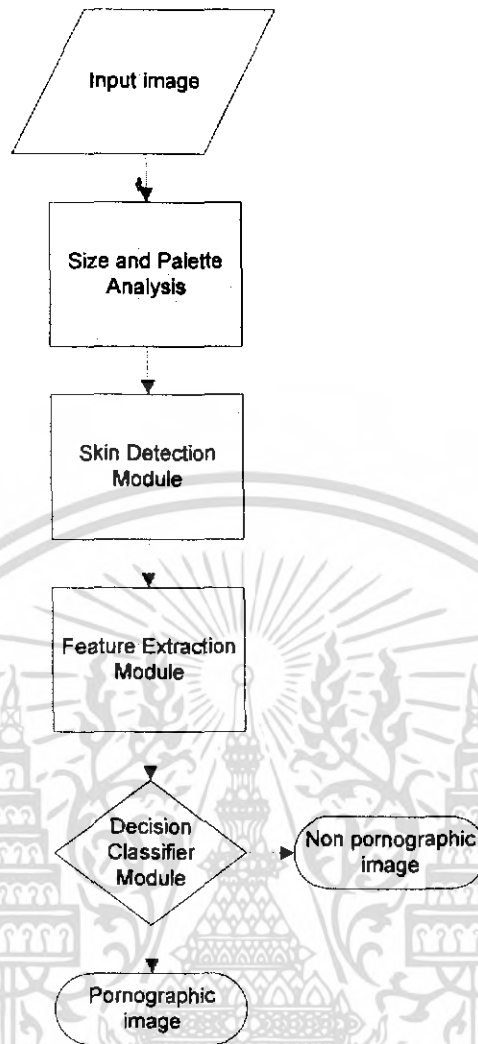
การตรวจหาภาพอนาจารภายในรูปภาพจะทำการระบุตำแหน่งของพิกเซลของผิวหนังและรวมกลุ่มไว้ด้วยกันเป็นกลุ่มของผิวหนัง โดยใช้รายละเอียดของสี, พื้นผิว และขอบ คุณลักษณะต่างๆถูกดึงมาจากกลุ่มของผิวหนังและใช้เป็นอินพุทเพื่อใส่เข้าไปในการจัดกลุ่ม

ความสัมพันธ์ระหว่างเปอร์เซ็นต์ของสีผิวกับรูปอนาจารมีค่อนข้างมาก ในขั้นตอนแรกต้องมีการแบ่งกลุ่มของสีผิว แต่การจำแนกสีของผิวหนังอาจมีความผิดพลาดได้ เนื่องจาก

- 1) ภาพมีคุณภาพต่ำ เช่น ภาพที่มี contrast ต่ำ
- 2) ภาพที่มีวัตถุที่มีสีใกล้เคียงกับสีผิวมนุษย์
- 3) ภาพผิวหนังที่มีสีผิวจางเนื่องจากการส่องสว่างและการสะท้อนของแสง

การตรวจสอบภาพอนาจารมี 4 ขั้นตอนหลักๆ คือ

1. Size and palette Analysis เป็นการกรองภาพแบบพื้นฐานที่สุด เช่น ภาพมีขนาดเล็กกว่าที่เราตั้งเอาไว้ และภาพที่มีจำนวนสีไม่มาก (น้อยกว่า 50 พิกเซล) เนื่องจากคุณสมบัติไม่เหมือนกับภาพอนาจาร
2. Skin Detection ในขั้นตอนนี้เราจะพิจารณาจากสี, พื้นผิว, และขอบ พิจารณาพื้นผิวที่เชื่อมต่อกัน
3. Feature Extraction การดึงคุณสมบัติ 6 อย่างจากข้อมูลภาพ
4. Decision Classifier เป็นการจัดประเภทของข้อมูลภาพ



รูปที่ 2.12 การทำงานของการตรวจสอบภาพอนาจาร

2.5.4.1 การกำจัดภาพขนาดเล็กกว่ากำหนด (Size and palette Analysis)

จากสมมติฐานภาพอนาจารควรมีขนาดใหญ่และใช้พื้นที่ส่วนใหญ่ของหน้าเอกสารเว็บบ ดังนั้นเพื่อจำกัดการตรวจสอบและเพิ่มประสิทธิภาพในการตรวจสอบในขั้นต้น ทำการตัดภาพที่มีขนาดเล็กกว่ากำหนด ออกแบบให้เท่ากับ 50x50 พิกเซล ภาพที่มีขนาดเล็กกว่าที่กำหนดจะไม่ถูกนำมาพิจารณา

2.5.4.2 การตรวจหาสีผิวมนุษย์ (skin detection module)

ในการตรวจหาสีผิวมนุษย์ที่ออกแบบมี 4 ขั้นตอนย่อยดังนี้

1. การตรวจหาพิกเซลในช่วงของสีผิว (Skin Tone Color Detection) จะพิจารณาจากค่า RGB ที่ถูกแปลงเป็นค่า YCbCr ว่าพิกเซลนั้นมีค่าอยู่ในช่วงสีผิวหรือไม่
2. การตรวจหาขอบเขตพื้นที่ของพิกเซลที่มีค่าอยู่ในช่วงสีผิว (Skin Region Expansion) จากการพิจารณาข้อที่ 1 อาจมีการรวมถึงพิกเซลข้างเคียงที่มีสีอยู่ในช่วงของสีผิว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3. การแบ่งพื้นที่ของสีผิวด้วยเส้นขอบ (Skin Region Segmentation) การตรวจหาขอบเขตจากข้อที่ 2 จะมีการพิจารณาร่วมกับกระบวนการตรวจหาขอบของ โซเบล (Sobel) พิกเซลที่อยู่บนขอบจะไม่ถูกตรวจจับว่าเป็นสีผิว การทำเช่นนี้เพื่อให้แน่ใจว่าเป็นพื้นที่ของผิวจริง
4. การกำจัดพื้นที่ของสีผิวที่มีขนาดเล็กกว่าที่กำหนดไว้ (Skin Blob Detection) กำหนดไว้ที่ 50x50 พิกเซล

2.5.4.3 การตรวจสอบคุณสมบัติของภาพ (Feature Extraction Module)

ภาพที่ถูกรตรวจสอบว่าเป็นภาพอนาจาร จะถูกพิจารณาคุณสมบัติ 6 ข้อ ดังตารางที่ 2.1

ลำดับ	คุณสมบัติ	รายละเอียด
1.	เปอร์เซ็นต์ของสีผิวในภาพ	อัตราของพิกเซลที่มีค่าอยู่ในช่วงของสีผิวทั้งหมดเทียบขนาดของภาพ
2.	เปอร์เซ็นต์ของพื้นที่สีผิวที่มีขนาดใหญ่ที่สุด	อัตราของพิกเซลในพื้นที่ที่มีสีผิวมากที่สุดเทียบกับขนาดของภาพ
3.	จำนวนชิ้นส่วนของพื้นที่ผิว	จำนวนของพื้นที่ที่มีสีผิวในภาพ
4.	อัตราความสูงของพื้นที่ผิว	ขนาดความสูงของพื้นที่ที่มีสีผิวขนาดใหญ่ที่สุดต่อขนาดความสูงของภาพ
5.	อัตราความกว้างของพื้นที่ผิว	ขนาดความกว้างของพื้นที่ที่มีสีผิวขนาดใหญ่ที่สุดต่อขนาดความสูงของภาพ
6.	อัตราส่วนของพื้นที่ผิวขนาดใหญ่เทียบกับจำนวนพื้นที่ผิวทั้งหมด	อัตราส่วนของพื้นที่ที่มีสีผิวขนาดใหญ่ที่สุดในภาพเทียบกับขนาดพื้นที่ที่มีสีผิวทั้งหมดในภาพ

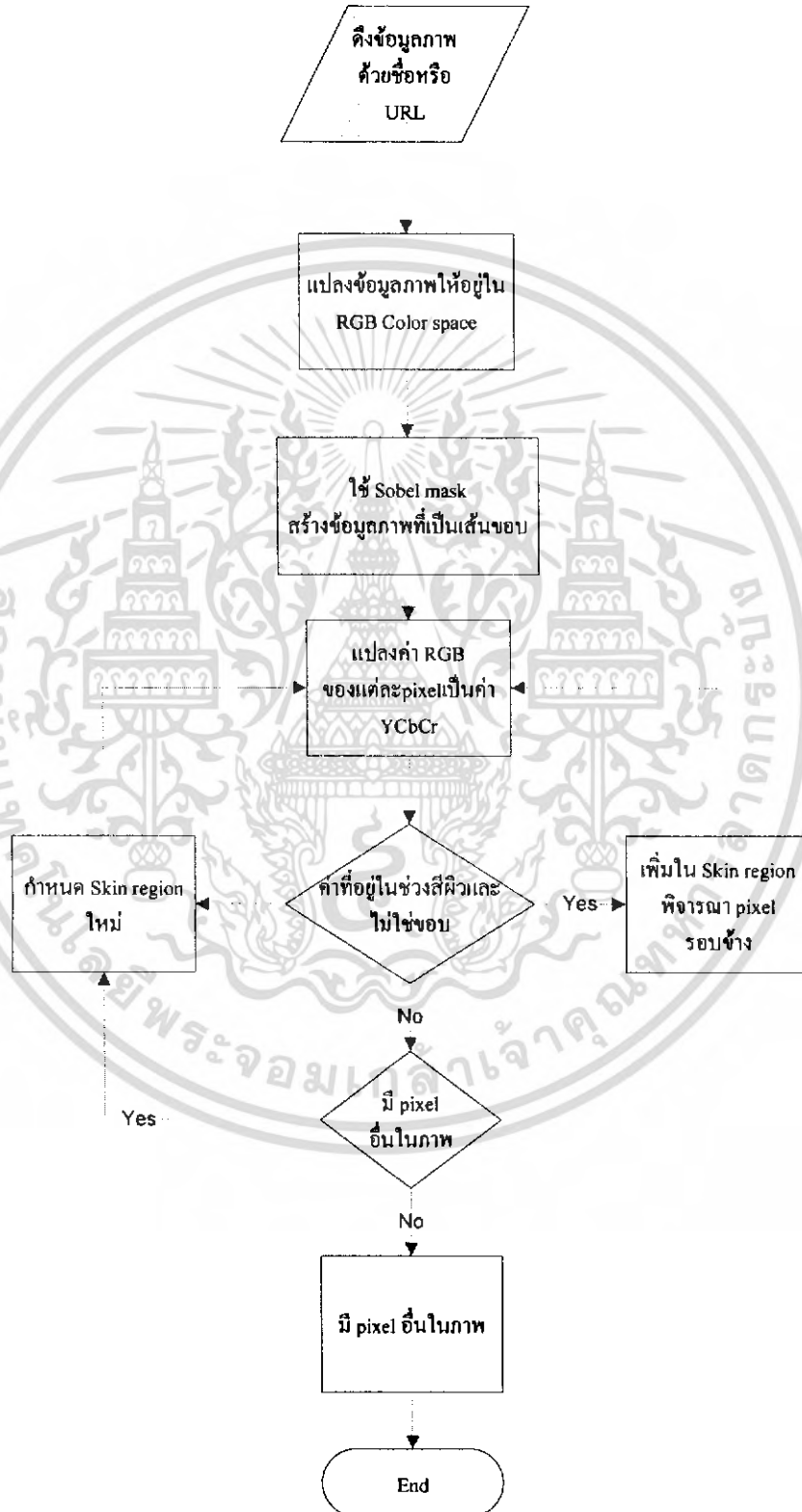
ตารางที่ 2.1 คุณสมบัติ 6 ประการที่ได้จากภาพ

คุณสมบัติทั้ง 6 ข้อสามารถใช้จำแนกภาพอนาจารได้ โดยภาพที่ไม่มีกลุ่มของสีผิวจะถูกมองว่าไม่เป็นภาพอนาจาร ซึ่งไม่มีคุณสมบัติข้อที่ 1 จากสมมติฐานภาพอนาจารจะต้องประกอบด้วยกลุ่มของสีผิวขนาดใหญ่ต่อเนื่องกัน ซึ่งสามารถพิจารณาได้จากคุณสมบัติข้อที่ 2 และ 3 ที่แสดงถึงอัตราส่วนของพื้นที่ผิวที่มีขนาดใหญ่ที่สุดและจำนวนชิ้นส่วนของพื้นที่ผิว ภาพที่มีจำนวนของพื้นที่ผิวมากอาจจะไม่ใช่ภาพอนาจาร ความสัมพันธ์ระหว่างขนาดและรูปร่างของกลุ่มสีผิว ถ้าเป็นภาพอนาจารจะมีกลุ่มสีผิวขนาดใหญ่ คุณสมบัติข้อ 5 และ 6 เป็นความสัมพันธ์ของ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัตราส่วนของความกว้างและความสูงของกลุ่มสีผิวต่อความกว้างและความสูงของภาพ มีแนวโน้มว่ากลุ่มสีผิวจะครอบคลุมพื้นที่ส่วนใหญ่ของภาพ

โครงการนี้ได้จัดทำโปรแกรมตรวจสอบภาพอนาจารที่สามารถวิเคราะห์คุณสมบัติของภาพได้ โดยมีขั้นตอนการทำงานดังรูปที่ 2.13



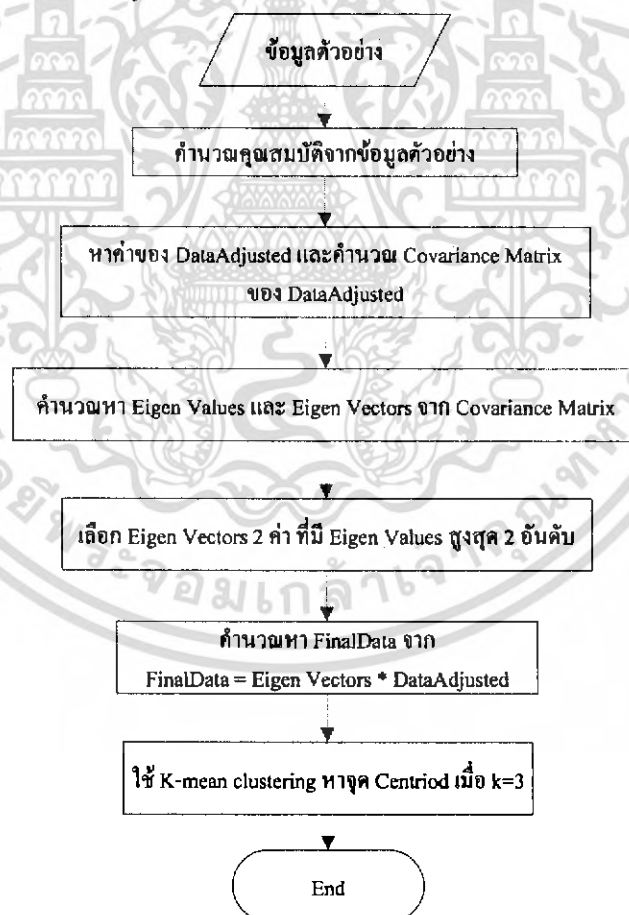
รูปที่ 2.13 การทำงานของ โปรแกรมการตรวจสอบภาพอนาจาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.5.4.4 การจัดประเภทของภาพด้วยคุณสมบัติ (Decision Classifier)

เมื่อพิจารณาคุณสมบัติของภาพทั้ง 6 ข้อ เราสามารถจำแนกประเภทของภาพได้ 3 กลุ่ม ได้แก่ ภาพอนาจาร , ภาพที่ไม่ใช่ภาพอนาจารแต่มีส่วนประกอบของสีผิว และภาพที่ไม่มีส่วนประกอบของสีผิว เพื่อให้สามารถพิจารณาคุณสมบัติที่มีความสำคัญมากที่สุดจำเป็นต้องทำการรวมคุณสมบัติและลดคุณสมบัติบางประการ โดยไม่ทำให้ข้อมูลสูญหายมากเกินไปด้วยหลักการการวิเคราะห์ส่วนประกอบที่ต้องทำการคำนวณหาไอแกนเวกเตอร์และไอแกนแวลูส์ของชุดข้อมูลตัวอย่าง แล้วทำการหาจุดเซ็นทรอยด์ 3 จุดที่แสดงถึงข้อมูลภาพ 3 ประเภทด้วยหลักการของ K-mean Clustering และเมื่อต้องการจำแนกภาพสามารถทำได้โดยการคำนวณหาระยะทางที่ใกล้ที่สุดเทียบกับจุดเซ็นทรอยด์ด้วยหลักการของ K-Nearest Neighbor (KNN)

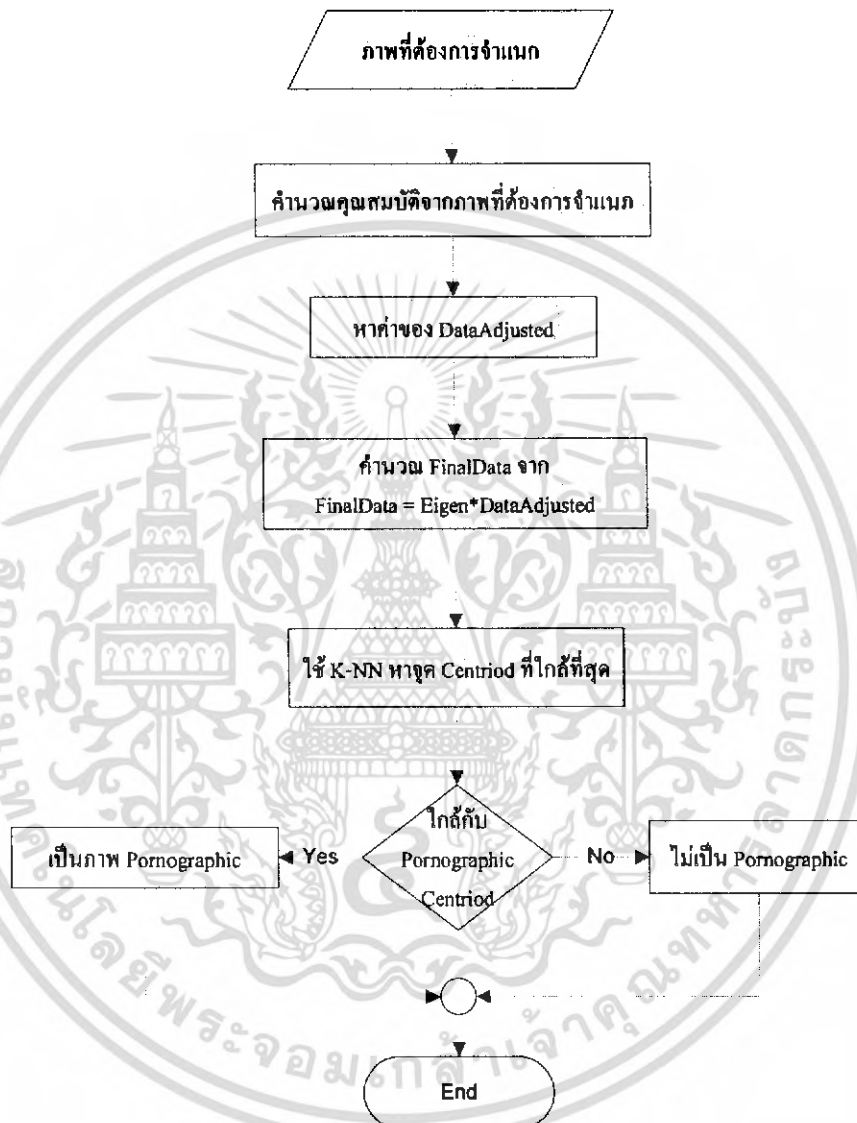
โครงการนี้ได้จัดทำโปรแกรมจัดประเภทของภาพด้วยคุณสมบัติ ที่สามารถจำแนกประเภทของภาพด้วยคุณสมบัติของภาพได้ โดยมีขั้นตอนการคำนวณหาไอแกนเวกเตอร์และไอแกนแวลูส์ของชุดข้อมูลตัวอย่าง แล้วทำการหาจุดเซ็นทรอยด์ 3 จุดที่แสดงถึงข้อมูลภาพ 3 ประเภทด้วยหลักการของ K-mean Clustering ดังรูปที่ 2.14



รูปที่ 2.14 การทำงานของการคำนวณหาไอแกนเวกเตอร์และไอแกนแวลูส์ของชุดข้อมูลตัวอย่างและ การหาจุดเซ็นทรอยด์ 3 จุด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อคำนวณหาไอแกนเวกเตอร์และไอแกนแวลูส์ของชุดข้อมูลตัวอย่าง แล้วทำการหาจุดเซ็นทรอยด์ 3 จุด ได้จะกระบวนการข้างต้นแล้ว เมื่อต้องการจำแนกภาพสามารถทำได้โดยการคำนวณหาระยะทางที่ใกล้ที่สุดเทียบกับจุดเซ็นทรอยด์ด้วยหลักการของ K-Nearest Neighbor (KNN) มีขั้นตอนการทำงานดังรูปที่ 2.15



รูปที่ 2.15 การจำแนกภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

หลักการพื้นฐานบนเครือข่าย

(Networking Concept)

3.1 โพรโทคอลเอชทีทีพี (HTTP: HyperText Transfer Protocol)

โพรโทคอลนี้สร้างขึ้นสำหรับบริการที่เรียกว่า WWW (World Wide Web) ในเครือข่ายอินเทอร์เน็ตโดยเฉพาะ โพรโทคอลนี้จะเป็นตัวกำหนดวิธีการส่งข้อมูลหรือไฟล์ระหว่างเครื่องคอมพิวเตอร์ลูกข่าย (Client) และ เครื่องแม่ข่าย (Server) รวมถึงกำหนดกฎระเบียบในการติดต่อด้วยโพรโทคอลนี้ถูกออกแบบมาให้มีความกระชับรัด สามารถทำงานได้รวดเร็ว มีการะบวนการทำงานที่ไม่ซับซ้อน และมีคำสั่งที่ใช้งานไม่มากนัก แต่สามารถรองรับข้อมูลได้ทุกแบบ ไม่ว่าจะเป็นข้อมูลทั่วไปที่เข้ารหัสแบบ MIME หรือข้อมูลที่เป็นกราฟิก เช่น ไฟล์ที่เป็น GIF หรือ JPEG เป็นต้น

3.1.1 วิธีการติดต่อของโพรโทคอล HTTP

โพรโทคอล HTTP อยู่บนพื้นฐานของระบบเครือข่ายไคลเอนต์ / เซิร์ฟเวอร์ (Client / Server) ที่ต้องมีการร้องขอรับบริการจากไคลเอนต์ (request) และการตอบสนองหรือการให้บริการของเซิร์ฟเวอร์ (response) จึงสามารถแบ่งการทำงานออกเป็น 2 ด้านคือ ด้านเว็บเซิร์ฟเวอร์ และด้านไคลเอนต์ โดยไคลเอนต์จะติดต่อเข้ามายังเซิร์ฟเวอร์และอ้างถึงแอดเดรสของเซิร์ฟเวอร์โดยใช้รูปแบบของ URL ส่วนด้านเซิร์ฟเวอร์จะส่งข้อมูลกลับมาในรูปแบบที่เป็นภาษา HTML (HyperText Markup Language) โดยที่โพรโทคอล HTTP ใช้วิธีการเข้ารหัสในแบบ MIME เป็นมาตรฐานของการทำงาน โพรโทคอลนี้ถูกออกแบบมาให้สามารถรับส่งข้อมูลผ่าน Proxy หรือ Firewall ต่าง ๆ ได้ โดยอาศัยการเชื่อมต่อผ่านทางโพรโทคอล TCP/IP อีกทีหนึ่ง โดยใช้พอร์ตหมายเลข 80 เป็นช่องทางมาตรฐานในการติดต่อ ในทางปฏิบัติจะใช้พอร์ตหมายเลขอื่นก็ได้ ในปัจจุบันเว็บเบราว์เซอร์ทั่วไปจะกำหนดค่ามาตรฐานไว้ที่พอร์ต 80 ดังนั้นหากมีการกำหนดไว้ที่พอร์ตอื่น จะทำให้เกิดความลำบากต่อผู้ใช้ที่ต้องระบุหมายเลขพอร์ตลงใน URL ด้วยด้วยเหตุที่การทำงานของโพรโทคอล HTTP เป็นแบบไคลเอนต์และเซิร์ฟเวอร์ ดังนั้นการติดต่อสื่อสารใดๆผ่านโพรโทคอลนี้จำเป็นต้องมีเครื่องตัวแม่กับตัวลูก การสื่อสารจะสมบูรณ์ได้ การติดต่อกันระหว่างไคลเอนต์ไปยังเซิร์ฟเวอร์ผ่านโพรโทคอล HTTP มีขั้นตอนดังนี้

ขั้นแรก : Open Socket

ไคลเอนต์ จะสร้างการเชื่อมต่อ (Connection) กับเซิร์ฟเวอร์ผ่านซ็อกเก็ต (Socket) เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้ไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขั้นที่สอง : Request

ไคลเอนต์ส่งคำร้องขอข้อมูลไปยังเซิร์ฟเวอร์

ขั้นที่สาม : Information Transfer

เซิร์ฟเวอร์จะไปหาข้อมูลที่ไคลเอนต์ต้องการ

ขั้นที่สี่ : Response

เซิร์ฟเวอร์ส่งข้อมูลตอบสนอง (Response) กลับมายังไคลเอนต์เสมอ

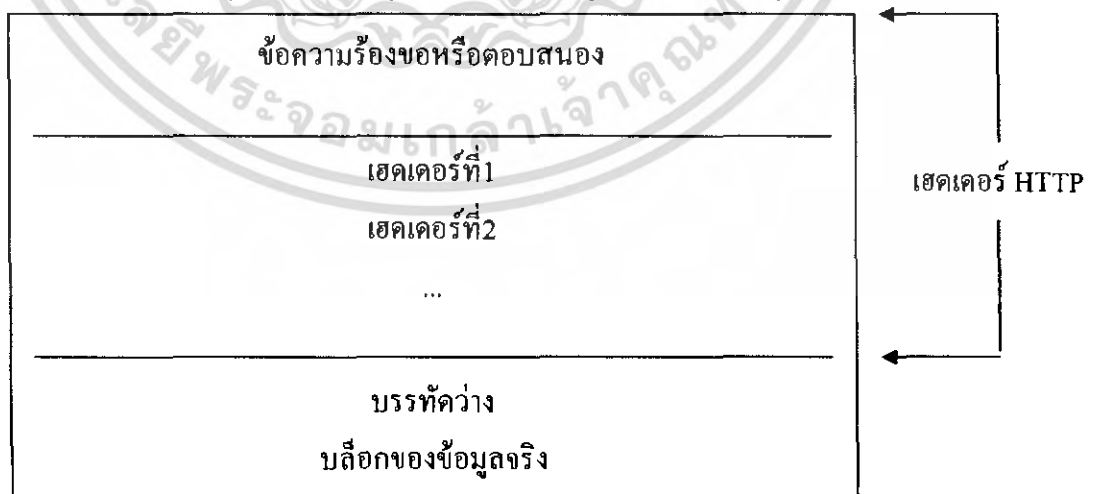
ขั้นสุดท้าย : Close Socket

ปลดการเชื่อมต่อของซ็อกเก็ตของทั้งสองฝั่งออก

ด้วยการทำงานของโพรโตคอล HTTP ที่มีการเชื่อมต่อในระยะเวลาเพียงสั้น หรือที่เรียกว่าเป็นโพรโตคอลแบบ Connectionless ในลักษณะดังกล่าว ทำให้ในช่วงเวลาหนึ่ง ๆ เซิร์ฟเวอร์ที่ให้บริการ WWW สามารถรองรับไคลเอนต์ได้จำนวนมากพร้อม ๆ กัน เพราะไม่มีใครได้ทำการเชื่อมต่ออย่างถาวรในการร้องขอรับบริการจากไคลเอนต์ และการตอบสนองหรือการให้บริการจากเซิร์ฟเวอร์นั้นย่อมต้องมีการรับส่งข้อมูลระหว่างกัน แต่ข้อมูลที่รับส่งกันในแต่ละครั้งไม่ได้มีเฉพาะข้อมูลเพียงอย่างเดียว แต่ละฝ่ายจะมีส่วนเฮดเดอร์ HTTP (HTTP header) เข้าไปในส่วนต้นของข้อมูลที่รับ-ส่งกันด้วย ซึ่งเฮดเดอร์ HTTP จะเป็นตัวบอกว่าข้อมูลที่ส่งมานี้เป็นข้อมูลอะไร เป็นข้อมูลการร้องขอจากไคลเอนต์ หรือเป็นข้อมูลตอบสนองจากเซิร์ฟเวอร์

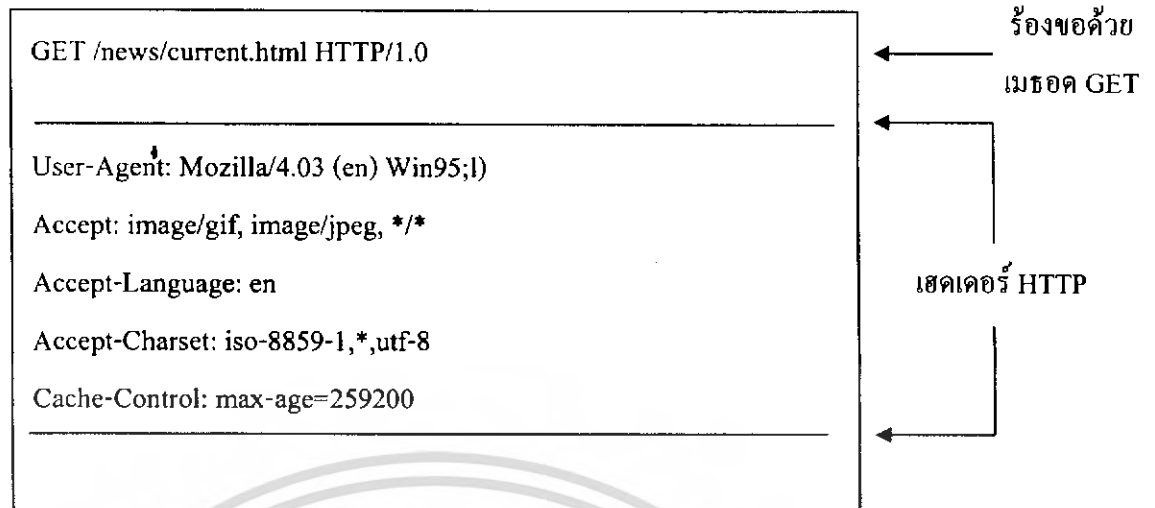
3.1.2 โครงสร้างของโพรโตคอล HTTP

โครงสร้างของข้อมูล HTTP จะแบ่งออกเป็น 2 ส่วนใหญ่ ๆ คือ ส่วนเฮดเดอร์ หรือเรียกว่า Metadata จะเป็นส่วนเก็บข้อมูลที่จำเป็นต้องใช้ภายในโพรโตคอล ส่วนที่สองเป็นส่วนของข้อมูลจริงที่ต้องการรับส่ง ดังแสดงในรูปที่ 3.1 และมีรูปแบบการส่งข้อมูลทั้ง 2 แบบ ดังรูปที่ 3.2, 3.3

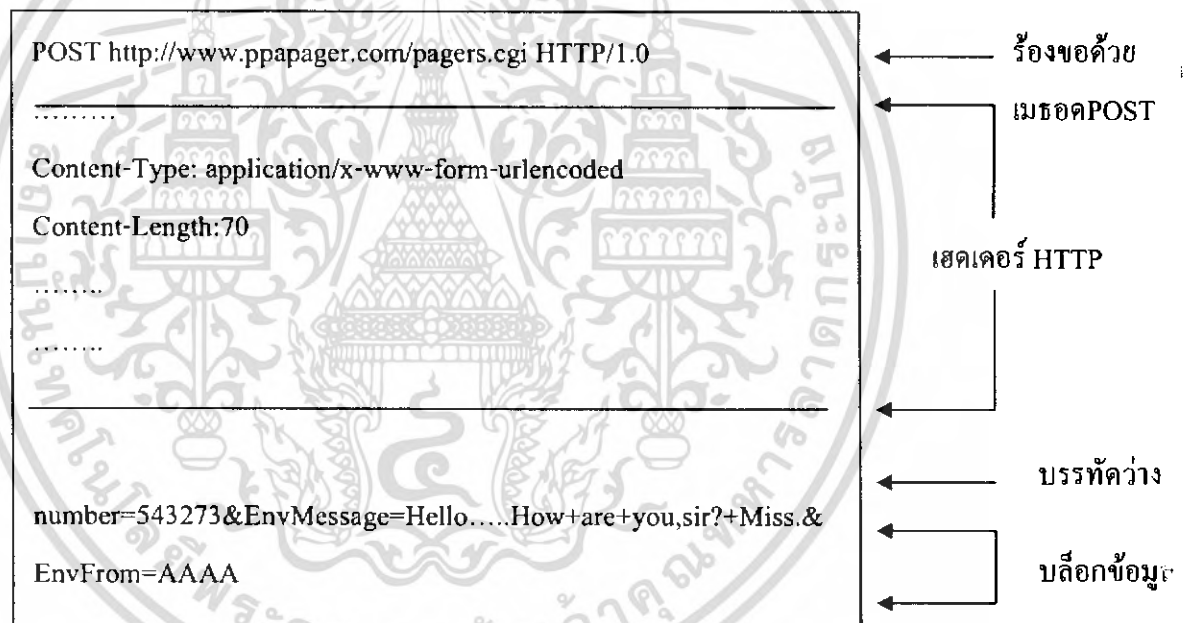


รูปที่ 3.1 โครงสร้างของข้อมูลที่ส่งผ่านโพรโตคอล HTTP

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.2 ตัวอย่างข้อความร้องขอด้วยเมธอด GET



รูปที่ 3.3 ตัวอย่างข้อความร้องขอด้วยเมธอด POST

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.2 พร็อกซีส่วนตัว (Personal Web Proxy)

3.2.1 พร็อกซี (Proxy) คืออะไร

พร็อกซี (Proxy) คือ โปรแกรมที่ทำหน้าที่เป็นได้ทั้ง เซิร์ฟเวอร์ และ โคลแอนด์ สำหรับจุดประสงค์ในการสร้างการร้องขอเพื่อเครื่องโคลแอนด์อื่นๆ การร้องขอจะถูกให้บริการ โดยการส่งเข้าไปภายในพร็อกซี ซึ่งในบางกรณีจะมีการเปลี่ยนไปยังเครื่องเซิร์ฟเวอร์อื่น พร็อกซีจะต้องสามารถจัดการได้ทั้งแบบ เซิร์ฟเวอร์ และ โคลแอนด์ที่ต้องการรายละเอียดดังกล่าว ปัจจุบันได้แบ่งพร็อกซีออกเป็นหลายประเภทแต่สามารถแบ่งเป็นประเภทใหญ่ๆ ได้ 2 แบบ คือ

Transparent Proxy คือ พร็อกซีที่จะไม่ถูกปรับเปลี่ยนการร้องขอหรือการตอบรับเมื่อได้ถูกร้องขอเพื่อการพิสูจน์ยืนยันตัวตนของพร็อกซีและระบุตัวตน (proxy authentication และ identification) ผู้ใช้งานไม่จำเป็นต้องเช็คค่าอะไรในเครื่องคอมพิวเตอร์ของตัวเอง ตัวของ Transparent Proxy นั้นจะใช้หลักของ iptables ในการ redirects ผู้ที่ติดต่อผ่านมาทางพอร์ต http ให้ไปยัง พร็อกซีเซิร์ฟเวอร์ ได้เอง โดยอัตโนมัติ

Non-Transparent Proxy คือ พร็อกซีที่สามารถปรับเปลี่ยนการร้องขอและตอบรับเพื่อให้บริการสำหรับผู้ใช้งาน เช่น ตัวแทนผู้ให้บริการ, การเปลี่ยนรูปของสื่อกลางประเภทต่างๆ, protocol reduction หรือ anonymity filtering ยกเว้นในที่ที่ใช้ Transparent Proxy หรือ Non-Transparent Proxy ที่อยู่บนสถานะที่แน่นอนแล้ว HTTP proxy ต้องการที่จะนำมาใช้ ได้กับทั้งสองชนิดนี้

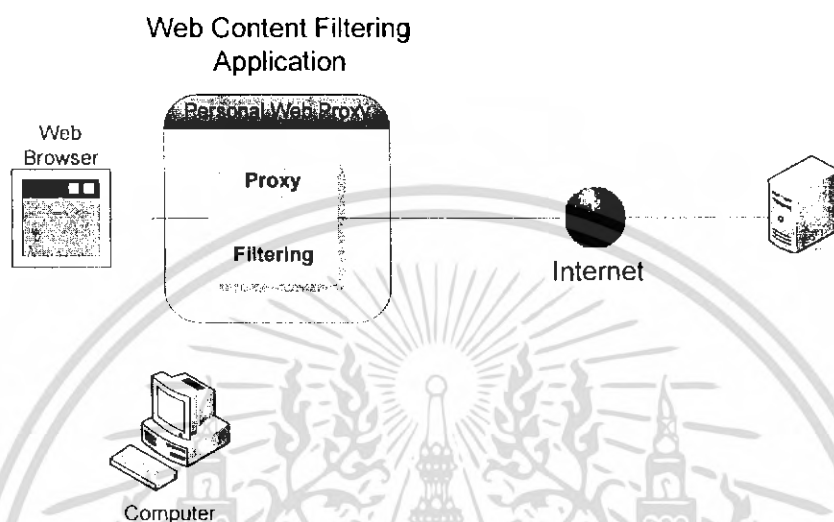
Proxy server คือ คอมพิวเตอร์ที่มีบริการต่างๆ บนระบบเน็ตเวิร์ค อนุญาตให้โคลแอนด์สามารถทำการเชื่อมต่อไปยังบริการเน็ตเวิร์คอื่นๆ ได้ โดยโคลแอนด์จะเชื่อมต่อไปที่ proxy server ก่อน จากนั้นร้องขอการเชื่อมต่อ ไฟล์ หรือทรัพยากรอื่นๆ บนเซิร์ฟเวอร์ต่างๆกัน ซึ่งพร็อกซีจะต้องหาทรัพยากรนั้น ส่งคืนให้กับโคลแอนด์ จากการเชื่อมต่อไปยังเซิร์ฟเวอร์ที่ระบุไว้ หรือคอยให้บริการเองบนแคช (Cache) ที่มีอยู่ แต่บางกรณี พร็อกซีอาจเปลี่ยนการร้องขอของโคลแอนด์ หรือเปลี่ยนการตอบกลับจากทางฝั่งเซิร์ฟเวอร์ ขึ้นอยู่กับจุดประสงค์ที่ตั้งไว้ของการใช้งาน

พร็อกซีที่ทำการคัดกรองข้อมูลเนื้อหาบนเว็บ เรียกว่า เว็บพร็อกซี (Web Proxy) โดยเว็บพร็อกซีจำนวนมากจะเปลี่ยนรูปแบบของเว็บเพจที่จะส่งใหม่ เพื่อจุดประสงค์บางอย่างตามที่ต้องการ ผู้ควบคุมดูแลเน็ตเวิร์คจะนำพร็อกซีมาใช้เพื่อต้องการสกัดกั้นไวรัสคอมพิวเตอร์ หรือเนื้อหาของโปรแกรมที่มีเจตนาร้ายต่อผู้ใช้

ในหลายๆ องค์กร อาทิ เช่น บริษัท โรงเรียน จะใช้ Proxy server บังคับใช้บนเน็ตเวิร์คตามนโยบาย หรือเพื่อต้องการความปลอดภัย ปราศจากโปรแกรมประสงค์ร้าย หรือเพื่อบริการแคชในตัว

3.2.2 ฟร็อกซี่ส่วนตัว (Personal Web Proxy) คืออะไร

ฟร็อกซี่ส่วนตัว (Personal Web Proxy) คือฟร็อกซี่ที่อยู่บนเครื่องของผู้ใช้ เมื่อผู้ใช้จะทำการร้องขอข้อมูลของหน้าเว็บเพจไปที่เว็บเซิร์ฟเวอร์ต่างๆ จะต้องส่งและรับข้อมูลผ่านฟร็อกซี่ที่ตั้งอยู่ที่เครื่องดังรูปที่ 3.4 นั่นคือ ฟร็อกซี่ที่อยู่ในเครื่องนี้ทำหน้าที่เป็น Port Forwarding นั่นเอง



รูปที่ 3.4 รูปแบบของ Personal Web Proxy

3.3 การติดต่อระหว่างโปรเซส (Inter-Process Communication)

การติดต่อระหว่างโปรเซส (Inter-Process Communication หรือ IPC) เป็นเทคนิคต่างๆ ที่ใช้สำหรับแลกเปลี่ยนข้อมูลกันระหว่าง 2 เธรด (Thread) หรือมากกว่าใน 1 โปรเซส (Process) โดยโปรเซสจะถูกรันบนคอมพิวเตอร์ 1 เครื่องหรือหลายเครื่องที่เชื่อมต่อกันก็ได้บนเน็ตเวิร์ค เทคนิค IPC ถูกแบ่งออกตามไปตามหลายวิธีการ เช่น การส่งข้อความหากัน, การส่งสัญญาณให้ทำงานพร้อมกัน, ใช้งานบนหน่วยความจำร่วมกัน และการเรียกโปรซีเจอร์ระยะไกล (Remote Procedure Call หรือ RPC เป็นบริการที่อนุญาตให้โปรแกรมที่ทำงานบนคอมพิวเตอร์เครื่องหนึ่งสามารถเรียกใช้งานโปรแกรมย่อยที่คอมพิวเตอร์อีกเครื่องได้ โดยใช้การส่งข้อมูลไปยังเครื่องที่ต้องการและรับค่าผลลัพธ์ที่ได้กลับมา) วิธีของ IPC ที่ใช้จะขึ้นอยู่กับ bandwidth ความสามารถในการติดต่อสื่อสารกันระหว่างเธรด และชนิดของข้อมูลการส่งหากัน

3.3.1 คลาส (Classes) ของการติดต่อระหว่างโปรเซสที่ใช้

ได้นำมาประยุกต์ใช้กับโปรแกรมเพื่อติดต่อระหว่างในส่วนของ อินเตอร์เฟส (User Interface) กับ Windows Service

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

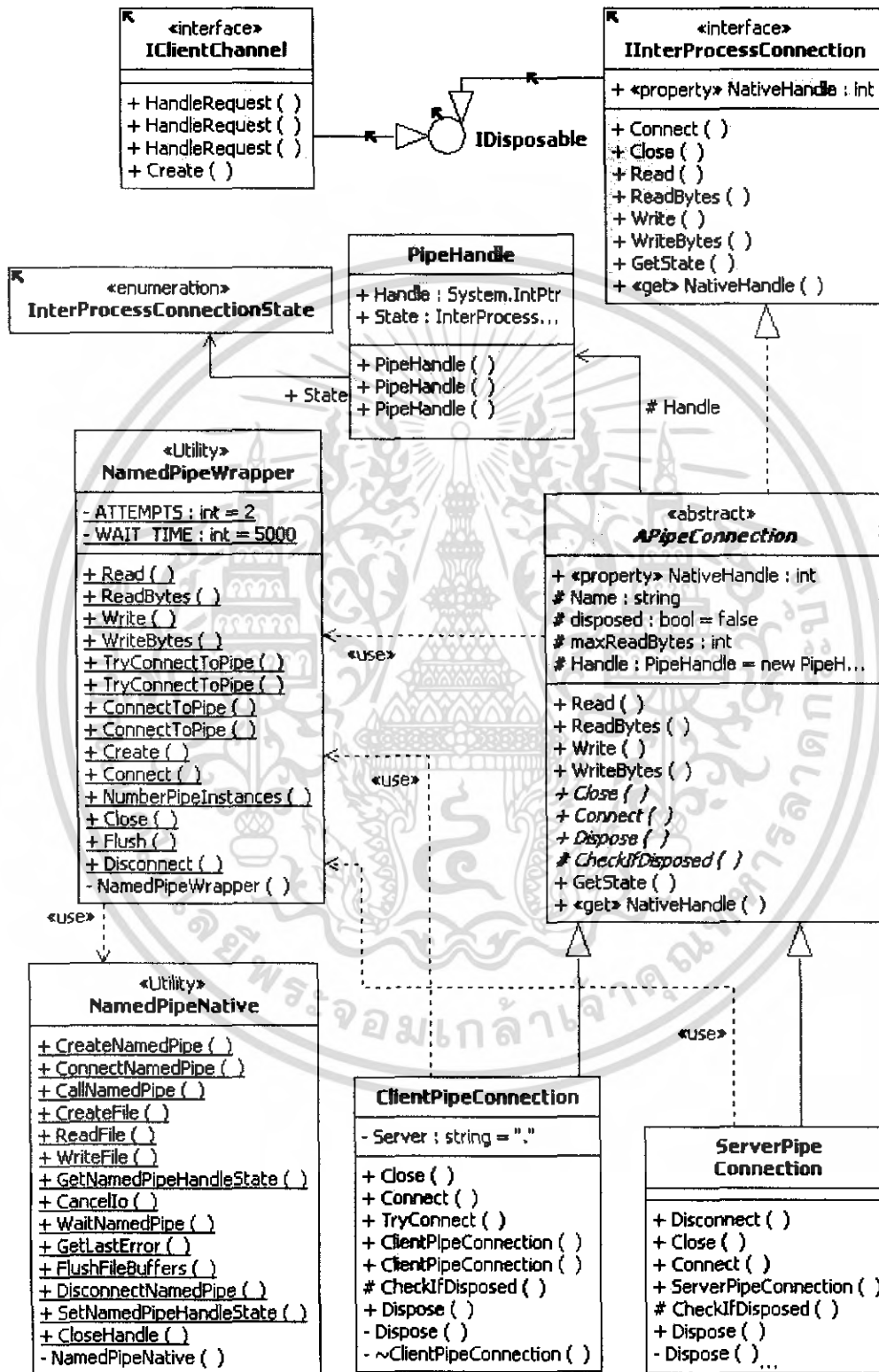
มีคลาส และเมธอดที่อยู่ในส่วนหนึ่งของ Named Pipes Solution ตามไดอะแกรมดังรูปที่ 3.5 จะแสดงคลาส และความสัมพันธ์ดังกล่าว ซึ่งจะเห็นว่าอินเตอร์เฟสอยู่หลายส่วน เช่น IClientChannel และ IinterProcessConnection ซึ่งทั้งหมดได้ถูกคอมไพล์ใน AppModule.InterProcessComm interfaces เหล่านี้ได้ถูกบรรจุอยู่ในabstract ของ Name Pipes ฝั่ง Client ที่เกี่ยวข้องกับ IPC ตามพื้นฐานหลักการของ object-oriented ในเรื่อง “Loose coupling” ฝั่ง Client application จะใช้อินเตอร์เฟส เมื่อทำการแลกเปลี่ยน Messages กับฝั่ง Server ซึ่งอนุญาตให้ปรับเปลี่ยน IPC protocol ถ้าจำเป็น คลาสต่อไปนี้จะเป็คลาสหลักที่รับผิดชอบในส่วนของ .NET Named Piped

- NamedPipeNative: คลาสนี้จะไปเรียกใช้เมธอด ใน kernel32.dll สำหรับ Name Pipes Communication ซึ่งมันจะกำหนดค่าคงที่ของ error code และค่าของ parameter method
- NamedPipeWrapper: คลาสนี้จะเป็นตัวที่ครอบคลุม NamedPipesNative มันจะใช้เมธอด ใน kernel32.dll ไปควบคุม Named Pipes Function
- APipeConnection: เป็น Abstract class ที่มีเมธอด สำหรับการสร้างการเชื่อมต่อ Named Pipes การอ่านและเขียนข้อมูลคลาสนี้จะถูกสืบทอดจากคลาส ClientPipeConnection และ ServerPipeConnection ซึ่งถูกใช้โดย Client และ Server application ตามลำดับ
- ClientPipeConnection: จะถูกใช้โดย Client Application ไปติดต่อกับ Server โดยใช้ Named Pipes
- ServerPipeConnection: อนุญาตให้ Named Pipes server สร้างการเชื่อมต่อและแลกเปลี่ยน ข้อมูลกับ client
- PipeHandle: จะ hold ในส่วนของ Operating system native handle และ สถานะปัจจุบันของการเชื่อมต่อ Pipe

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3.2 Named Pipes UML static diagram

Static diagram ของ Inter Process Communication ที่นำมาใช้จะเป็นดังรูป 3.5



รูปที่ 3.5 คลาสไดอะแกรมของ Inter Process Communication

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

เทคนิคที่ใช้ในการเขียนโปรแกรม (Programming Technique)

4.1 การเขียนโปรแกรมแบบไม่ประสานเวลา (Asynchronous Programming)

การเขียนโปรแกรมแบบ Asynchronous Programming ทำให้การเชื่อมต่อจะไม่ถูกปิดกั้น โดยตัวแอปพลิเคชันระหว่างที่รอคอยกระบวนการทำงานของระบบเครือข่ายเสร็จสมบูรณ์ แทนที่จะทำการปิดกั้นเซิร์ฟเวอร์ที่ทำงานอยู่ปัจจุบัน ซึ่งจะใช้รูปแบบของ .NET asynchronous ในกระบวนการติดต่อในระบบเครือข่ายบน thread อื่นๆ ระหว่างที่โปรแกรมอยู่ในระหว่างการทำงานใน thread แรก เมื่อ thread ที่ 2 ได้ทำงานเสร็จเรียบร้อยแล้ว มันจะส่ง event object ไปที่เซิร์ฟเวอร์ซึ่งเป็นการบอกสถานะของ task นั้นๆ

เหตุการณ์ (Event) สามารถที่จะส่งสัญญาณออกไปในกรณีที่เมธอดนั้นกำลังรอคอยที่จะตอบกลับ หรือไม่มีการส่งสัญญาณตอบกลับ ถ้าเหตุการณ์ไม่ได้ถูกส่งสัญญาณเมธอดที่รอคอยอยู่จะถูกปิดกั้น (block) จนกว่าสัญญาณถูกส่งมา

เราจะมีชนิดของเหตุการณ์อยู่ 2 ชนิด นั่นคือ

- Automatic event มันสามารถทำให้ตัวมันเองกลับไปสถานะไม่ส่งสัญญาณ (non-signaled state) เมื่อไรก็ตามที่มันถูกส่งสัญญาณออกไป
- Manual reset event มันจะส่งสัญญาณจนกระทั่งมันได้ถูกสั่งให้รับการ reset ไปยังสถานะไม่ส่งสัญญาณ (non-signaled state)

.NET จะมี ManualResetEvent class ซึ่งเราสามารถเรียกใช้ในโปรแกรมได้ หากเราต้องการที่จะดำเนินการต่อ ให้เราทำการตั้งค่า ManualResetEvent object ไปที่สถานะส่งสัญญาณ (signaled state) จะทำให้มันหยุดปิดกั้น (Block) ทุกๆเมธอดที่ทำการรอคอยอยู่ และโปรแกรมสามารถดำเนินการต่อได้ เราจะใช้ object นี้เพื่ออำนวยความสะดวกของการเขียนโปรแกรม asynchronous programming ได้

รูปแบบการเขียนโปรแกรมแบบ Asynchronous programming เหมาะสมสำหรับโปรแกรมที่ไม่สามารถรอคอยเพื่อการทำงานบนเครือข่ายได้สมบูรณ์ก่อนที่จะทำการต่อไป

รูปแบบการเขียนโปรแกรมแบบ Asynchronous programming จะใช้เมธอดที่เรียกว่า callback ในการส่งผลของการทำงาน .NET Framework จะมี method ที่จัดการเกี่ยวกับการกำหนดค่าเริ่มต้น แต่ต้องการเมธอด callback ในการจบการทำงาน เช่น เมธอด Connect () ใช้

BeginConnect () ในการเริ่มการติดต่อไปที่อุปกรณ์เครือข่ายและในการทำงานร่วมกับเมธอด callback เพื่อให้เกิดการเชื่อมต่อที่สมบูรณ์

.NET Framework มีคลาสที่ใช้สำหรับการทำงาน Asynchronous นั่นคือ AsyncCallback สำหรับใช้ในโปรแกรมเพื่อที่จะทำให้กระบวนการทำงานแบบ Asynchronous สมบูรณ์

```
AsyncCallback aCallback = new AsyncCallback (AsyncCallback);
```

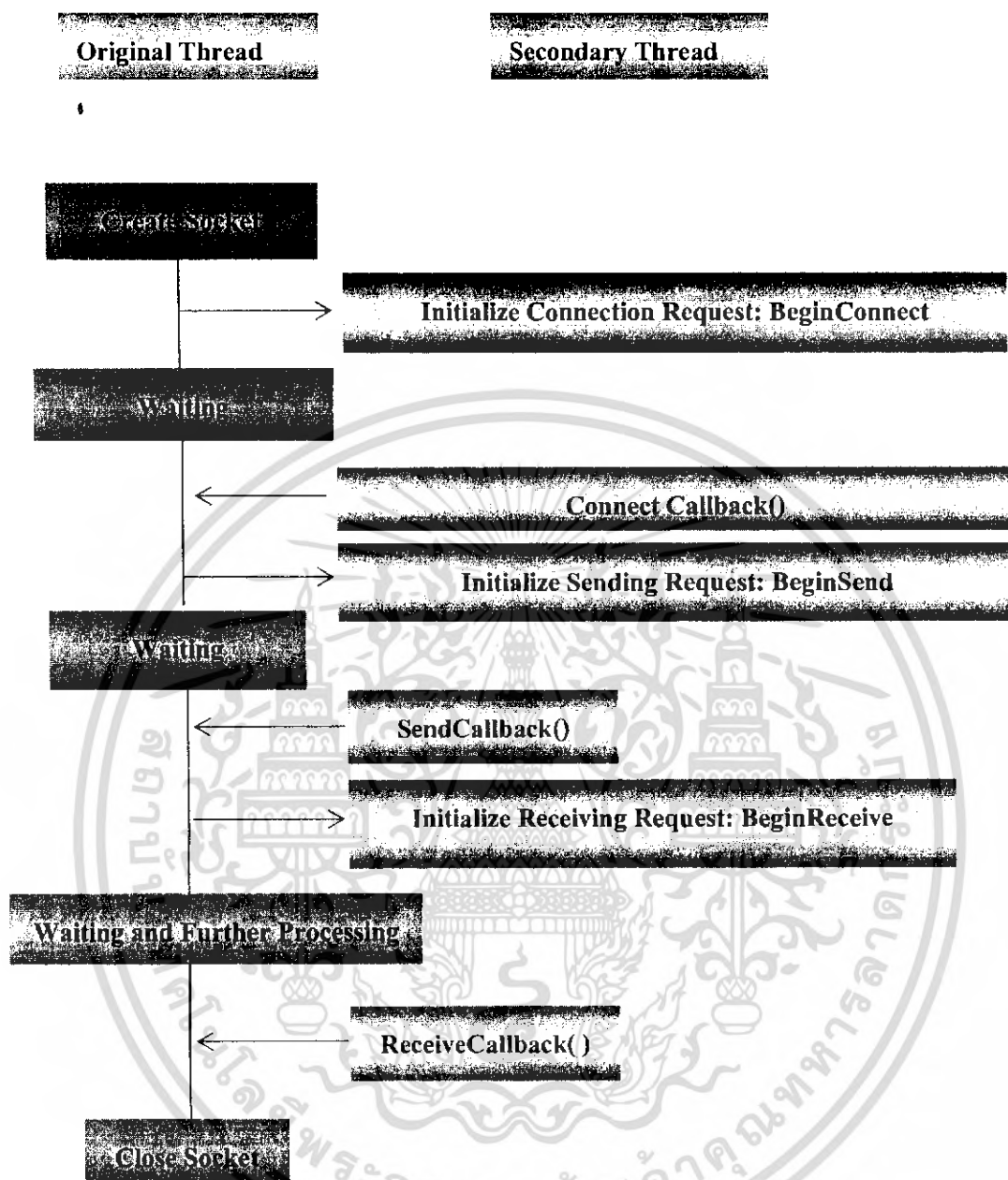
ในรูปแบบนี้จะใช้ Multiple threads ไปยังกระบวนการเชื่อมต่อบนเครือข่าย เพื่อความง่ายในการ synchronization ระหว่าง เธรด ต่างๆ เราจะใช้คลาส ManualResetEvent เพื่อจะระงับการทำงานของคลาสหลักและสัญญาณเมื่อการทำงานสามารถดำเนินการต่อได้

ต่อไปเราจะอธิบายการทำงานเป็นแผนผังการทำงานของทั้งฝั่งไคลเอ็นต์ และฝั่งเซิร์ฟเวอร์



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.1.1 Asynchronous Client Application



รูปที่ 4.1 แผนภาพขั้นตอนการทำงานของไคลเอนต์แอปพลิเคชัน

แผนภาพดังรูปที่ 4.1 แสดงขั้นตอนการทำงานของทางฝั่งไคลเอนต์แอปพลิเคชัน ประกอบด้วย 3 งานที่จะทำให้ Asynchronous สามารถทำงานได้ เป็นการเชื่อมต่อไปยังฝั่งเซิร์ฟเวอร์ ทำการส่งข้อมูล และรับข้อมูลกับทางเซิร์ฟเวอร์

เนื่องจากเราไม่สามารถทำงานกับซ็อกเก็ต หากเราไม่สามารถทำการหยุดการเชื่อมต่อระหว่างเซิร์ฟเวอร์ หรือ Deadlock สามารถที่จะเกิดขึ้นได้หากทำการรับข้อมูลแทนที่จะควรจะต้องส่งข้อมูล เราจะใช้คลาส `ManualEventReset` ที่จะทำให้ Synchronize ต่าง เทรด กันเกิดขึ้นมา โดย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

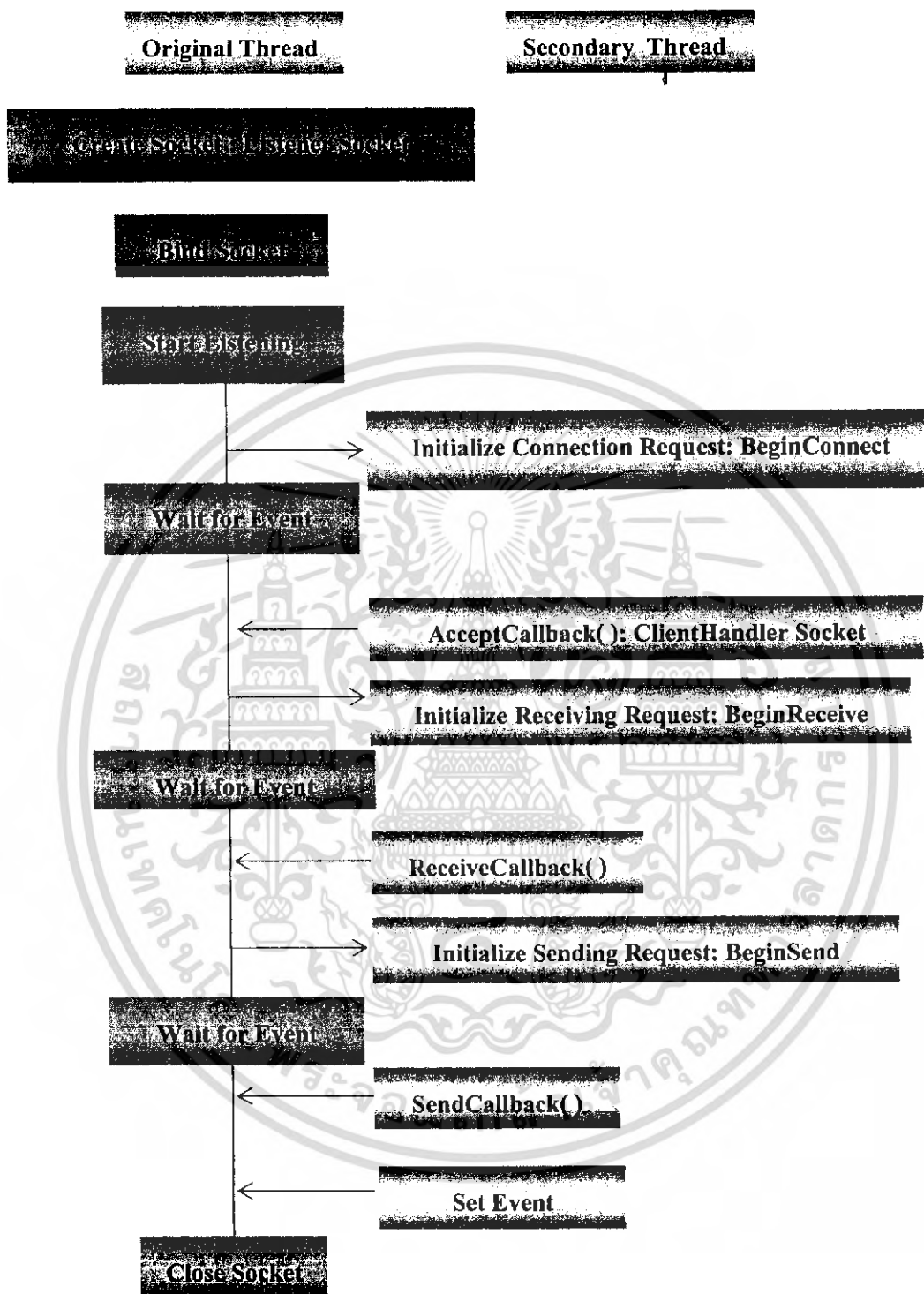
ManualEventReset จะมีสถานะ set และ reset (เมธอด Set (), Reset ()) โดยแต่ละงานประกอบไปด้วย การเชื่อมต่อ การส่ง และการรับข้อมูล ซึ่งสามารถใช้เมธอด WaitOne () ที่จะทำให้ เธรดหลักสามารถรอานที่เกี่ยวข้องอย่างสมบูรณ์ (เมธอดนี้จะปิดกั้นเธรดปัจจุบันจนกว่าที่จะได้ส่งสัญญาณ)

โดยงานแรกของเราคือการเชื่อมต่อไปที่เซิร์ฟเวอร์ ใช้เมธอด BeginConnect () ซึ่งเป็นการเริ่มการเชื่อมต่อให้เป็นการร้องเรียกการเชื่อมต่อแบบ Asynchronous จากนั้นทำการสร้าง AsyncCallback เพื่อให้ไปที่เมธอดที่ต้องการเมื่อเกิดการ Callback และในเมธอดที่ Callback นี้ควรที่จะเรียกเมธอด EndConnect () ในการหยุดการร้องเรียกการเชื่อมต่อ และ return socket ที่ทำการเชื่อมต่อ และสุดท้ายเราจะทำการเรียกเมธอด Set () ของ ManualResetEvent ซึ่งจะเป็นการบอกให้กับเธรดหลักว่าการเชื่อมต่อเสร็จสิ้นแล้ว

ต่อไปเราจะทำการติดต่อสื่อสารกับเซิร์ฟเวอร์ในการส่ง และรับข้อมูล ซึ่งใช้เมธอด BeginSend () ที่จะใช้ในการส่งข้อมูลแบบ Asynchronous ไปยัง ซ็อกเก็ตที่ทำการเชื่อมต่อ ซึ่งเมธอดนี้จะเรียกฟังก์ชันที่ใช้ในการ callback ผ่านยัง AsyncCallback และทำการเรียกเมธอด EndSend () เมื่อหยุดการร้องเรียกการส่งแบบ Asynchronous จากนั้นทำการ Set () ของ ManualResetEvent

เมธอด BeginReceive () ทำการเริ่มการรับข้อมูลแบบ Asynchronous จาก ซ็อกเก็ตซึ่งเมื่อเกิดการ Callback เมธอด BeginReceive () จะทำการคืนค่า (return) จำนวนของไบต์ ดังนั้นเราสามารถที่จะทำการเช็คได้ว่าหากมีบางข้อมูลส่งกลับมาในคิว (queue)

4.1.2 Asynchronous Server Application



รูปที่ 4.2 แผนภาพขั้นตอนการทำงานของเซิร์ฟเวอร์แอปพลิเคชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากรูป 4.2 การทำงานจะเริ่มจากการสร้างซ็อกเก็ต และจากนั้นจะทำการรับฟังPort ที่จะมาเกี่ยวข้องกับคีย์ โปรแกรม Asynchronous Server ต้องเรียก BeginAccept () เพื่อจะรับการติดต่อ และเมื่อได้รับการเชื่อมต่อเรียบร้อยแล้ว มันจะไปเรียกฟังก์ชัน BeginReceive () และ BeginSend () เพื่อส่งข้อมูลและรับข้อมูลจากซ็อกเก็ตของไคลเอ็นต์

ภายหลัง BeginAccept () ได้ถูกเรียกแล้วเราจะตั้งสถานะให้อยู่ในสถานะการคอย (waiting state) ดังนั้นมันจะแตกต่างจากเรดในโปรแกรมจะยังคงทำงานต่อไป เมื่อเครื่องไคลเอ็นต์พยายามที่ติดต่อฝั่งเซิร์ฟเวอร์ ซึ่งถ้าเราไม่ทำเช่นนี้โปรแกรมจะจบการทำงานก่อนเครื่องไคลเอ็นต์ ซึ่ง ManualResetEvent จะถูกใช้เพื่อที่จะตั้งสถานะในการคอย

ฟังก์ชันBeginAccept () จะต้องใช้2ตัวแปรคือตัวแรกคือตัวแปรของAsynchronous (เพื่อการอ้างอิงถึงฟังก์ชัน callback) และอีกตัวแปรหนึ่งคือตัวแปรที่จะถูกใช้ส่งข้อมูลไปที่ฟังก์ชัน callback ฟังก์ชัน BeginAccept () จะไปเรียกฟังก์ชัน AcceptCallback () เมื่อการเชื่อมต่อได้ถูกเชื่อมต่อกับซ็อกเก็ต ตัวฟังก์ชัน Callback นี้จะ ไปเรียกฟังก์ชัน EndAccept () ในการส่งซ็อกเก็ตใหม่กลับไปทางไคลเอ็นต์

การเขียนโปรแกรมแบบ Asynchronous socket จะใช้ threadsสำหรับการทำงานต่าง asynchronous กัน thread หลักจะมีไว้สำหรับกำหนดค่าเริ่มต้นของ listener socket thread ที่2 มีไว้สำหรับรับตอบรับเรดที่ได้ผ่านเข้ามาเรดสุดท้ายจะใช้สำหรับรับและส่งข้อมูล เราสามารถร้องขอค่า thread ID จากเรด โดยใช้ฟังก์ชัน GetCurrentThreadID () หลังจากที่ได้เกิดการเชื่อมต่อโดยการเรียก EndAccept () ซ็อกเก็ตใหม่จะสามารถสื่อสารกับไคลเอ็นต์ โดยการเรียกใช้ฟังก์ชัน BeginReceive () และฟังก์ชัน BeginSend () ฟังก์ชัน ReceiveCallback () จะไปเรียกฟังก์ชัน EndReceive () เพื่อที่จะทำให้การทำงานแบบAsynchronousเสร็จสมบูรณ์ โดยจะมี BeginSend () ไว้ส่งข้อมูลกลับไปยังclient ไม่เช่นนั้นเราจะต้องเรียก BeginReceive () เพื่อที่จะรับข้อมูลที่ออกมา

เมื่อฟังก์ชัน SendCallback () ทำงานเสร็จเรียบร้อยแล้วโดยการเรียก EndSend () จากนั้นจะไปปิดตัวที่ควบคุมซ็อกเก็ต และไปตั้งค่า ManualResetEvent บนเรดหลัก

4.2 การสร้าง และการติดตั้งวินโดว์เซอร์วิสแอปพลิเคชัน (Windows Service Application) โดย Visual Studio.NET

วินโดว์เซอร์วิสแอปพลิเคชัน (Windows Service Application) ได้กลายเป็นส่วนหนึ่งของมาตรฐานของรูปแบบโปรเจกต์พื้นฐาน ซึ่งประโยชน์ของวินโดว์เซอร์วิส (Windows Service) คือวินโดว์เซอร์วิสสามารถเริ่มการทำงานของโปรแกรมก่อนที่ผู้ใช้ (User) คนใดคนหนึ่งจะเข้าสู่ระบบ (หากมีการปรับตั้งค่าให้เริ่มโปรแกรมเมื่อก่อนวินโดว์ทำการ Boot Up) และประโยชน์อีกอย่างหนึ่งคือสามารถรันโปรแกรมได้ตลอดเวลา ตรงกันข้ามที่วินโดว์ยังทำการรันอยู่ โดยผู้ใช้ทั่วไปที่ไม่ใช่ admin จะไม่ทำการปิดเซอร์วิส หรือ โพรเซสของโปรแกรมนี้อีก

4.2.1 เขียนโปรแกรมวินโดว์เซอร์วิส (Coding Windows Service)

ขั้นตอนแรกของการสร้างวินโดว์เซอร์วิส (ในที่นี้ยกตัวอย่างใช้ C#.NET) ให้เลือกเมนู New Project จากนั้นเลือกชนิด (Type) ของโปรเจกต์เป็น Windows Service

หลังจากสร้างโปรเจกต์เป็นวินโดว์เซอร์วิสใหม่แล้ว คลาส (Class) ที่ใช้จะถูกสืบทอด (Inheritance) มาจากคลาส System.ServiceProcess.ServiceBase ซึ่งถูกสร้างมาให้แล้ว พร้อมด้วย 4 เมธอด (Method) พื้นฐานได้แก่

```
protected override void OnStart(String[] args)
protected override void OnStop()
protected override void OnPause()
protected override void OnContinue()
```

4.2.2 การเพิ่มคลาส Installer ในโปรแกรม

วินโดว์เซอร์วิสแอปพลิเคชัน ต่างจากวินโดว์แอปพลิเคชัน ที่จะต้องถูกติดตั้งโดยใช้ installUtil.exe ในคลาสใหม่นี้ ให้สร้าง instance ของคลาส System.ServiceProcess.ServiceProcessInstaller และอีก 1 instance ของคลาส System.ServiceProcess.ServiceInstaller สำหรับแต่ละวินโดว์เซอร์วิสในโปรเจกต์ โดยแต่ละออบเจกต์ของ ServiceInstaller ถูกสร้างความสัมพันธ์กับ 1 วินโดว์เซอร์วิส พิจารณาตามคุณสมบัติ (Property) ServiceName ส่วนคุณสมบัติอื่นสามารถที่จะถูกกำหนดค่าตามลักษณะของเซอร์วิสตามต้องการ เช่น โหมด start (Automatic หรือ Manual)

4.2.3 การใช้เครื่องมือรันเพื่อช่วยในการติดตั้งวินโดวส์เซอร์วิส

ตอนนี้ผู้พัฒนาสามารถคอมไพล์โปรเจกต์ไปเป็น ตัวโปรแกรมที่สามารถรันได้ และติดตั้ง เซอร์วิสบนเครื่องคอมพิวเตอร์ โดยใช้เครื่องมือของ .NET installUtil.exe ซึ่งคำสั่ง Command line ที่ต้องใช้มีดังนี้

`installUtil MyWinService.exe` สำหรับการติดตั้งวินโดวส์เซอร์วิสแอปพลิเคชัน

`installUtil /u MyWinService.exe` สำหรับการยกเลิกการติดตั้งของเซอร์วิส

4.2.4 การรวมส่วนประกอบของตัวจับเวลา (Timer Component) กับวินโดวส์เซอร์วิส

เนื่องมาจากตัวจับเวลาเป็นสิ่งที่ต้องการในกำหนดการของงานเรา ซึ่งมี 3 ทางในการสร้าง ตัวจับเวลาใน .NET Framework:

System.Timers.Timer class

System.Threading.Timer class

System.Windows.Forms.Timer class

คลาสที่ 3 เป็นตัวควบคุมสำหรับแอปพลิเคชันกับ interface ดังนั้นในการใช้งานไม่เหมาะกับเซอร์วิส ส่วนในคลาสที่ 2 ทำงานกับ TimerCallBack และ เธรด (Thread) ในการควบคุมตัวจับเวลา ดังนั้นในคลาสที่ 1 ได้ถูกนำมาใช้เกี่ยวข้องกับเมธอด และคุณสมบัติ ที่จะควบคุมการทำงานโดยตรง และเพิ่มเหตุการณ์ (Event) โดยเรียก Timer.Elapsed เมื่อถึงช่วงเวลาในรอบซึ่งถูกตั้งไว้ เช่นตัวอย่าง โค้ดดังรูปที่ 4.3

```
private void aTimer_Elapsed(object sender, ElapsedEventArgs e)
{
    if (DateTime.Now.Second == 50)
    {
        //MessageBox.Show("Now: second is 50");
        count++;
        //Create File on c:\
        StreamWriter file = new StreamWriter("c:\\Service" + count + ".txt");
        //Open File or Program: targetFile
        string targetFile = @"c:\\Service" + count + ".txt";
        System.Diagnostics.Process.Start(targetFile);
    }
}
```

รูปที่ 4.3 ตัวอย่างส่วนของโค้ดของเมธอด เมื่อได้รับเหตุการณ์ครบรอบ interval ของตัวจับเวลา

สมมติว่ามีกรตั้งค่า interval = 1000 (1000 millisecond:1 second) หมายความว่าทุกๆรอบ 1 วินาที จะเกิดเหตุการณ์ aTimer_Elapsed ดังตัวอย่างนี้ จะทำการเช็คค่าวินาทีปัจจุบันมีค่าเท่ากับ 50 หรือไม่ ถ้าเท่าจะทำการสร้างไฟล์ชื่อ Service0.txt ไว้ที่ C:\ ในครั้งแรก (เนื่องจากค่าเริ่มต้นของ count ถูกตั้ง เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ไว้เท่ากับ 0 และจะบวก 1 ทุกๆ 50 วินาที) จากนั้นจะเปิดไฟล์ Service0.txt ที่สร้างนี้ด้วยโปรแกรม notepad โปรแกรมจะทำงานแบบนี้ไปเรื่อยๆ ในทุกๆ 50 วินาที

โดยขอบเขตของ Timer สามารถตั้งค่า และผูกกับเมธอดของวิน โคว์เซอร์วิสได้ดังนี้

คุณสมบัติ Interval ถูกใช้ในการตั้งค่าเป็นช่วงเวลา ในหน่วย millisecond ที่จะเรียกเหตุการณ์ Elapsed

เมธอด Start() ตั้งค่าคุณสมบัติ Enabled ของ timer ให้เท่ากับ true คืออนุญาตให้ตัวจับเวลา timer เริ่มทำการนับได้เพื่อส่งเหตุการณ์ event

เมธอด Stop() ตั้งค่าคุณสมบัติ Enabled ของ timer ให้เท่ากับ false มีผลให้ตัวจับเวลา timer หยุดส่งเหตุการณ์เมื่อครบช่วงเวลา

4.2.5 การทดลองใช้โปรแกรม Personal Web Proxy เมื่อเป็นวินโดว์เซอร์วิส

เมื่อเราได้ลองเขียนโปรแกรมให้ทำตัวเป็นวิน โคว์เซอร์วิสแล้ว (จากหัวข้อย่อ 4.2.1 ถึง 4.2.4 เรื่องการสร้างและติดตั้งวิน โคว์เซอร์วิส) ต่อ ไปจะนำ โปรแกรม Personal Web Proxy ที่ได้ทำการพัฒนาเบื้องต้น จากวินโดว์แอปพลิเคชัน (Windows Application) เปลี่ยนเป็นวินโดว์เซอร์วิส (Windows Service) จากนั้นจึงทดลองใช้งานว่าจะทำงานได้ปกติเหมือนกับก่อนการเปลี่ยนเป็นวินโดว์เซอร์วิสหรือไม่ โดยมีขั้นตอนการทดลองดังนี้

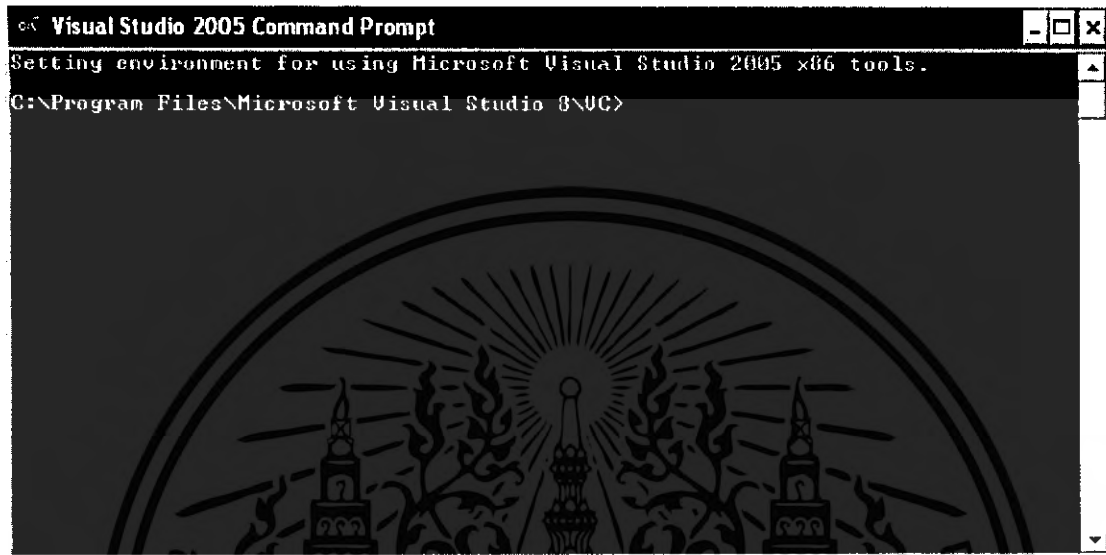
นำโปรแกรมวินโดว์เซอร์วิสเบื้องต้นที่ได้ลองเขียนขึ้นมาข้างต้น มารวมกับตัวโปรแกรม Personal Web Proxy เพื่อให้โปรแกรมทำงานเป็นวินโดว์เซอร์วิส โดยเริ่มจากนำไฟล์คลาสทั้งหมด (Form1.cs, Form1.resx, Form1.Designer.cs) ที่จำเป็นต้องใช้ในการรันของโปรแกรมพรีอ็อกซ์มาไว้ที่ namespace เดียวกับโปรแกรมวินโดว์เซอร์วิส ทำการเพิ่มโค้ดในเมธอด OnStart (string [] args) ของดังรูปที่ 4.4 สั่งให้โปรแกรมวินโดว์เซอร์วิสสร้าง instance ของคลาสพรีอ็อกซ์ (Class Form1) เพื่อใช้ในการรันตัวพรีอ็อกซ์เมื่อวินโดว์เซอร์วิสทำการเริ่ม (Start) ทำงาน

```
protected override void OnStart(string[] args)
{
    // TODO: Add code here to start your service.
    aTimer.Enabled = true;
    Application.EnableVisualStyles();
    Application.SetCompatibleTextRenderingDefault(false);
    Application.Run(new Form1());
}
```

รูปที่ 4.4 โค้ดที่เพิ่มเพื่อสั่งให้โปรแกรม Personal Web Proxy ทำงาน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากนั้นทำการ built project ของวินโดวเซอร์วิสทำให้ได้ไฟล์ที่ใช้ในการรันโปรแกรมวินโดวเซอร์วิส (WindowsService1.exe) ทำการติดตั้งโปรแกรมวินโดวเซอร์วิส (WindowService1.exe) โดยใช้เครื่องมือ installUtil.exe ของ .NET เริ่มจากเข้าไปที่ Visual Studio 2005 Command Prompt ดังรูปที่ 4.5



รูปที่ 4.5 Command Prompt เมื่อเข้ามาที่ Visual Studio 2005 Command Prompt

รัน InstallUtil.exe เพื่อทำการติดตั้ง โปรแกรมวินโดวเซอร์วิส โดยใช้คำสั่งดังนี้
 >Installutil yourproject.exe {path ของ โปรแกรมวินโดวเซอร์วิสที่ต้องการติดตั้ง}
 เช่น InstallUtil C:\Debug\WindowService1.exe ดังรูปที่ 4.6 และได้ผลลัพธ์ดังรูปที่ 4.7 หากทำการติดตั้งโปรแกรมวินโดวเซอร์วิสสำเร็จ
 หมายเหตุ: หากต้องการยกเลิกการติดตั้งวินโดวเซอร์วิสนี้ ให้ใช้คำสั่ง
 InstallUtil /u C:\Debug\WindowService1.exe

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

Visual Studio 2005 Command Prompt
Setting environment for using Microsoft Visual Studio 2005 x86 tools.
C:\Program Files\Microsoft Visual Studio 8\VC>InstallUtil c:\Debug\WindowsService1.exe

```

รูปที่ 4.6 การใช้คำสั่งในการติดตั้งโปรแกรมวินโดวเซอร์วิส

```

Visual Studio 2005 Command Prompt
Setting environment for using Microsoft Visual Studio 2005 x86 tools.
C:\Program Files\Microsoft Visual Studio 8\VC>InstallUtil c:\Debug\WindowsService1.exe
Microsoft (R) .NET Framework Installation utility Version 2.0.50727.42
Copyright (c) Microsoft Corporation. All rights reserved.

Running a transacted installation.

Beginning the Install phase of the installation.
See the contents of the log file for the c:\Debug\WindowsService1.exe assembly's progress.
The file is located at c:\Debug\WindowsService1.InstallLog.
Installing assembly 'c:\Debug\WindowsService1.exe'.
Affected parameters are:
  logtoconsole =
  assemblypath = c:\Debug\WindowsService1.exe
  logfile = c:\Debug\WindowsService1.InstallLog
Installing service Service1...
Service Service1 has been successfully installed.
Creating Eventlog source Service1 in log Application...

The Install phase completed successfully, and the Commit phase is beginning.
See the contents of the log file for the c:\Debug\WindowsService1.exe assembly's progress.
The file is located at c:\Debug\WindowsService1.InstallLog.
Committing assembly 'c:\Debug\WindowsService1.exe'.
Affected parameters are:
  logtoconsole =
  assemblypath = c:\Debug\WindowsService1.exe
  logfile = c:\Debug\WindowsService1.InstallLog

The Commit phase completed successfully.
The transacted install has completed.

```

รูปที่ 4.7 ผลลัพธ์เมื่อทำการติดตั้งโปรแกรมวินโดวเซอร์วิสสำเร็จ

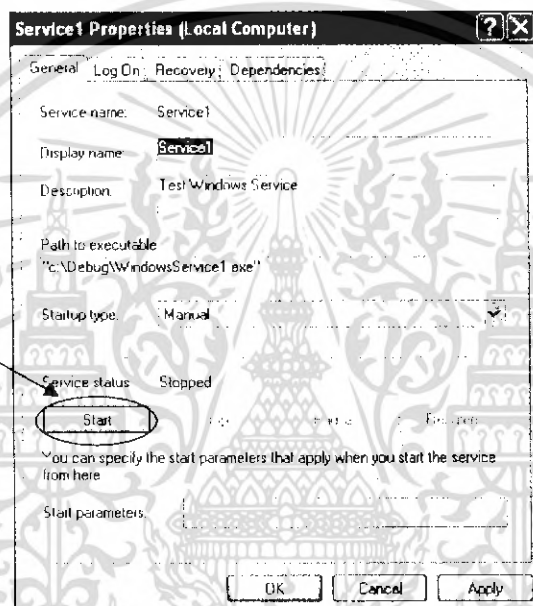
เมื่อเข้าไปที่ Service Control Manager (โดยการกด Run... จากนั้นพิมพ์ services.msc) จะเห็นได้ว่าโปรแกรมวินโดวเซอร์วิสที่เราได้ทำการติดตั้งเรียบร้อยแล้วดังรูปที่ 4.8 (ชื่อของวินโดวเซอร์วิสที่สร้าง คือ Service1 และตั้ง StartUp Type เริ่มต้นเป็น Automatic)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Service Name	Description	Status	Startup Type	Log On As
Shell Hardware Det...	Provides n...	Started	Automatic	Local System
Service1	Test Windo...		Automatic	Local System
Server	Supports fil...	Started	Automatic	Local System
Security Center	Monitors s...	Started	Automatic	Local System

รูปที่ 4.8 วินโดว์เซอร์วิสที่ถูกติดตั้งเรียบร้อยแล้ว

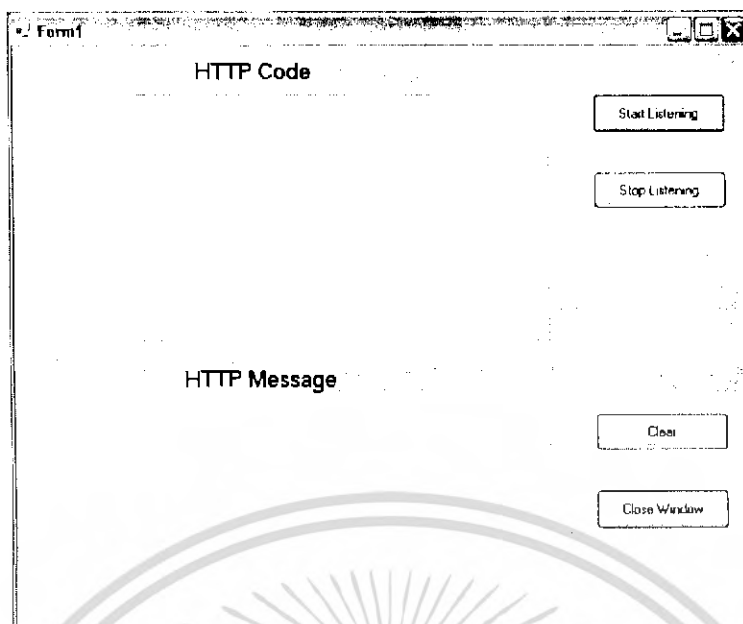
สามารถทำการเริ่ม (Started) ให้วินโดว์เซอร์วิสทำงานได้ 2 วิธี คือ 1. ทำการ Restart เครื่องคอมพิวเตอร์ หรือ 2. เข้าไปกดปุ่ม Start เองบน Properties ของวินโดว์เซอร์วิสใน Services Control Manager ด้วยตัวเองดังรูปที่ 4.9



รูปที่ 4.9 กดปุ่ม Start ใน Properties เพื่อเรียกให้วินโดว์เซอร์วิสตัวนี้รัน

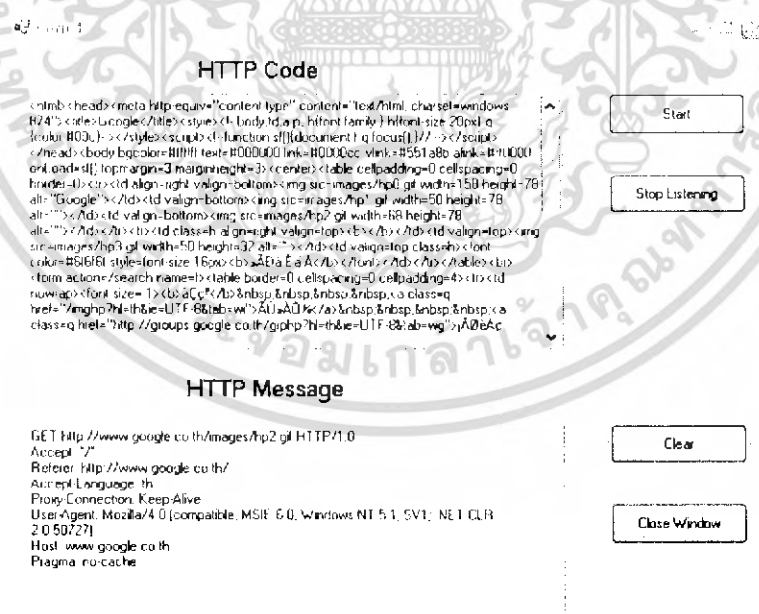
เมื่อวินโดว์เซอร์วิสเริ่มทำงานแล้ว จะปรากฏหน้าต่างเหมือนโปรแกรม Personal Web Proxy (Windows Application) ขณะเริ่มรันโปรแกรมดังรูปที่ 4.10

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



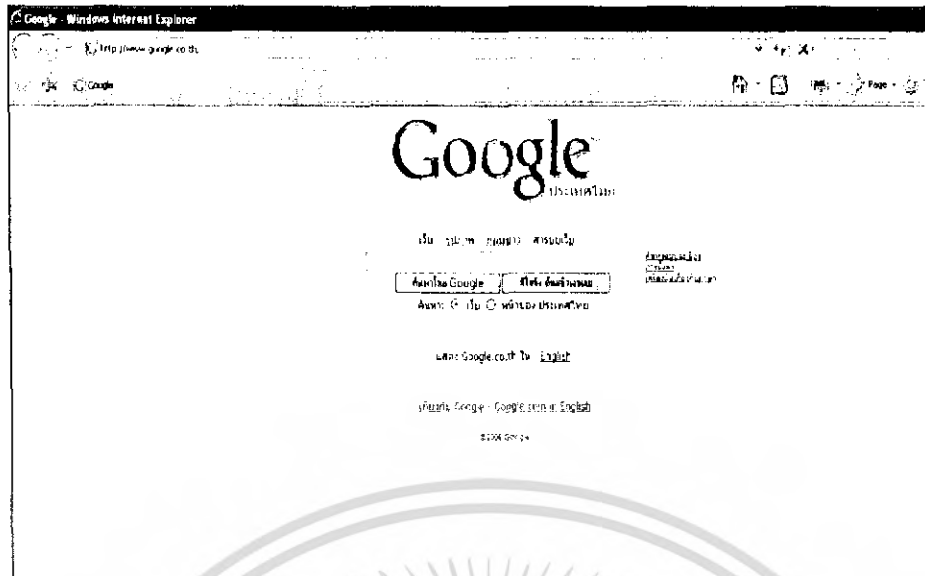
รูปที่ 4.10 หน้าต่างควบคุมโปรแกรม เมื่อเริ่ม (Started) วินโดว์เซอร์วิสชื่อ Service1

เมื่อได้ทดลองใช้งานพร้อมๆ พบว่าได้ผลออกมาถูกต้องเหมือนกับ โปรแกรม Personal Web Proxy ก่อนที่จะถูกนำมาทำเป็นวินโดว์เซอร์วิสดังรูปที่ 4.11 และส่งคืนข้อมูลมาที่เว็บเบราว์เซอร์ถูกต้องตามที่ได้อธิบายจากเว็บเซิร์ฟเวอร์ดังรูปที่ 4.12



รูปที่ 4.11 HTTP Message ที่ผ่านโปรแกรมเพื่อส่งคืนข้อมูลจากเว็บเซิร์ฟเวอร์ ให้กับเว็บเบราว์เซอร์ ตามการร้องขอ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.12 หน้าเว็บเบราว์เซอร์ที่ได้รับข้อมูลส่งกลับจากการร้องขอผ่านโปรแกรม

4.3 นิพจน์ปรกติ (Regular Expression)

นิพจน์ปรกติ (Regular Expression) หมายถึง รูปแบบของลำดับ หรือ กลุ่มของสัญลักษณ์ที่ใช้แทนลำดับ หรือกลุ่มของอักขระตามที่ต้องการ โดยนิพจน์ปรกติมักถูกใช้ในเครื่องมือที่ใช้เขียนโปรแกรม มีประโยชน์ช่วยในการค้นหา และจัดการเกี่ยวกับข้อความ กลุ่มอักขระตามที่ต้องการ ซึ่งขึ้นอยู่กับรูปแบบของนิพจน์ปรกติที่เรากำหนดไว้

ภาษาที่ใช้ในการเขียนโปรแกรมส่วนมากจะสนับสนุนการใช้นิพจน์ปรกติไว้จัดการกับข้อความ หรือที่เรียกว่าสตริง ยกตัวอย่างเช่น ภาษาเพิร์ล (Perl) ซึ่งหลักการของนิพจน์ปรกติเป็นที่แพร่หลายครั้งแรกบนระบบปฏิบัติการยูนิกซ์ (Unix)

4.3.1 หลักการพื้นฐาน

นิพจน์ปรกติมักจะถูกเรียกว่า รูปแบบ (Pattern) เป็นการแสดงเครื่องหมายที่อธิบาย หรือแทนถึงกลุ่มของอักขระ โดยไม่ต้องมีการเก็บแต่ละตัวอักขระทั้งหมดไว้ ยกตัวอย่างเช่น Billy, Belly และ Berlly สามารถถูกแทนได้โดยรูปแบบ "B(i|e?)lly"

การเลือกอย่างใดอย่างหนึ่ง จะใช้เครื่องหมาย "|" (Vertical bar) เช่น "gray|grey" ซึ่งสามารถจับคู่ได้ทั้ง gray และ grey

เครื่องหมายวงเล็บ "(") ใช้ในการกำหนดขอบเขต รวมกลุ่มของตัวอักขระเข้าไว้ด้วยกัน หรือลำดับความสำคัญของเครื่องหมายได้ ยกตัวอย่างเช่น "gr(a|e)y" สามารถจับคู่ได้ทั้ง gray และ grey

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวบ่งบอกถึงจำนวนของตัวอักษร สามารถกำหนดได้ โดยวางไว้อยู่หลังตัวอักษรหรือกลุ่มของอักษรที่ต้องการ ซึ่งมีเครื่องหมายดังนี้

- ? (Question mark) หมายถึง ตัวอักษรตัวหน้าสามารถมีได้ 0(ไม่มี) หรือ 1 ตัวเท่านั้น เช่น “colou?r” แทนได้ทั้ง colour และ color
- * (Asterisk) หมายถึง ตัวอักษรตัวหน้าสามารถมีได้ 0(ไม่มี), 1 หรือมากกว่า 1 ตัวก็ได้ เช่น “go*gle” แทนได้ทั้ง ggle, gogle, google, gooogle และ อื่นๆ
- + (Plus sign) หมายถึง ตัวอักษรตัวหน้าจะต้องมีอย่างน้อย 1 ตัวขึ้นไป เช่น “go+gle” แทนได้ทั้ง gogle, google, gooogle และ อื่นๆ (แต่ไม่สามารถแทน ggle ได้)

โดยโครงสร้างเหล่านี้สามารถถูกนำมาใช้รวมกันได้ เช่น “((great)*grand)?((fajmo)ther)” แทนได้ทั้ง father, mother, grand father, grand mother, great grand father, great grand mother, great great grand father, great great grand mother, great great great grand father, great great great grand mother เป็นต้น

4.3.2 สัญลักษณ์ที่ใช้ในนิพจน์ปกติ

. ใช้แทนตัวอักษรอะไรก็ได้ 1 ตัว (ยกเว้นตัวที่บ่งบอกถึงการขึ้นบรรทัดใหม่ “\n”) เช่น “a.cd” แทน abcd และ “a.d” แทน abcd

\ ใช้ใส่หน้าตัวอักษรพิเศษ หรือสัญลักษณ์ต่างๆ เพื่อบอกให้รู้ว่าเครื่องหมายพิเศษที่ใช้ไม่ได้เป็นส่วนหนึ่งของไวยากรณ์ หรือเรียกอีกอย่างหนึ่งว่า escaping ยกตัวอย่างเช่น “a|b” แทนได้กับตัวอักษร a หรือ |b จะเห็นได้ว่าตัวอักษรพิเศษ “\” ที่ตามหลังสัญลักษณ์นี้จะไม่มี ความหมายพิเศษ มองเป็นตัวอักษรธรรมดา ในที่นี้คือ ตัว “|” (ตัวอักษร หรือสัญลักษณ์พิเศษที่ต้อง escape มีดังนี้ ^, \$, (,), ., [,], *, ?, +, \, และ {)

[] ใช้ระบุตำแหน่งในคำว่า ในตำแหน่งนี้จะประกอบไปด้วยตัวอักษรอะไรได้บ้าง เช่น [abc] แทน “a”, “b” และ “c”, [a-z] แทนตัวพยัญชนะภาษาอังกฤษตัวเล็กได้ทั้งหมด, [abcv-z] แทน “a”, “b”, “c”, “v”, “w”, “x”, “y” และ “z” เป็นต้น

[^] ใช้ระบุตำแหน่งในคำว่า ในตำแหน่งนี้จะต้องไม่ปรากฏตัวอักษรใด เช่น [^abc] แทนตัวอักษรใดก็ได้ที่ไม่ใช่ “a”, “b” และ “c”, [^a-z] แทนตัวอักษรใดก็ได้ที่ไม่ใช่ตัวพยัญชนะภาษาอังกฤษตัวเล็ก เป็นต้น

^ ใช้ระบุตำแหน่งเริ่มต้นของข้อความ เช่น “^big” แทน “big dog”, “big cat” และอื่นๆ ที่มีคำว่า big เป็นคำเริ่มต้นของข้อความ

\$ ใช้ระบุตำแหน่งสุดท้ายของข้อความ เช่น “dog\$” แทน “big dog”, “little dog” และอื่นๆ

ที่มีคำว่า dog ลงท้ายสุดของข้อความ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

{ } ใช้แทนจำนวนครั้งที่ซ้ำกันได้ เช่น “ab{3}” หมายถึง ให้มีตัวอักษร b จำนวน 3 ตัว นั่นคือ “abbb”, “ab{3,5}” หมายถึง ให้มีตัวอักษร b จำนวน 3-5 ตัว นั่นคือ “abbb”, “abbbb” หรือ “abbbbb”

\n ใช้แทนว่าขึ้นบรรทัดใหม่

\t ใช้แทน Tab หรือ ย่อหน้า

\r ใช้แทน return หรือ CR (Carriage return)

\w ใช้แทน word หรือ ตัวอักษรภาษาอังกฤษทั้งตัวเล็กและตัวใหญ่ ตัวเลข 0-9 รวมไปถึง เครื่องหมาย “_” ซึ่งให้ผลลัพธ์เช่นเดียวกับ [a-zA-Z_0-9]

\W ใช้อะไรก็ได้ที่ไม่ใช่ word หรือ ตัวอักษรภาษาอังกฤษทั้งตัวเล็กและตัวใหญ่ ตัวเลข 0-9 และ “_” ให้ผลลัพธ์เช่นเดียวกับ [^a-zA-Z_0-9]

\s ใช้แทนช่องว่าง หรือ Whitespace ให้ผลลัพธ์เช่นเดียวกับ [\n\t\v]

\S ใช้แทนอะไรก็ได้ที่ไม่ใช่ช่องว่าง หรือ Whitespace ให้ผลลัพธ์เช่นเดียวกับ [^\n\t\v]

\xhh ใช้แทนเลขฐานสิบหก (โดย hh ในที่นี้คือ เลขฐานสิบหก)

4.3.3 ตัวอย่างการใช้งาน

“at” จับคู่กับ ตัวอักษร 3 ตัว โดยตัวแรกจะเป็นตัวอักษรอะไรก็ได้ตามด้วย at เช่น hat, bat, cat เป็นต้น

“^[^h]at” หมายถึง กำนี้ต้องขึ้นต้นด้วยตัวอักษรอะไรก็ได้ที่ไม่ใช่ h ตามด้วย at

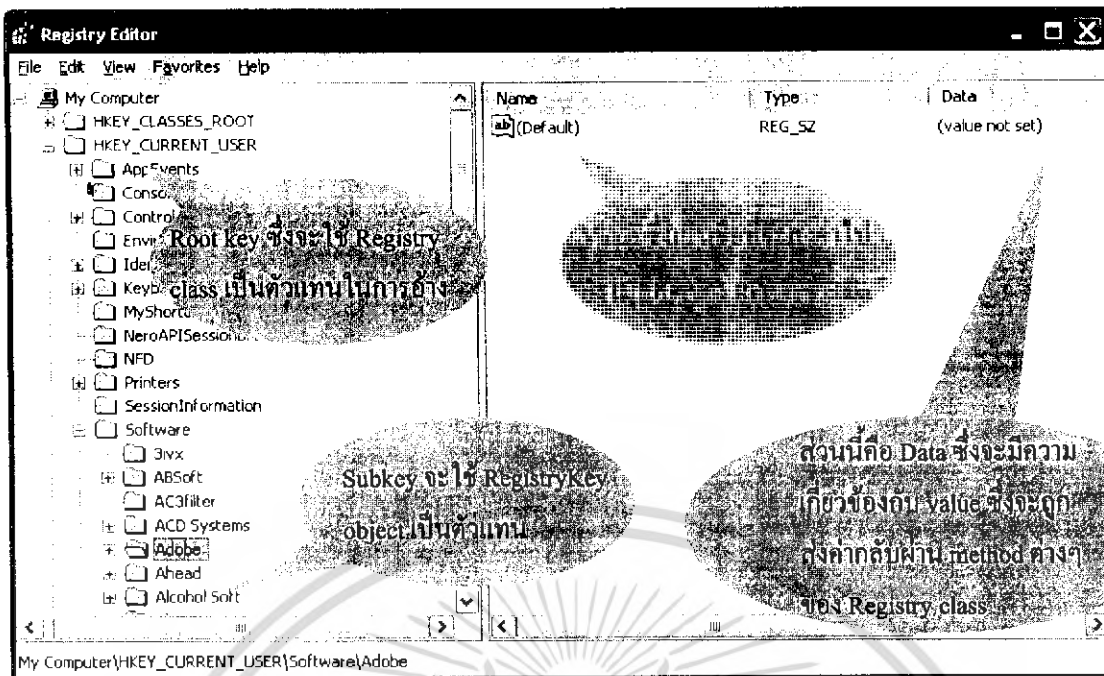
“^[1-9][0-9]*\$” หมายถึง ขึ้นต้นด้วยตัวเลข 1-9 และอาจต่อด้วยตัวเลข 0-9 ก็ได้ ในกรณีนี้หากเป็นเลข 0 โดดๆ จะไม่ได้ จะผ่านตั้งแต่ 1, 2, 3, 4 เป็นต้นไป และท้ายตัวเลขชุดนี้ต้องมีตัวอะไรตาม

“^(0|[1-9][0-9]*)\$” ความหมายเหมือนกับ “^[1-9][0-9]*\$” แต่จะสามารถแทนเลข 0 โดดๆได้ด้วย

“^(0|?[1-9][0-9]*)\$” ความหมายเหมือนกับ “^(0|[1-9][0-9]*)\$” เพียงแต่ ถ้าไม่ขึ้นต้นด้วยเลข 0 สามารถมีเครื่องหมาย ลบ ได้ หรือจะไม่มีเครื่องหมายลบ ก็ได้ เครื่องหมาย ? บ่งบอกว่า จะมีหรือไม่มี ก็ได้

“^[0-9]{1,3}([0-9]{3})*(\.[0-9]{1,2})?\$” หมายถึง ต้องขึ้นต้นด้วยตัวเลข 0-9 จำนวน 1-3 ตัว หรืออาจตามด้วยเครื่องหมายคอมม่า “,” พร้อมกับตัวเลข 0-9 อีก 3 ตัว (ต่อกันก็ชุดก็ได้) และสุดท้ายอาจจะต่อด้วยทศนิยม 1 หรือ 2 ตำแหน่ง หรือไม่ก็ได้ เช่น “0”, “110”, “20,325”, “3.08”, “199,789.00”, “123,456,789.99” เป็นต้น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.13 Registry Editor

ในการเขียน รีจิสตรี ใน .NET จะใช้ 2 class ในการทำงานที่เกี่ยวกับรีจิสตรีนั่นคือ Microsoft.Win32.RegistryKey และ Microsoft.Win32.Registry ซึ่ง RegistryKey class จะถูกใช้เป็นตัวแทนของ subkey ในรีจิสตรี และ Registry class จะถูกใช้เป็นตัวแทนของ root key ในรีจิสตรีเท่านั้น (root key เปรียบเสมือน sub key ชนิดพิเศษซึ่งอยู่ระดับบนขึ้นไป)

4.4.2 การเลือก Subkey

ไม่ว่าเราจะต้องการค่าไหนในรีจิสตรี จะต้องเลือกบาง subkey มาก่อนซึ่งในการเลือก subkey เราจะต้องเลือก subkey root ขึ้นมาก่อน ในการเลือก root key เราจะต้องการสร้าง RegistryKey ขึ้นมาใหม่เพื่อที่จะเข้าถึง subkey ที่เราต้องการ

ในตอนที่จะเลือก root key ซึ่งทั้งหมดอยู่ใน Microsoft.Win32.Registry ในการเลือกนั้นจะต้องเลือกใน Microsoft.Win32.Registry class ซึ่งมีดังนี้

Field Name	Use	Description
ClassesRoot	to access HKEY_CLASSES_ROOT	ข้อมูล file-types, components และอื่นๆ
CurrentConfig	to access HKEY_CURRENT_CONFIG	ข้อมูลของฮาร์ดแวร์ปัจจุบัน
CurrentUser	to access HKEY_CURRENT_USER	ข้อมูลของผู้ใช้ปัจจุบัน
LocalMachine	to access HKEY_LOCAL_MACHINE	การปรับแต่งค่าของ current machine
Users	to access HKEY_USERS	ข้อมูลการปรับแต่งค่าของผู้ใช้

ตารางที่ 4.1 รายละเอียดของ subkey root

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4 การปรับแต่งค่าในรีจิสตรี (Registry Configuration)

รีจิสตรี คือ ที่เก็บค่าที่เกี่ยวข้องกับโปรแกรมต่างๆ ในวินโดวส์ เช่น การตั้งค่าต่างๆของผู้ใช้ ค่าดีฟอลต์ของระบบ รีจิสตรีจะจัดการเมื่อโปรแกรมกำลังปิดการทำงานและทางผู้ใช้งานที่ต้องการที่จะเก็บข้อมูลเพื่อนำมาใช้อีกครั้งตอนที่เรียกโปรแกรมขึ้นมา

4.4.1 หลักการที่เกี่ยวข้อง

รีจิสตรีเปรียบเสมือนฐานข้อมูลขนาดใหญ่ ข้อมูลทั้งหมดจะถูกเก็บในรูปแบบที่คล้ายกัน คือในรูปแบบของ Key-value ซึ่งหมายถึงว่าจะมี key และมี value ที่เกี่ยวข้องกันเช่น

Key: "ForeColor"

Value: "Red"

อธิบายได้ว่าเรามี Key คือ ForeColor และมีค่าคือ Red ซึ่งจะมีข้อแตกต่างในรูปแบบที่ใช้ในการเก็บรีจิสตรี ดังนั้นตามตัวอย่างข้างต้น ForeColor คือค่าที่มีข้อมูลคือ Red

เพื่อที่จะเข้าใจ โครงสร้างของรีจิสตรีให้ไปทดลองศึกษาที่เครื่องของท่านเองโดย Start -> Run -> regedit จะปรากฏหน้าต่างใหม่ขึ้นมาซึ่งจะมีการแสดงค่าต่างๆเป็นแบบโครงสร้าง ไดรากทอรีในทางฝั่งซ้าย ซึ่งจะเก็บรายละเอียดของ values และ data ต่างๆ

วินโดวส์จะเก็บข้อมูลในรีจิสตรีค่อนข้างเยอะมาก นอกจากนี้ข้อมูลของ โปรแกรมที่ได้ติดตั้งไปของผู้ใช้คนนั้นๆจะเก็บค่าปรับแต่งต่างๆในรีจิสตรีด้วยเช่นกัน ข้อมูลต่างๆที่เก็บไว้ในรีจิสตรี ต้องจัดการให้ดี เพราะวารีจิสตรีจะถูกจัดเก็บไว้แยกกับไฟล์เตอร์ที่เราใช้กันอยู่

อธิบายการเขียนโปรแกรมที่เกี่ยวข้องกับรีจิสตรี

ก่อนอื่นเราต้องทำความเข้าใจคำพื้นฐานของ 4 อีลีเมนต์ที่เก็บอยู่ในรีจิสตรีคือ มี root key, subkey , value , data ซึ่งจะแสดงให้เห็นดังรูป 4.13

เมื่อเราเลือก root key แล้วเราต้องทำการสร้าง RegistryKey object เพื่อที่จะเรียก static RegistryKey.OpenSubKey หรือ เมธอด RegistryKey.CreateSubKey บน root key ที่เราเลือกมา การสร้าง subkey ขึ้นมาใหม่

วิธีการสร้าง subkey ด้วยเมธอด RegistryKey.CreateSubKey คือ

```
public RegistryKey CreateSubKey (string subkey);
```

เช่น

...

```
RegistryKey MyReg = Registry.CurrentUser.CreateSubKey
```

```
("SOFTWARE\\SomeCompany\\SomeApp\\SomeVer");
```

...

ในที่นี้เราจะสร้าง subkey ใหม่ใน CurrentUser root key ที่จะส่งค่า RegistryKey object กลับมา

ข้อสังเกต

- ถ้ามี subkey นี้อยู่แล้ว subkey นั้นจะส่งค่านี้กลับมาและไม่มีการสร้างค่าใหม่ ดังนั้นในทางปฏิบัติ เราจึงสามารถใช้เมธอด CreateSubKey ในการสร้างและการเปิด subkey
- เมธอด CreateSubKey สามารถส่งค่า subkey กลับมาในการเขียนค่า
- ถ้าเราเรียก CreateSubKey อีกครั้ง (ภายหลังจากการสร้าง subkey) ค่า child subkeys ใหม่จะถูกสร้างขึ้น

4.4.3 การเปิดค่า Subkey ที่มีอยู่แล้ว

ในการเปิดค่า subkey ที่มีอยู่แล้ว เราจะใช้เมธอด RegistryKey.OpenSubKey เช่น

```
RegistryKey MyReg = Registry.CurrentUser.OpenSubKey (...);
```

OpenSubKey จะมีอยู่ 2 โอเวอร์โหลด คือ

```
public RegistryKey OpenSubKey (string name);
```

ซึ่งจะส่งค่า subkey ที่อ่านได้อย่างเดียว

```
public RegistryKey OpenSubKey (string name, bool writable);
```

ซึ่งจะส่งค่า subkey ในโหมดที่ถูกระบุไว้ว่าสามารถเขียนได้หรืออ่านได้อย่างเดียว โหมดที่ใช้ในการเขียนต้องตั้งให้เป็น true และถ้าต้องการให้อ่านได้อย่างเดียวให้เซตเป็น false โดยที่ 2 เมธอดนั้น เราจะต้องส่ง path ที่จะเข้าถึงลงไป ถ้าพบจะส่งค่า RegistryKey object กลับมาถ้าไม่พบจะส่งค่า null กลับมา

เช่น

...

```
RegistryKey MyReg = Registry.CurrentUser.OpenSubKey
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```
("SOFTWARE\\SomeCompany\\SomeApp\\SomeVer", true);
```

...

ส่งค่า RegistryKey object กลับมาจาก CurrentUser root key กับการระบุ path ที่จะใช้ในการเขียน
กับ subkey ที่มีอยู่

4.4.4 การเรียกค่าที่เก็บอยู่มาใช้

เมื่อเราได้ค่า RegistryKey object แล้วเราสามารถเข้าถึงค่า subkey และ data ได้โดยการใช้
เมธอด RegistryKey.GetValue ดังนี้

```
someRegistryKeyobject.GetValue (...);
```

GetValue มีอยู่ 2 โอเวอร์โหลดคือ

```
public object GetValue (string name);
```

เมธอด นี้จะส่งค่าที่เป็นตัวแทนของ subkey หรือ ค่า null ถ้า value ไม่มีอยู่

```
public object GetValue (string name, object defaultValue);
```

เมธอด นี้จะส่งค่าที่เป็นตัวแทนของ subkey หรือค่า default ที่ระบุไว้โดย default Value

จะเห็นว่า GetValue จะส่งข้อมูลกลับมาเป็นชนิด object ดังนั้นเราสามารถทำการ cast ชนิดของ
ข้อมูลที่ต้องการ

4.4.5 การปรับแต่งค่าของข้อมูล

ในการเซตค่าเราจะใช้เมธอด RegistryKey.SetValue ดังนี้

```
public void SetValue (string name, object value);
```

string name จะเก็บ value ไว้ส่วนค่าของ object value จะเก็บ data ที่เราต้องการ แล้วจะแปลงค่าเป็น
DWORD, binary หรือ string นี่ก็เหตุผลว่าทำไม เมธอด RegistryKey.GetValue จะส่งข้อมูล
ออกมาใน แบบ type object

4.4.6 การปิด Subkey

เมื่อเราได้เข้าถึง subkey แล้วเราจะต้องปิดการใช้งานด้วย โดยการใช้เมธอด

RegistryKey.Close สำหรับการปิด subkey เช่น

```
RegistryKey MyReg = Registry.CurrentUser.CreateSubKey
```

```
("SOFTWARE\\SomeCompany\\SomeApp\\SomeVer");
```

```
int nSomeVal = (int)MyReg.GetValue("SomeVal", 0);
```

```
MyReg.SetValue ("SomeValue", nSomeVal+1);
```

```
MyReg.Close ();
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.4.7 การลบ Subkey

ในการลบ subkey จะมี 2 โอเวอร์โหลดของเมธอด RegistryKey.DeleteSubKey

```
public void DeleteSubKey(string subkey);
```

จะทำการลบ subkey ที่ระบุไว้

```
public void DeleteSubKey(string subkey, bool throwOnMissingSubKey);
```

จะทำการลบ subkey ที่ระบุไว้ และถ้าไม่พบ subkey จะ throwOnMissingSubKey ถ้าเซตค่าเป็น true

4.4.8 การลบ Entire tree

การลบ subkey และ subkey ย่อยๆรวมทั้ง ค่าอื่นๆทั้งหมด เราจะต้องใช้เมธอด RegistryKey.DeleteSubKeyTree เช่น

```
public void DeleteSubKeyTree(string subkey);
```

4.4.9 การลบ Value

ในการลบ Values จะมีเมธอดอยู่ 2 โอเวอร์โหลดคือ

```
public void DeleteValue(string name);
```

โดยให้เราระบุ Value กับตัวแปร name ซึ่งจะเป็นตัวที่ถูกลบ และ

```
public void DeleteValue(string name, bool throwOnMissingValue);
```

เมธอดนี้ก็เช่นกันแต่จะมีการ thrown exception ถ้าค่า throwOnMissingValue เป็น true

4.4.10 การรับค่า List of subkeys

การรับค่า List of subkeys จะใช้เมธอด RegistryKey.GetSubKeyNames คือ

```
public string[] GetSubKeyNames();
```

4.4.11 การรับค่า list of value names

การรับค่า list of value names จะใช้เมธอด RegistryKey.GetValueNames

```
public string[] GetValueNames();
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 ส่วนต่อขยายบนไฟร์ฟอกซ์ (Firefox Extension)

ในการเขียน Extension ให้ Firefox นั้นเราจะต้องรู้ XUL , JavaScript และภาษาอื่นที่จำเป็นจะต้องใช้ในการเขียน Extension

4.5.1 โครงสร้างของไฟล์ Extension

โครงสร้างของไฟล์ Extension จะประกอบไปด้วย

extension.xpi:

```
extension.rdf
chrome/extension.jar
components/extension.dll
components/extension.js
defaults/extension.something
defaults/preferences/extension.js
```

โดยที่ไฟล์เหล่านี้จะอยู่ใน xpi ไฟล์ จะต้องมี jar ไฟล์อยู่ในโฟลเดอร์ chrome โดยที่ส่วนประกอบของ XPCOM จะกำหนดให้ไฟล์ preference อยู่ที่ defaults/preferences ในส่วนของ Extension Manager จะทำการตรวจสอบไฟล์เหล่านี้ว่าอยู่ที่ไหนในตอนเริ่มโปรแกรม การติดตั้งการลบโปรแกรมที่ติดตั้งแล้ว เพื่อที่จะดำเนินการกับไฟล์อาทิเช่น chrome โดยอัตโนมัติและการไหลคค่าดีฟอลต์ของ prefs เป็นต้น

4.5.2 ไฟล์ extension.rdf

โครงสร้างของไฟล์นี้โดยรวมๆจะเป็นเช่นนี้

```
<?xml version="1.0"?>
<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:em="http://www.mozilla.org/2004/em-rdf#">
  <Description about="urn:mozilla:extension:manifest">
    <!-- properties -->
  </Description>
</RDF>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.3 ส่วนต่างๆที่ไฟล์ extension.rdf เป็นคุณสมบัติที่ต้องมี

em:id คือส่วนนี้เปรียบเสมือนค่า id ของโปรแกรมนั้นๆที่เราจะเขียนขึ้น

```
<em:id>{daf44bf7-a45e-4450-979c-91cf07434c3d}</em:id>
```

em:version คือส่วนนี้คือส่วนที่เรียกว่า Firefox Version Format (FVF) เพื่อระบุว่าโปรแกรมที่ใช้ อยู่คือโปรแกรมเวอร์ชันที่เท่าไร

```
<em:version>4.6</em:version>
```

em:targetApplication คือส่วนนี้คือเป็นส่วนที่ระบุว่าเป็นโปรแกรมที่จะใช้ เป็น Extension ของ Firefox ซึ่งจะมีส่วนของ minVersion และ maxVersion คือระบุว่าจะใช้กับ Firefox เวอร์ชันไหนได้บ้าง ซึ่งนอกจากนี้ถ้าอยากให้ Extension นี้ใช้ได้กับ Thunderbird ได้เราสามารถระบุค่า id เปลี่ยนไปตามที่ Thunderbird ต้องการ ซึ่งตัวอย่างนี้จะใช้กับ Firefox เท่านั้น เช่น

```
<em:targetApplication>
```

```
<Description>
```

```
<em:id>{ec8030f7-c20a-464f-9b0e-13a3a9e97384}</em:id>
```

```
<em:minVersion>0.9</em:minVersion>
```

```
<em:maxVersion>2.0.0.1</em:maxVersion>
```

```
</Description>
```

```
</em:targetApplication>
```

em:name คือส่วนนี้จะป็นชื่อของ Extension ของเรา

```
<em:name>My Extension</em:name>
```

em:file คือส่วนนี้จะเปรียบเสมือน URIs เพื่อบอกว่าไฟล์ที่อยู่ใน .jar ว่าประกอบด้วยอะไรบ้างและ อยู่ที่ไหนบ้าง เช่น

```
<em:file>
```

```
<Description about="urn:mozilla:extension:file:myext.jar">
```

```
<em:package>content/myext/</em:package>
```

```
<em:locale>locale/en-US/myext/</em:locale>
```

```
<em:skin>skin/classic/myext/</em:skin>
```

```
</Description>
```

```
</em:file>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5.4 ส่วนต่างๆที่ไฟล์ extension.rdf เป็นคุณสมบัติเพิ่มเติม

em:description คือส่วนของรายละเอียดของไฟล์ Extension ที่จะแสดงรายละเอียดออกมาในหน้า user interface เช่น

```
<em:description>Advanced foo tools.</em:description>
```

em:creator คือเป็นที่ระบุชื่อผู้เขียน โปรแกรมนี้ เช่น

```
<em:creator>Prueak & Worawit</em:creator>
```

em:contributor คือเป็นที่ระบุชื่อผู้ที่ทำให้การสนับสนุนที่เกี่ยวข้องกับโปรแกรมนี้ เช่น

```
<em:contributor>NECTEC and SIPA</em:contributor>
```

em:homepageURL เป็นหน้า home page ของผู้ผลิต เช่น

```
<em:homepageURL>http://www.ce.kmitl.ac.th/</em:homepageURL>
```

em:updateURL คือส่วนที่เกี่ยวข้องกับการอัปเดตโปรแกรม extension นี้ ซึ่งถ้าได้ enable ไว้ตัวโปรแกรมจะไปเช็คว่ามีเวอร์ชันที่เป็นเวอร์ชันใหม่กว่าหรือไม่ถ้ามีเวอร์ชันที่ใหม่กว่าจะทำการอัปเดตให้ ซึ่งทางฝั่งเซิร์ฟเวอร์จะต้องส่งไฟล์ในรูปแบบ text/rdf หรือ หรือมีตัวเช็คอัปเดตว่าทำงานหรือไม่ ทาง Firefox จะทำการส่งค่าเหล่านี้ผ่านทาง URL ในกรณีที่ได้ให้มีการโต้ตอบแบบ RDF Dynamically ยกตัวอย่างเช่น PHP หรือ CGI รูปแบบเป็นดังนี้

%ITEM_ID% เป็น id ของ Extension ที่จะทำการอัปเดต

%ITEM_VERSION% เป็น version ของ Extension ที่จะทำการอัปเดต

%APP_ID% เป็น id ของ Extension ปัจจุบัน

%APP_VERSION% เป็น version ของ Extension ปัจจุบัน

เช่น

```
<em:updateURL>http://www.ce.kmitl.ac.th/update.cgi?id=%ITEM_ID%&version=%ITEM_VERSION%</em:updateURL>
```

em:optionsURL คือ URL ของ Extension เช่น

```
<em:optionsURL>chrome://myext/content/options.xul</em:optionsURL>
```

em:aboutURL คือ URL ที่เกี่ยวกับ Extension นี้ เช่น

```
<em:aboutURL>chrome://myext/content/about.xul</em:aboutURL>
```

em:iconURL คือ URL ของไอคอนขนาด32*32ที่จะไปแสดงบน Extension Manager UI เช่น

```
<em:iconURL>chrome://myext/skin/icon.png</em:iconURL>
```

4.5.5 ตัวอย่างในการสร้างไฟล์ Extension

1. สร้างไฟล์ content โครงสร้างเป็น เช่นนี้

c:\myextensions\

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- + WebContentFiltering
 - + content
 - + contents.rdf
 - + WebContentFiltering-Overlay.xul
 - contents.rdf เขียนโปรแกรมดังนี้

```

<?xml version="1.0"?>
<RDF:RDF xmlns:RDF="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:chrome="http://www.mozilla.org/rdf/chrome#">
  <RDF:Seq about="urn:mozilla:package:root">
    <RDF:li resource="urn:mozilla:package:WebContentFiltering"/>
  </RDF:Seq>
  <RDF:Description about="urn:mozilla:package:WebContentFiltering"
    chrome:displayName="Web Content Filtering"
    chrome:author="Prueak and Worawit"
    chrome:authorURL="http://www.ce.kmitl.ac.th"
    chrome:name="Web Content Filtering"
    chrome:extension="true"
    chrome:settingsURL="chrome://WebContentFiltering/content/settings/settings.xul"
    chrome:description="Filter content and image from Web Site.">
  </RDF:Description>
  <RDF:Seq about="urn:mozilla:overlays">
    <RDF:li resource="chrome://browser/content/browser.xul"/>
  </RDF:Seq>
  <!-- overlay information for Mozilla Firebird-->
  <RDF:Seq about="chrome://browser/content/browser.xul">
    <RDF:li>chrome://WebContentFiltering/content/WebContentFiltering-Overlay.xul</RDF:li>
  </RDF:Seq>
</RDF:RDF>

```

อธิบายโปรแกรมเพิ่มเติม

<RDF:Seq about="chrome://browser/content/browser.xul">
 เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

<RDF:li>chrome://WebContentFiltering/content/WebContentFiltering-Overlay.xul</RDF:li>

ส่วนนี้เป็นส่วนของการอ้างอิง path ที่ต้องใช้ในการแสดงผลออกหน้าจอเช่นพวก GUI ต่างๆ

- สร้างไฟล์ WebContentFiltering-Overlay.xul

```
<?xml version="1.0"?>
```

```
<overlay id="helloworldOverlay"
```

```
  xmlns="http://www.mozilla.org/keymaster/gatekeeper/there.is.only.xul">
```

```
<menupopup id="menu_ToolsPopup">
```

```
  <menuitem label="Web Content Filtering" position="1" />
```

```
</menupopup>
```

```
</overlay>
```

ส่วนนี้จะป็นลักษณะของ GUI ในการแสดงผล ส่วนนี้ต้องทำการศึกษาภาษา XUL ว่าแต่
 ะแท็กที่ใช้ต้องใช้อะไรในการอ้างอิง

2. สร้างไฟล์ install.rdf ซึ่งการเขียนได้อธิบายไว้ในข้างต้นแล้ว

```
<?xml version="1.0"?>
```

```
<RDF xmlns="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
```

```
  xmlns:em="http://www.mozilla.org/2004/em-rdf#">
```

```
<Description about="urn:mozilla:install-manifest">
```

```
<em:id>{12a1584b-2123-473d-8752-e82e74e3cb1b}</em:id>
```

```
<em:version>1.0.0</em:version>
```

```
<em:targetApplication>
```

```
<Description>
```

```
<em:id>{ec8030f7-c20a-464f-9b0e-13a3a9e97384}</em:id>
```

```
<em:minVersion>1.0</em:minVersion>
```

```
<em:maxVersion>2.0.0.1</em:maxVersion>
```

```
</Description>
```

```
</em:targetApplication>
```

```
<em:name>Web Content Filtering</em:name>
```

```
<em:description>Filter content and image from Web Site.</em:description>
```

```
<em:creator>Prueak and Worawit CE@KMITL</em:creator>
```

```
<em:homepageURL>http://www.ce.kmitl.ac.th/</em:homepageURL>
```

```
<em:contributor>NECTEC and SIPA</em:contributor>
```

```
<em:file>
```

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

<Description about="urn:mozilla:extension:file:WebContentFiltering.jar">
  <em:package>content/</em:package>
</Description>
</em:file>
</Description>
</RDF>

```

3. สร้างไฟล์ WebContentFiltering.jar เกิดจากการ zip ไฟล์แล้วเปลี่ยนเป็นนามสกุล .jar โดยมีโครงสร้างดังนี้

WebContentFiltering.jar

```

+- content/
+- contents.rdf
+- WebContentFiltering-Overlay.xul

```

4. สร้างไฟล์ WebContentFiltering.xpi เกิดจากการ zip ไฟล์แล้วเปลี่ยนเป็นนามสกุล .xpi โดยมีโครงสร้างดังนี้

helloworld.xpi

```

+- install.rdf
+- chrome/
+- WebContentFiltering.jar

```

5. เมื่อสร้างเสร็จเรียบร้อยแล้วจะมีไฟล์โครงสร้างตามรูปนี้

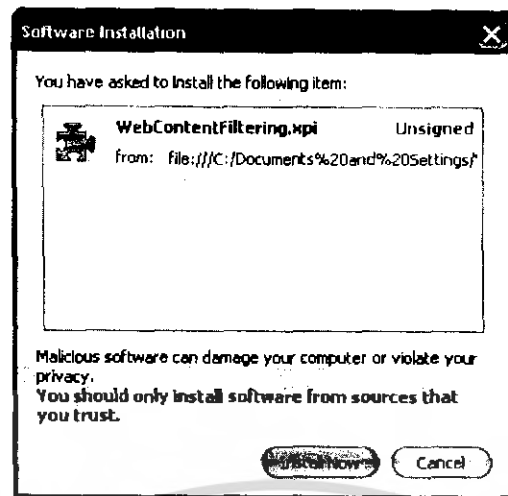
```

+- WebContentFiltering
+- content
+- contents.rdf
+- helloworld-Overlay.xul
+- chrome
+- WebContentFiltering.jar
+- install.rdf
+- WebContentFiltering.xpi

```

ซึ่งไฟล์ที่จะต้องเอาไปใช้คือ .xpi สร้างเปิดโดยใช้ Firefox ได้และจะทำการติดตั้งได้ ซึ่งเมื่อได้ติดตั้งจะได้ผลดังรูปที่ 4.14, 4.15, 4.16

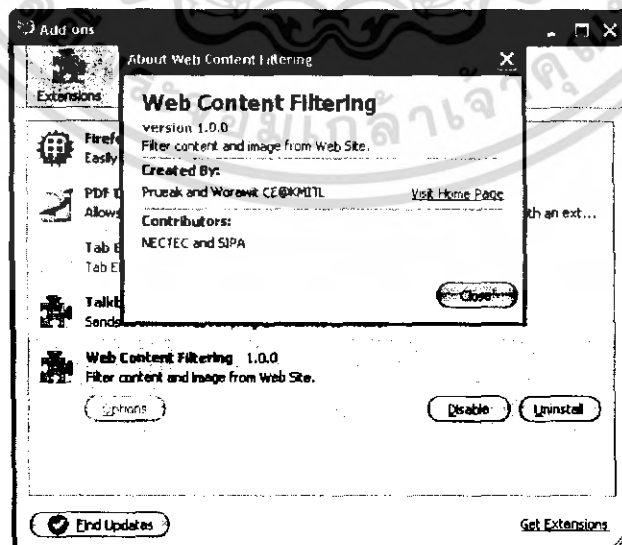
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.14 การติดตั้ง Extension



รูปที่ 4.15 Extension ที่ติดตั้งเสร็จเรียบร้อยแล้ว



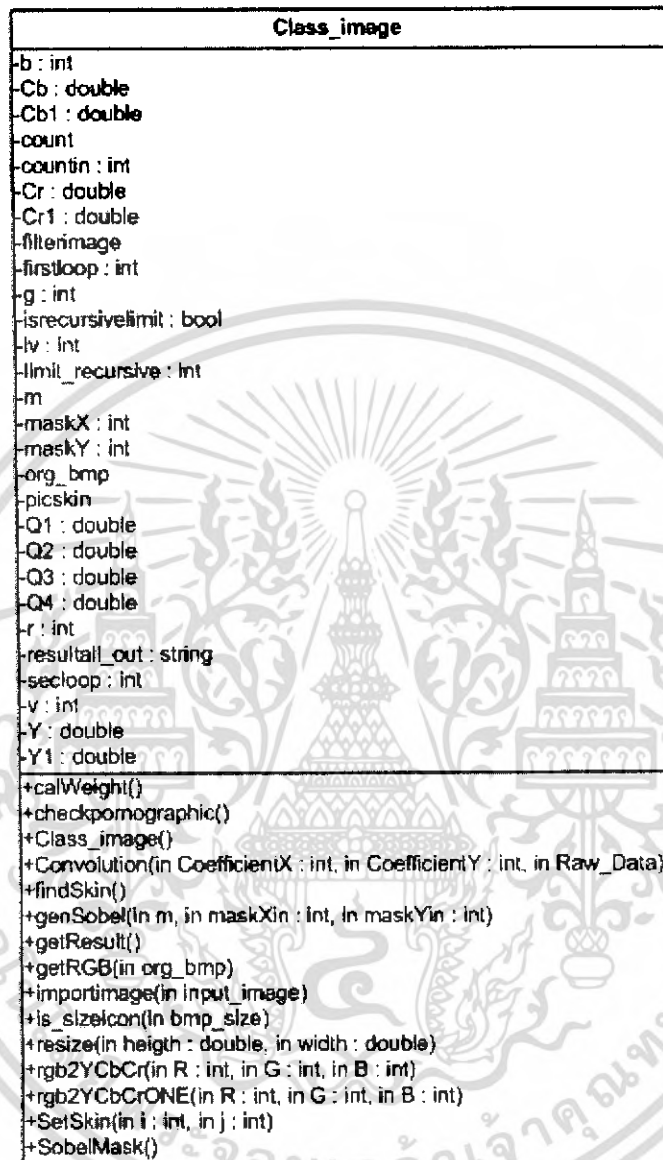
รูปที่ 4.16 Extension ที่ได้แสดงขึ้นมาในเมนู Add-ons และทำการกด About

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

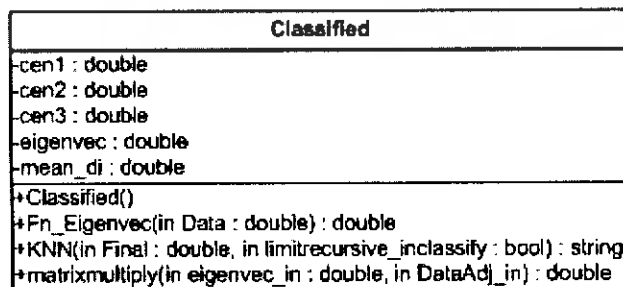
4.6 คลาสไดอะแกรมของโปรแกรมคัดกรองข้อมูล (Class Diagram)

คลาสไดอะแกรมทั้งหมดของโปรแกรมนี้อจะเป็นไปดังรูปที่ 4.17 – 4.32

4.6.1 คลาสที่เกี่ยวข้องกับWindowsService1



รูปที่ 4.17 คลาสชื่อ Class_image



รูปที่ 4.18 คลาสชื่อ Classified

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

PersonalWebProxy
+count : int +countAll : int +countPic : int +pfnWorkerCallBack +SocketEvent -privilegeDN -sListener -m_clientCount : int -m_workerSocketList -web_page
+OnClientConnect(in asyn) : void +PersonalWebProxy() +ReceiveCallBack(in asyn) : void +SendCallBack(in ar) : void +Start_Server() : void +WaitForData(in soc, in clientNumber : int) : void -RequestHTTPHeader(in szData : string, in uri : string) : <unspecified> -ResponseHTTPHeader(in resp) : string -CloseSocketsListening() : void -CloseSockets(in workerSocket) : void

รูปที่ 4.19 คลาสชื่อ PersonalWebProxy

«struct»PrivilegeDN
+listPrivilegeDN +InitialListPrivilege() : void +IsPrivilegeDN(in url : string) : bool

รูปที่ 4.20 สตักชื่อ PrivilegeDN

Program
-Main() : void

รูปที่ 4.21 คลาสชื่อ Program

ProjectInstaller
+serviceInstaller1 -components -serviceProcessInstaller1 +ProjectInstaller() #Dispose() : void -InitializeComponent() : void

รูปที่ 4.22 คลาสชื่อ ProjectInstaller

Service1
+pipeManager -components +Service1() #Dispose(in disposing) : void #OnStart(in args : string) : void #OnStop() : void -InitializeComponent() : void

รูปที่ 4.23 คลาสชื่อ Service1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

WebPage
-anWord -code : string -countAllLink : double -countImg : double -countImgLink : double -countLinkNonRela : double -countLinkRela : double -countMalefWord : double -countMeta : double -countParam : double -countScript : double -countTxtLink : double -url : string
+DeletePopup(in htmlCode : string) : string +DetectTag() : void +FindMaleficWord(in str : string) : long +GetCode() : string +GetCountAllLink() : double +GetCountImg() : double +GetCountImgLink() : double +GetCountLinkNonRela() : double +GetCountLinkRela() : double +GetCountMalefWord() : double +GetCountMeta() : double +GetCountParam() : double +GetCountScript() : double +GetCountTxtLink() : double +GetUri() : string +IsMaleficWeb() : bool +SetCode(in code : string) : void +SetCountAllLink(in countAllLink : double) : void +SetCountImg(in countImg : double) : void +SetCountImgLink(in countImgLink : double) : void +SetCountLinkNonRela(in countLinkNonRela : double) : void +SetCountLinkRela(in countLinkRela : double) : void +SetCountMalefWord(in countMalefWord : double) : void +SetCountMeta(in countMeta : double) : void +SetCountParam(in countParam : double) : void +SetCountScript(in countScript : double) : void +SetCountTxtLink(in countTxtLink : double) : void +SetUri(in url : string) : void +WebPage(in url : string, in htmlCode : string)

รูปที่ 4.24 คลาสชื่อ WebPage

«struct» WordPool
+listWordPool
+InitialListWordPool() : void

รูปที่ 4.25 สตักชื่อ WordPool

WServiceNamedPipe
+gobal_level : string +goodList : string -disposed : bool
#Dispose(in disposing : bool) : void +~WServiceNamedPipe() +IsInGoodList(in url : string) : bool +PipeListener() : void +WServiceNamedPipe(in name : string, in outBuffer : uint, in inBuffer : uint, in maxReadBytes : int, in secure : bool)

รูปที่ 4.26 คลาสชื่อ WServiceNamePipe

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.6.2 คลาสที่เกี่ยวข้องกับไฟล์ WindowsControl

License_Agreement
-components -groupBox1 -groupBox2 -label1 -label2
+License_Agreement() #Dispose(in disposing : bool) : void -InitializeComponent() : void

รูปที่ 4.27 คลาสชื่อ License_Agreement

IniFileConfig
+path : string
+IniFileConfig(in INIPath : string) +IniReadValue(in Section : string, in Key : string) : string +IniWriteValue(in Section : string, in Key : string, in Value : string) : void -WritePrivateProfileString(in section : string, in key : string, in val : string, in filePath : string) : long -GetPrivateProfileString(in section : string, in key : string, in def : string, in retVal, in size : int, in filePath : string) : int

รูปที่ 4.28 คลาสชื่อ iniFileConfig

CheckPassword
+chk_password_correct : bool +getPwd -component -btn_chkcancel -btn_chkok -label1 -registry_program : string -txt_chkpassword
+CheckPassword() #Dispose(in disposing : bool) -btn_chkok_Click(in sender : object, in e) : void -btn_chkcancel_Click(in sender : object, in e) : void -InitializeComponent() : void

รูปที่ 4.29 คลาสชื่อ CheckPassword

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

WindowsControl
+notifyIcon1
+pipeManager
+refListView_UrlBlocked
+refListView_UrlPassed
+registry_levelMode : string
+registry_program : string
#setReg
-appPath : string
-btn_add_goodlist
-btn_apply
-btn_change_password
-btn_delete_goodlist
-btn_License
-btn_vlewlog
-chkbl_goodlist
-col_timeblock
-col_timepass
-col_urlblock
-col_urlpass
-components
-contextMenuStrip1
-en_d
-en_dt
-getReg
-groupBox1-10
-label_date1
-label_date2
-label1-20
-lb_high
-lb_high_des1
-lb_high_des2
-lb_low
-lb_low_des1
-lb_medium
-lb_medium_des1
-lb_medium_des2
-levelMode
-listviewBlock
-listviewPass
-logo
-MenuItem1
-monthCalendar1
-pathlog : string
-pathprogram : string
-scroll_Level
-selectedDate
-serviceController1
-tabControl1
-tabPage1
-tabPage2
-tabPage3
-tabPage4
-tabPage5
-tabPage6
-timer
-txt_addgoodlist
-txt_confirmpassword
-txt_newpassword
-txt_oldpassword

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

```

+Append_UrlBlocked(in uri : string) : void
+Append_UrlPassed(in uri : string) : void
+SendGoodListRegistryToService() : void
+WindowsControl()
#Dispose(in disposing : bool) : void
-btn_add_goodlist_Click(in sender : object, in e) : void
-btn_apply_Click(in sender : object, in e) : void
-btn_change_password_Click(in sender : object, in e) : void
-btn_cleargoodlist_Click(in sender : object, in e) : void
-btn_delete_goodlist_Click(in sender : object, in e) : void
-btn_License_Click(in sender : object, in e) : void
-btn_viewlog_Click(in sender : object, in e) : void
-CloseBrowser() : void
-InitializeComponent() : void
-installFirefox() : void
-MenuItem1_Click(in sender : object, in e) : void
-MenuItemShutdown_Click(in sender : object, in e) : void
-monthCalendar1_DateSelected(in sender : object, in e) : void
-notifyIcon1_DoubleClick(in sender : object, in e) : void
-SaveLogInFile() : void
-SetBlackOnHighMode() : void
-SetBlackOnLowMode() : void
-SetBlackOnMediumMode() : void
-SetGrayOnHighMode() : void
-SetGrayOnLowMode() : void
-SetGrayOnMediumMode() : void
-SetSilverOnHighMode() : void
-SetSilverOnLowMode() : void
-SetSilverOnMediumMode() : void
-setUnUseableProxy() : void
-setUseableProxy() : void
-StartService() : void
-StopService() : void
-tabControl1_SelectedIndexChanged(in sender : object, in e) : void
-timer_Tick(in sender : object, in e) : void
-trackBar1_Scroll(in sender : object, in e) : void
-WindowsControl_FormClosing(in sender : object, in e) : void

```

รูปที่ 4.30 คลาสชื่อ WindowsControl

WControlNamedPipe
-disposed : bool
+Dispose(in disposing : bool) : void
+PipeListener() : void
+WControlNamedPipe(in name : string, in outBuffer : uint, in inBuffer : uint, in maxReadBytes : int, in secure : bool)
--WControlNamedPipe()

รูปที่ 4.31 คลาสชื่อ WControlNamePipe

ViewLog
-btn_close
-col_status
-col_timeblock
-col_urlblock
-components
-label_dateLog
-label1
-listView_log
+ShowLogFile(in pathLogFile : string) : void
+ViewLog()
+ViewLog(in day : string, in month : string, in year : string)
#Dispose(in disposing : bool) : void
-InitializeComponent() : void
-btn_close_Click(in sender : object, in e) : void

รูปที่ 4.32 คลาสชื่อ ViewLog

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่นุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Program
-Main() : void

รูปที่ 4.32 คลาสชื่อ Program



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

การทดลองและผลการทดลอง

5.1 การทดสอบส่วนที่เป็นเนื้อหาของเว็บเพจ

ในการทดลองส่วนของคัตกรองข้อความ เราจะใช้ปัจจัยในการพิจารณาว่าในหน้าเว็บเพจ 1 หน้า นั้น มีอะไรที่จะสามารถบ่งบอกได้ว่าเว็บไหนเป็นเว็บไซต์ที่ลามกอนาจาร และเว็บไหนเป็นเว็บที่ไม่ลามกอนาจาร ซึ่งมีทั้งหมด 10 ประเภท ดังนี้

- 1) จำนวน IMG Tag
- 2) จำนวน META Tag
- 3) จำนวน POOL (จำนวนทั้งหมดที่เป็นคำไม่เหมาะสมใน Title Tag และ META Tag)
- 4) จำนวน SCRIPT Tag
- 5) จำนวนลิงค์ทั้งหมด
- 6) จำนวนลิงค์ที่เป็นรูปภาพทั้งหมด
- 7) จำนวนลิงค์ที่เป็นตัวหนังสือทั้งหมด
- 8) จำนวนลิงค์ที่ไม่ใช่ที่อยู่ภายในเว็บตัวเอง (Non-Relative Link)
- 9) จำนวนลิงค์ที่อยู่ภายในเว็บตัวเอง (Relative Link)
- 10) จำนวน PARAM Tag (เพื่อนับจำนวนสื่อมัลติมีเดียทั้งหลาย)

5.1.1 ทำการศึกษาคำอนาจารที่จะต้องใช้ (Learning)

เมื่อได้ปัจจัยดังกล่าวแล้ว เราได้ใช้เว็บไซต์ที่ลามกอนาจาร ทั้งต่างประเทศ และในประเทศ ที่จะนำมาทำการศึกษา(Learning) ดังตารางที่ 5.1

ประเภท	จำนวน(เว็บไซต์)
จำนวนเว็บไซต์ที่ลามกอนาจาร	150

ตารางที่ 5.1 จำนวนเว็บไซต์ที่ลามกอนาจาร รวมทั้งต่างประเทศและในประเทศ

เมื่อทำการศึกษาจากเว็บดังกล่าวแล้วพบว่า ได้คำหยาบที่เสี่ยงต่อการเป็นเว็บไซต์ลามกอนาจารดังนี้ (แสดงอยู่ในรูป Regular Expression)

`[^w]+adulter[^w]+, (.*)amateurcouple(.*)[^w]+anal[^w]+, [^w]+anilingus(.*),`

`[^w]+anus[^w]+, [^w]+ass[^w]+, (.*)bdsm(.*), (.*)blowjob(.*), [^w]+bondage[^w]+,`

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

[^w]+boob(s)?[^w]+, [^w]+boobie(s)?[^w]+, (.*borrachas(.*) , (.*bulldoglist(.*) ,
 [^w]+busty[^w]+, [^w]+cam(s)?[^w]+, (.*cfnm(.*) , [^w]+cock(s)?[^w]+, (.*creampie(.*) ,
 [^w]+cum[^w]+, [^w]+cumshot(s)?[^w]+, (.*cybersex(.*) , (.*cybersexual(.*) ,
 [^w]+dildo(s)?[^w]+, (.*femdom(.*) , (.*fuck(.*) , (.*handjob(.*) , (.*hentai(.*) ,
 (.*interracial(.*) , (.*maledom(.*) , (.*malestripper(.*) , (.*masturbat(.*) , [^w]+mature(.*) ,
 (.*milf(.*) , [^w]+nude(.*) , [^w]+nudism(.*) , (.*nudistas(.*) , [^w]+nudity(.*) , [^w]+oral(.*) ,
 (.*orgasm(.*) , [^w]+pantie(.*) , (.*penis(.*) , [^w]+pantie(.*) , [^w]+piss(.*) , (.*pom(.*) ,
 (.*pued(.*) , (.*purecfnm(.*) , [^w]+puss(.*) , [^w]+putas(.*) , [^w]+rimjob(.*) ,
 [^w]+rimming(.*) , (.*rubias(.*) , (.*shemale(.*) , (.*tetazas(.*) , (.*tgp(.*) , [^w]+tits[^w]+ ,
 (.*titworld(.*) , (.*upskirt(.*) , (.*xnxx(.*) , (.*xxx(.*) , (.*yobt(.*) และ (.*ztod(.*)

5.1.2 นำกลุ่มทดลองมาทำการทดสอบ (Testing)

จำนวนเว็บไซต์ที่ไม่ลามกอนาจาร และที่เป็นเว็บไซต์ลามกอนาจารที่นำมาทดสอบ ดัง
 ตารางที่ 5.2 และ ตารางที่ 5.3 ตามลำดับ

ประเภท	จำนวน(เว็บไซต์)
ศิลปะ	19
ธุรกิจ	52
ชื่อของ	107
คอมพิวเตอร์	150
บันเทิง	151
อาหาร	33
สุขภาพ	36
บ้าน	50
ข่าวและสื่อ	143
ภูมิภาค	21
วิทยาศาสตร์	25
สังคม	140
สังคม	34

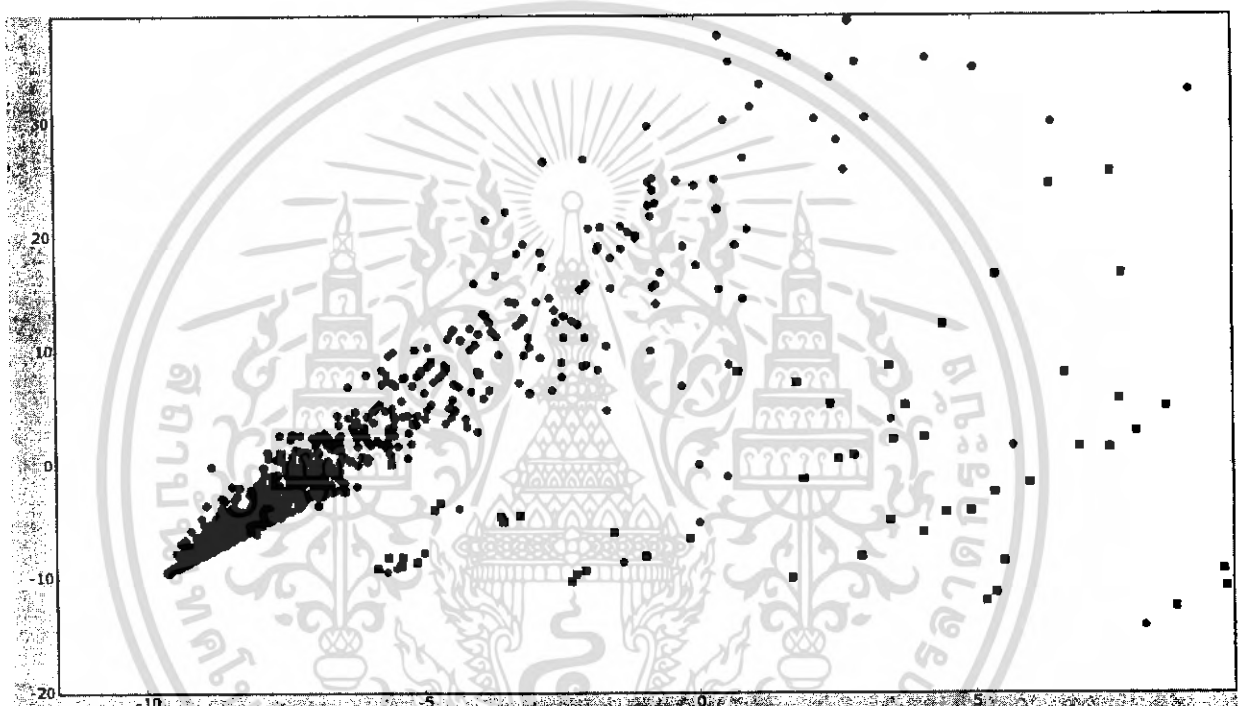
ตารางที่ 5.2 จำนวนเว็บไซต์ที่ไม่ลามกอนาจาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ประเภท	จำนวน(เว็บไซต์)
จำนวนเว็บไซต์ที่เป็นเว็บไซต์ลามกอนาจาร	201

ตารางที่ 5.3 จำนวนเว็บไซต์ที่เป็นเว็บไซต์ลามกอนาจารรวม

เมื่อนำเว็บไซต์เหล่านี้มาทำการหาสถิติทั้ง 10 ประเภทดังที่ได้กล่าวไปแล้ว จากนั้นนำมาทำการลดมิติ (Dimension Reduction) ให้เหลือ 2 มิติ โดยใช้หลักการ Eigen ได้กราฟดังรูปที่ 5.1 ซึ่ง 1 จุดสี่เหลี่ยม ■ หมายถึง 1 เว็บไซต์ลามก และ 1 จุดวงกลม ● หมายถึง 1 เว็บไซต์ปกติไม่ลามก

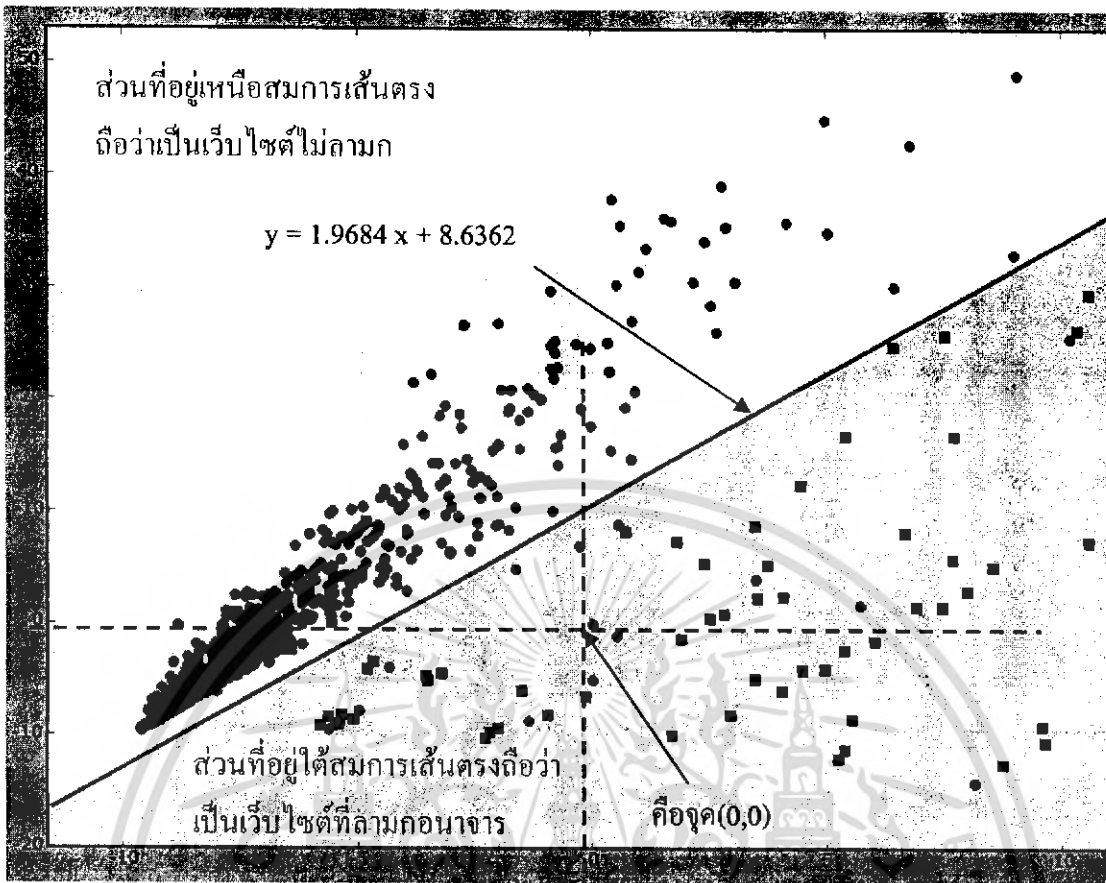


รูปที่ 5.1 กราฟของความสัมพันธ์ของคุณสมบัติของเว็บเพจ

เมื่อได้กราฟมาแล้ว เราจะต้องพิจารณาต่อว่าจะใช้หลักเกณฑ์ใดในการแยกแยะระหว่างเว็บไซต์ลามกอนาจาร กับไม่ลามกอนาจาร ซึ่งจากกราฟจะเห็นได้ว่าจุดสี่เหลี่ยม และจุดวงกลมนั้นแยกออกจากกันค่อนข้างชัดเจน เราจึงได้อาศัยสมการเส้นตรงในการแบ่งกลุ่มของประเภทเว็บไซต์เหล่านั้นดังรูปที่ 5.2 โดย

- ส่วนที่อยู่เหนือเส้นที่กำหนดจะเป็นเว็บไซต์ที่ไม่ใช่เว็บไซต์ลามกอนาจาร
- ส่วนที่อยู่ใต้เส้นที่กำหนดจะเป็นเว็บไซต์ที่ไม่ใช่เว็บไซต์ลามกอนาจาร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.2 สมการบนกราฟที่ใช้แบ่งแยกกลุ่มระหว่างเว็บอนาจาร และเว็บไม่อนาจาร

จากสมการเส้นตรง $y = mx + b$ หากเราให้ความชันคงที่ และเปลี่ยนแปลงแต่ค่า x, y จะได้ค่า b ที่แตกต่างกันออกไปซึ่งเราสามารถรู้ได้ว่าอยู่เหนือสมการเส้นตรงหรืออยู่ใต้สมการเส้นตรงซึ่งในที่นี้ได้สมการเส้นตรงเป็น $y = 1.9684x + 8.6362$ นั่นคือ

- หากค่า b มากกว่า 8.6362 (อยู่เหนือสมการเส้นตรง) ถือว่าเป็นเว็บไซต์ไม่ลามก
- หากค่า b น้อยกว่าหรือเท่ากับ 8.6362 (อยู่ใต้สมการเส้นตรง) ถือว่าเป็นเว็บไซต์ลามก

ผลที่ได้เป็นดังตารางที่ 5.4

ประเภทของเว็บ	จำนวนเว็บ	เว็บอนาจาร	เว็บไม่อนาจาร	ประสิทธิภาพ (%)	ความผิดพลาด (%)
เว็บอนาจาร	201	185	16	91.35	8.65
เว็บไม่อนาจาร	961	945	16	98.33	1.67
รวม	1162	1130	32	97.25	2.75

ตารางที่ 5.4 ผลการทดลองของกลุ่มตัวอย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.2 การทดสอบภาพอนาจาร

หลักการตรวจสอบภาพอนาจารที่ประกอบด้วยขั้นตอนการตรวจหาสีผิวมนุษย์ การตรวจสอบคุณสมบัติของภาพและการจัดประเภทของภาพด้วยคุณสมบัติ ซึ่งใช้หลักการของการประมวลผลภาพเข้ามาช่วย เพื่อให้สามารถพิจารณาคุณสมบัติของภาพตามขั้นตอนการตรวจสอบภาพอนาจาร

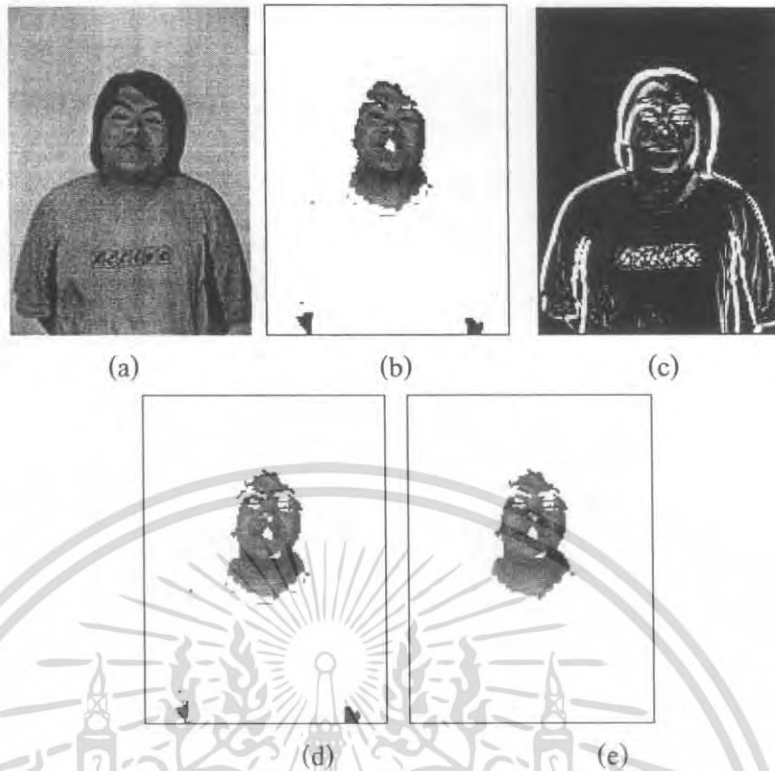
โครงการนี้ได้ใช้ภาษา C# ในการพัฒนาส่วนของโปรแกรมตรวจสอบภาพอนาจารตามขั้นตอนที่ได้กล่าวไว้ และทดสอบการทำงานของโปรแกรมที่ได้พัฒนาขึ้น ดังนี้

5.2.1 การตรวจหาสีผิวมนุษย์

โครงการนี้ทดสอบการตรวจหาสีผิวของมนุษย์โดยการพิจารณาพื้นที่ที่มีค่าของพิกเซลอยู่ในช่วงของสีผิวในภาพ ประกอบกับการใช้การตรวจหาขอบของภาพเข้าร่วมการพิจารณาด้วย ซึ่งผลการทดลองเป็นดังรูปที่ 5.3

วิธีการทดลอง

- นำภาพต้นฉบับมาแปลงเป็นภาพในรูปแบบสี RGB ดังรูปที่ 5.3 (a)
- แปลงข้อมูลภาพให้อยู่ในรูปแบบสี YCbCr แล้วผ่านการตรวจสอบช่วงของสีผิว ถ้าพิกเซลใดมีค่าไม่อยู่ในช่วงของสีผิวจะถูกแปลงให้เป็นสีขาว ดังรูปที่ 5.3 (b)
- ทำการสร้าง Sobel Edge Operator จากภาพต้นฉบับเพื่อตรวจสอบหาขอบภายในภาพ ดังรูปที่ 5.3 (c)
- นำภาพที่ผ่านการตรวจสอบช่วงสีผิวมาตรวจสอบหาขอบภายในภาพ พิกเซลใดที่มีสีอยู่ในช่วงสีผิวแต่อยู่บนขอบจะถูกตัดทิ้ง เพื่อทำการกรองส่วนที่เป็นผิวหนังจริง ดังรูปที่ 5.3 (d)
- ทำการตัดพื้นที่ผิวที่มีขนาดเล็กกว่าขนาดที่กำหนดไว้ออก เพื่อทำการตรวจสอบหาพื้นที่ของสีผิวที่มีขนาดใหญ่ในภาพ ดังรูปที่ 5.3 (e)



รูปที่ 5.3 (a) ภาพต้นฉบับ

(b) ภาพเมื่อผ่านการทำกระบวนการตัดสีผิว

(c) ภาพเมื่อผ่านการทำ Sobel Edge Operator กับภาพต้นฉบับ

(d) ภาพ เมื่อผ่านกระบวนการตัดสีผิวและผ่านการกรองด้วย Sobel Edge Operator

(e) ภาพเมื่อผ่านกระบวนการตัดเอาพื้นที่สีผิวที่มีขนาดเล็กออก

5.2.2 การตรวจสอบคุณสมบัติของภาพ

วิธีการทดลอง

- นำภาพที่ผ่านกระบวนการตรวจหาสีผิวของมนุษย์มาหาพื้นที่ผิวที่ต่อเนื่องกัน
- คำนวณเปอร์เซ็นต์ของจำนวนพิกเซลที่มีสีผิวทั้งหมดในภาพต่อจำนวนพิกเซลทั้งหมดของภาพ
- ทำการพิจารณาหาพื้นที่ผิวที่มีความต่อเนื่องกันซึ่งมีขนาดใหญ่ที่สุดในภาพ และคำนวณเปอร์เซ็นต์ของพื้นที่ที่ใหญ่ที่สุดต่อจำนวนพิกเซลทั้งหมดของภาพ
- นับจำนวนพื้นที่ผิวที่ต่อเนื่องกัน
- ทำการพิจารณาพื้นที่ผิวที่มีความต่อเนื่องกันซึ่งมีขนาดใหญ่ที่สุดในภาพ และคำนวณเปอร์เซ็นต์ของความสูงของพื้นที่นั้นต่อความสูงของภาพ
- ทำการพิจารณาพื้นที่ผิวที่มีความต่อเนื่องกันซึ่งมีขนาดใหญ่ที่สุดในภาพ และคำนวณ

เปอร์เซ็นต์ของความกว้างของพื้นที่นั้นต่อความกว้างของภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- ทำการพิจารณาพื้นที่ผิวที่มีความต่อเนื่องกันซึ่งมีขนาดใหญ่ที่สุดในภาพ และคำนวณหาเปอร์เซ็นต์ของพื้นที่นั้นต่อขนาดพื้นที่ที่มีสีผิวทั้งหมดในภาพ

จากการทดลองการตรวจสอบคุณสมบัติของภาพทั้ง 6 ข้อ ดังที่กล่าวในบทที่ 3 ผลที่ได้จากการทดสอบกับโปรแกรมที่พัฒนาขึ้นด้วยภาษา C# ได้ผลการทดลองดังตารางที่ 5.5

คุณสมบัติของภาพ	ภาพอนาจาร	ภาพไม่อนาจารแต่มีสีผิวเป็นส่วนประกอบ	ภาพที่ไม่มีสีผิวเป็นส่วนประกอบ
เปอร์เซ็นต์ของสีผิวในภาพ	66.7	11.59	0
เปอร์เซ็นต์ของพื้นที่สีผิวที่มีขนาดใหญ่ที่สุด	66.7	8.64	0
จำนวนชิ้นส่วนของพื้นที่ผิว	1	3	0
อัตราความสูงของพื้นที่ผิว	96.0	41.25	0
อัตราความกว้างของพื้นที่ผิว	94.0	38.33	0
อัตราส่วนของพื้นที่สีผิวขนาดใหญ่เทียบกับจำนวนพื้นที่ผิวทั้งหมด	100.0	74.53	0

ตารางที่ 5.5 ผลการทดลองการตรวจสอบคุณสมบัติของภาพ

จากตารางที่ 5.5 เป็นผลการทดลองโดยใช้ตัวอย่างภาพ 3 ภาพที่มีคุณสมบัติเป็นภาพอนาจาร, ภาพไม่อนาจารแต่มีสีผิวเป็นส่วนประกอบ (ตัวอย่างจากรูปที่ 5.3) และภาพที่ไม่มีสีผิวเป็นส่วนประกอบ จากการทดลองจะสังเกตเห็นว่าภาพอนาจารจะประกอบไปด้วยพื้นที่ของสีผิวขนาดใหญ่และครอบคลุมพื้นที่ส่วนใหญ่ของภาพ

5.2.3 การจัดประเภทของภาพด้วยคุณสมบัติ

จากการทดลองการตรวจสอบคุณสมบัติของภาพจะเห็นว่าบางคุณสมบัติไม่สามารถใช้เป็นตัวระบุได้ว่าภาพนั้นเป็นภาพอนาจารได้อย่างชัดเจน ดังนั้นจึงต้องมีการนำคุณสมบัติของภาพทั้ง 6 อย่างมาผ่านหลักการการวิเคราะห์ส่วนประกอบ (Principle Component Analysis) โดยนำคุณสมบัติทั้ง 6 มาพิจารณาร่วมกันเพื่อให้สามารถระบุภาพนั้นเป็นภาพอนาจารได้อย่างมีประสิทธิภาพ

วิธีการทดลอง

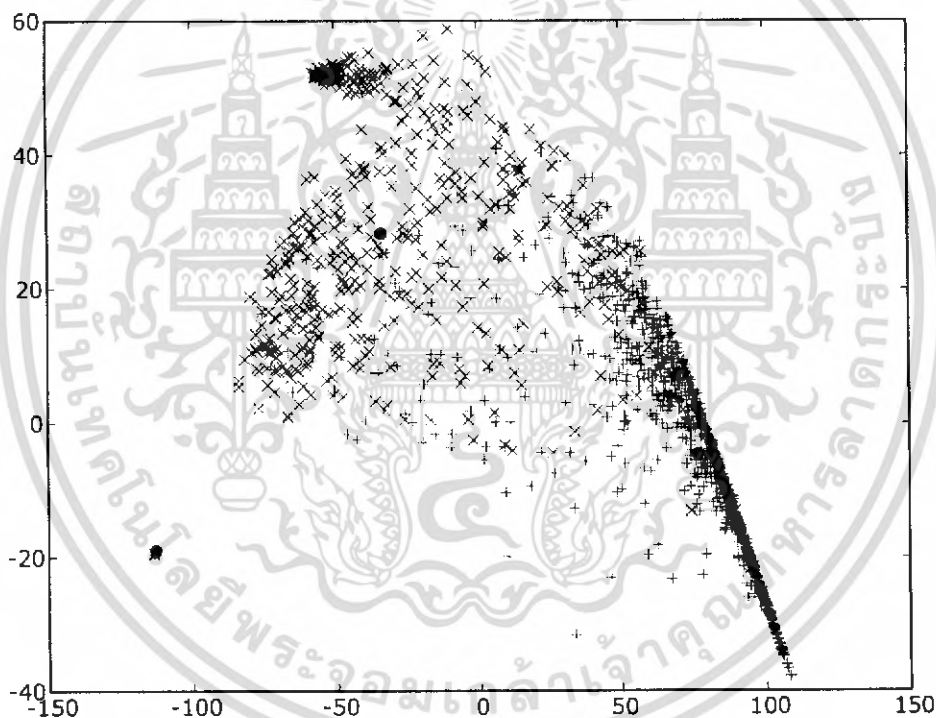
- นำตัวอย่างภาพอนาจาร 1,000 ภาพ และภาพที่ไม่ใช่ภาพอนาจารที่ประกอบด้วยภาพมนุษย์ที่ไม่ใช่ภาพอนาจาร, ภาพสัตว์, พืช, เมือง, ทิวทัศน์ และอื่นๆ 1,000 ภาพ มาผ่านกระบวนการการตรวจหาสีผิวมนุษย์และการตรวจสอบคุณสมบัติของภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- รวบรวมคุณสมบัติทั้ง 6 ข้อของภาพ 2,000 ภาพมาผ่านกระบวนการวิเคราะห์ส่วนประกอบ
- เลือกคุณสมบัติที่มีความสำคัญสูงสุด 2 คุณสมบัติที่ทำให้สูญเสียข้อมูลน้อยที่สุดมาผ่านกระบวนการจัดกลุ่ม (Clustering) ด้วยกระบวนการ K-mean เมื่อกำหนดค่า $K = 3$ เพื่อหาจุดเซ็นทรอยด์แบ่งข้อมูลเป็น 3 กลุ่ม
- ใช้ K-nearest neighbor algorithm กับทุกข้อมูลภาพเทียบกับจุดเซ็นทรอยด์

การทดลองที่ 1

จากการทดลองพัฒนาโปรแกรมการจัดประเภทของภาพด้วยคุณสมบัติด้วยภาษาไพธอนจะได้ผลดังรูปที่ 5.4



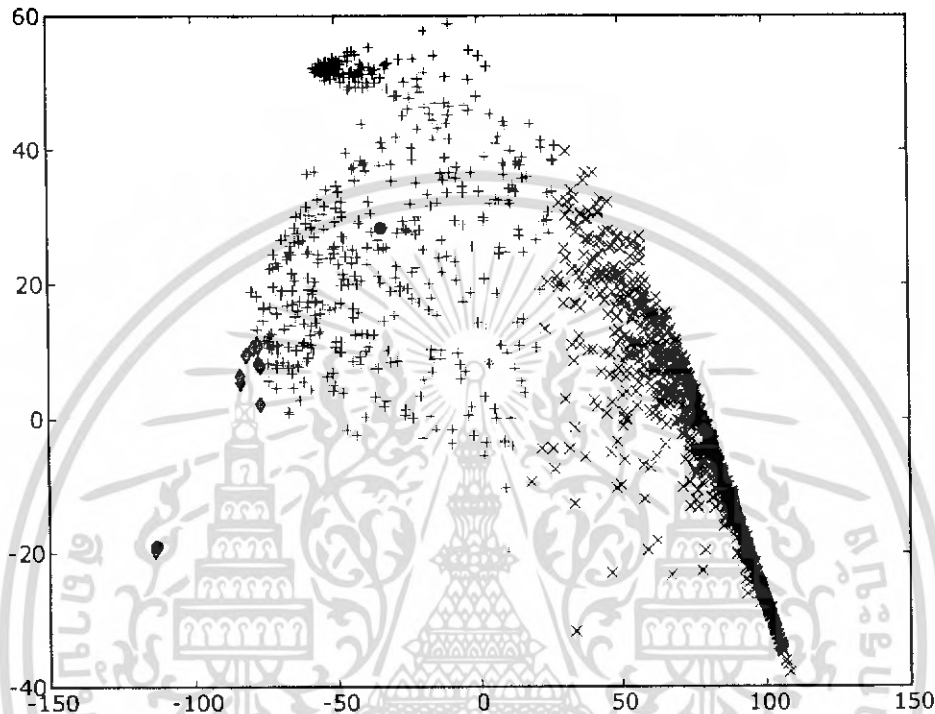
รูปที่ 5.4 กราฟของความสัมพันธ์ของคุณสมบัติของภาพ

จากรูปที่ 5.4 แสดงกราฟความสัมพันธ์ของคุณสมบัติของภาพตัวอย่าง 2,000 ภาพ โดยแบ่งออกเป็น 2 กลุ่ม ได้แก่ ภาพอนาจาร, ภาพไม่อนาจาร ซึ่งภาพอนาจารแสดงด้วยสัญลักษณ์ '+' ซึ่งรวมตัวกันอยู่ทางมุมล่างขวาของกราฟ และภาพไม่อนาจารแสดงด้วยสัญลักษณ์ 'x' ซึ่งรวมตัวกันอยู่ตรงกลางและมุมล่างขวาของกราฟ แต่ด้วยกระบวนการของการจัดกลุ่ม K-mean clustering สามารถแบ่งกลุ่มข้อมูลออกเป็น 3 กลุ่ม ได้แก่ ภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อนาจาร, ภาพไม่อนาจารแต่มีสีผิวเป็นส่วนประกอบ และภาพที่ไม่มีสีผิวเป็นส่วนประกอบ

การทดสอบการจัดกลุ่มข้อมูลด้วยตัวอย่างภาพที่นำมาวิเคราะห์หองค์ประกอบ 2,000 ภาพ ได้กราฟดังรูปที่ 5.5 และผลที่ได้ดังตารางที่ 5.6



รูปที่ 5.5 ผลการทดลองการจัดกลุ่มข้อมูลของตัวอย่างภาพ

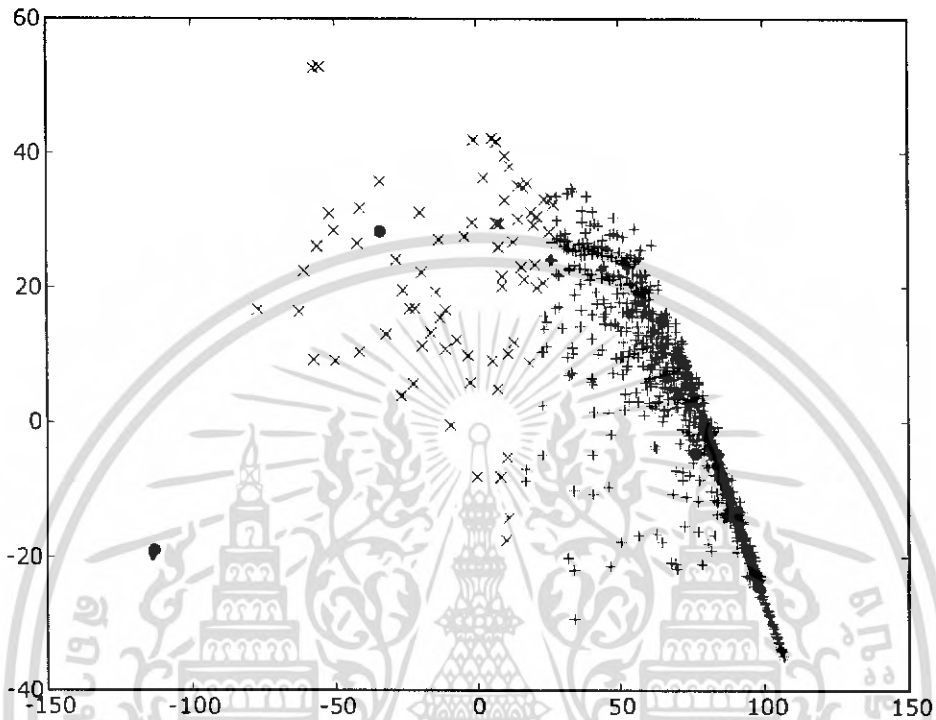
ประเภทของภาพ	จำนวนภาพ	ภาพอนาจาร	ภาพที่ไม่ใช่ภาพ อนาจาร	ประสิทธิภาพ (%)	ความผิดพลาด (%)
ภาพอนาจาร	1000	919	81	91.9	8.1
ภาพที่ไม่ใช่ภาพ อนาจาร	1000	65	935	93.5	6.5
รวม	2000	984	1016	92.7	7.3

ตารางที่ 5.6 ประสิทธิภาพของการจัดกลุ่มข้อมูลของตัวอย่างภาพ

การจัดกลุ่มข้อมูลมีการแบ่งออกเป็น 3 กลุ่ม ได้แก่ ภาพอนาจาร, ภาพที่ไม่ใช่ภาพอนาจารแต่มีสีผิวเป็นส่วนประกอบ และภาพที่ไม่มีสีผิวเป็นส่วนประกอบ แสดงดังรูปที่ 5.5 โดยภาพอนาจารแสดงแทนด้วยสัญลักษณ์ 'x', ภาพที่ไม่ใช่ภาพอนาจารแต่มีสีผิวเป็น

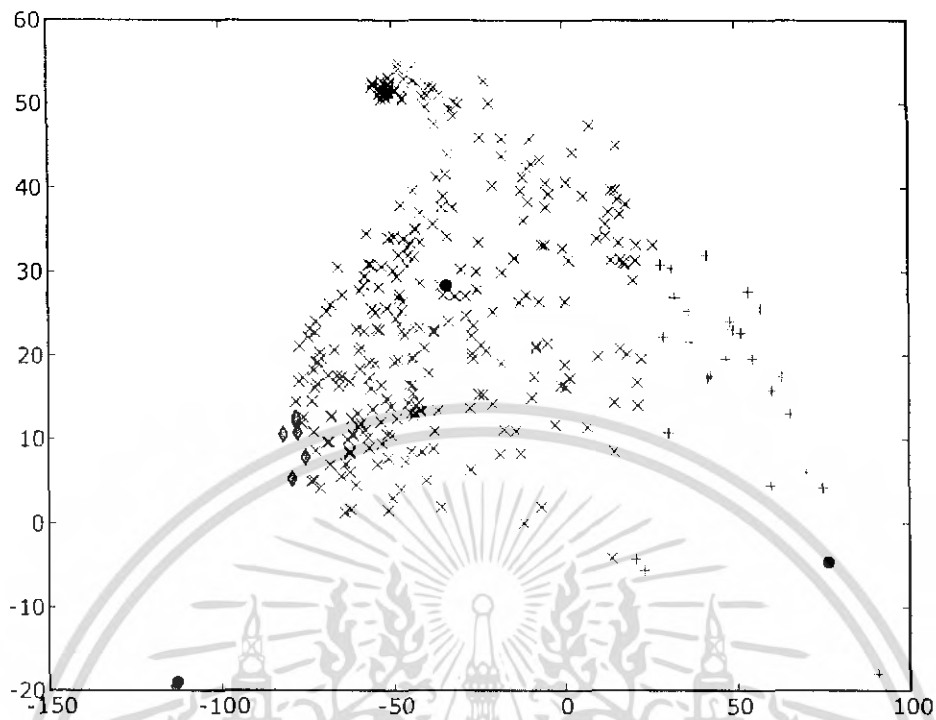
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนประกอบแสดงแทนด้วยสัญลักษณ์ '+' และภาพที่ไม่มีส่วนประกอบของสี่เหลี่ยมแสดงแทนด้วยสัญลักษณ์ '◇' และประสิทธิภาพของการทำงานแสดงดังตารางที่ 5.6
การทดลองที่ 2



รูปที่ 5.6 ผลการแบ่งกลุ่มของภาพอนาจารนอกกลุ่มตัวอย่าง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 5.7 ผลการแบ่งกลุ่มของภาพที่ไม่ใช่ภาพอนาจารนอกกลุ่มตัวอย่าง

ประเภทของภาพ	จำนวนภาพ	ภาพอนาจาร	ภาพที่ไม่ใช่ภาพ อนาจาร	ประสิทธิภาพ (%)	ความผิดพลาด (%)
ภาพอนาจาร	1000	924	76	92.4	7.6
ภาพที่ไม่ใช่ภาพ อนาจาร	500	26	474	94.8	5.2
รวม	1500	950	550	93.2	6.8

ตารางที่ 5.7 ประสิทธิภาพของการจัดกลุ่มข้อมูลของภาพนอกกลุ่มตัวอย่าง

การทดลองนี้ใช้ภาพที่ไม่ได้อยู่ในกลุ่มภาพตัวอย่างจำนวน 1,500 ภาพ ซึ่งประกอบด้วยภาพอนาจาร 1,000 ภาพ และภาพที่ไม่ใช่ภาพอนาจาร 500 ภาพ ด้วยการจัดกลุ่มโดยใช้จุดเซ็นทรอยด์ที่หาได้จากการทดลองที่ 1 โดยภาพอนาจารได้ผลดังรูปที่ 5.6 และภาพที่ไม่ใช่ภาพอนาจารได้ผลดังรูปที่ 5.7

จากรูปที่ 5.6 และ 5.7 แสดงการแบ่งกลุ่มของภาพอนาจารและภาพที่ไม่ใช่ภาพอนาจาร ภาพที่ถูกรวบรวมว่าเป็นภาพอนาจารแสดงด้วยสัญลักษณ์ '+', ภาพที่ถูกรวบรวมว่าเป็นภาพที่ไม่ใช่ภาพอนาจารแต่มีสีผิวเป็นส่วนประกอบแสดงด้วยสัญลักษณ์ 'x' และ ภาพที่ถูกรวบรวมว่าเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ส่วนประกอบของสีผิวแสดงแทนด้วยสัญลักษณ์ ‘◇’ และประสิทธิภาพการทำงานแสดงดังตารางที่ 5.7

ผลการทดสอบประสิทธิภาพ

โปรแกรมส่วนของ Image Detection สามารถตรวจสอบและคำนวณค่าน้ำหนักของรูปได้ โดยใช้เวลาเฉลี่ย 1.724 วินาทีต่อภาพ

5.3 ประสิทธิภาพการทำงาน

ทำการทดสอบประสิทธิภาพ โดยเปิดโปรแกรมคัดกรองข้อมูลเว็บ จากนั้นร้องขอข้อมูลผ่านเว็บเบราว์เซอร์ และทำการจับเวลาจนกว่า Progress bar จะสิ้นสุดการทำงาน ดังนี้

- 1) เว็บไซต์ที่ไม่สามารถนำมาทดสอบผ่านโปรแกรมทั้งหมด 100 เว็บไซต์ คัดกรองได้ อย่างถูกต้อง 97 เว็บไซต์ (97%) และผิดพลาด 3 เว็บไซต์ (3%)

จำนวนเว็บไซต์ที่ไม่สามารถ ที่คัดกรองได้ถูกต้อง (เว็บ)	โหมด Off (sec)	โหมด Medium (sec)	โหมด High (sec)
97	13.526	19.040	50.799

ตารางที่ 5.8 เวลาที่ใช้ ในการทำงานกับเว็บไซต์ที่ไม่สามารถที่คัดกรองได้ถูกต้อง

จำนวนเว็บไซต์ที่ไม่สามารถ ที่คัดกรองผิดพลาด (เว็บ)	โหมด Off (sec)	โหมด Medium (sec)	โหมด High (sec)
3	15.346	5.886	5.566

ตารางที่ 5.9 เวลาที่ใช้ ในการทำงานกับเว็บไซต์ที่ไม่สามารถที่คัดกรองผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจัย	เว็บไซต์ไม่онаจาร 100 เว็บไซต์	
	คัดกรองถูกต้อง 97 เว็บไซต์ (หน่วย)	คัดกรองผิดพลาด 3 เว็บไซต์ (หน่วย)
จำนวนเฉลี่ยรูป (IMG)	123.540	24.666
จำนวนเฉลี่ย META	4.479	5.666
จำนวนเฉลี่ยคำหาขาบ	0	2
จำนวนเฉลี่ยสคริปต์ (SCRIPT)	12.418	4
จำนวนเฉลี่ยลิงค์ทั้งหมด	155.551	73
จำนวนเฉลี่ยลิงค์รูป	31.377	18
จำนวนเฉลี่ยลิงค์ตัวหนังสือ	127.030	55
จำนวนเฉลี่ยลิงค์ภายนอกเว็บ	75.163	54.333
จำนวนเฉลี่ยลิงค์ภายในเว็บ	83.142	18.666
จำนวนเฉลี่ย PARAM	3.530	0

ตารางที่ 5.10 จำนวน Tag เฉลี่ยที่ได้จากเว็บไซต์ไม่อนาจาร 100 เว็บไซต์

- 2) เว็บไซต์ลามกอนาจาร ที่นำมาทดสอบทั้งหมด 100 เว็บไซต์ คัดกรองได้อย่างถูกต้อง 86 เว็บไซต์ (86%) และผิดพลาด 14 เว็บไซต์ (14%)

จำนวนเว็บไซต์อนาจาร ที่คัดกรองได้ถูกต้อง (เว็บ)	โหมด Off (sec)	โหมด Medium (sec)	โหมด High (sec)
86	22.240	10.747	10.467

ตารางที่ 5.11 เวลาที่ใช้ ในการทำงานกับเว็บไซต์อนาจารที่คัดกรองได้ถูกต้อง

จำนวนเว็บไซต์อนาจาร ที่คัดกรองผิดพลาด (เว็บ)	โหมด Off (sec)	โหมด Medium (sec)	โหมด High (sec)
14	27.44	27.824	76.732

ตารางที่ 5.12 เวลาที่ใช้ ในการทำงานกับเว็บไซต์อนาจารที่คัดกรองผิดพลาด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปัจจัย	เว็บไซต์อาจารย์ 100 เว็บไซต์	
	คัดกรองถูกต้อง 86 เว็บไซต์ (หน่วย)	คัดกรองผิดพลาด 14 เว็บไซต์ (หน่วย)
จำนวนเฉลี่ยรูป (IMG)	81.094	250.071
จำนวนเฉลี่ย META	5.305	1.714
จำนวนเฉลี่ยคำหาบ	4	0
จำนวนเฉลี่ยสคริปต์ (SCRIPT)	4.705	4.714
จำนวนเฉลี่ยลิงก์ทั้งหมด	462.305	350.5
จำนวนเฉลี่ยลิงก์รูป	36.976	42.5
จำนวนเฉลี่ยลิงก์ตัวหนังสือ	429.894	308
จำนวนเฉลี่ยลิงก์ภายนอกเว็บ	140.894	52.071
จำนวนเฉลี่ยลิงก์ภายในเว็บ	327.317	298.428
จำนวนเฉลี่ย PARAM	3.364	0

ตารางที่ 5.13 จำนวน Tag เฉลี่ยที่ได้จากเว็บไซต์อาจารย์ 100 เว็บไซต์

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 6

บทวิจารณ์และสรุป

6.1 บทสรุป

โครงการนี้เป็นโครงการใหม่ที่ได้จัดทำขึ้นมา โดยประกอบด้วย 2 ส่วนหลักๆ ดังนี้

1. Port Forwarding ทำหน้าที่เป็นตัวกลางในการส่งผ่านข้อมูลระหว่างทางฝั่งผู้ใช้เว็บเบราว์เซอร์ (Web Browser) กับทางฝั่งเว็บเซิร์ฟเวอร์ (Web Server)
2. Filtering ทำหน้าที่ตรวจสอบข้อมูลเนื้อหาบนหน้าเว็บไซต์ที่ได้รับจากทางฝั่งเว็บเซิร์ฟเวอร์ ก่อนที่จะส่งต่อไปยังฝั่งผู้ใช้เว็บเบราว์เซอร์ โดยการกรองข้อมูลเนื้อหาบนหน้าเว็บเพจนี้ใช้เทคนิคต่างๆ คือ
 - 2.1. Image Detection ซึ่งเป็นส่วนที่นำภาพที่หามาได้ นำมาผ่านกระบวนการประมวลผลของภาพ เพื่อตรวจสอบว่าภาพนั้นๆ เป็นภาพอนาจารหรือไม่ จากการทดลองพบว่า ยังมีความผิดพลาดอยู่ ในส่วนของกระบวนการตัดสีผิว ซึ่งอาจมีความผิดพลาดเกิดขึ้นเนื่องจาก
 - ภาพมีคุณภาพต่ำ
 - ภาพมีวัตถุที่มีสีใกล้เคียงกับสีผิวมนุษย์ เช่น ภาพของพืชที่มีสีเขียวเหลือง, ภาพคนที่ใส่เสื้อที่มีสีใกล้เคียงกับสีผิว, ภาพใบหน้าคนขนาดใหญ่
 - ภาพผิวหนังที่มีสีผิวจางเนื่องจากการส่องสว่างและการสะท้อนของแสง
 - ภาพขนาดใหญ่ที่ประกอบขึ้นจากภาพขนาดเล็กหลายภาพ
 อย่างไรก็ตามประสิทธิภาพการทำงานของโปรแกรมตรวจสอบภาพอนาจารเป็นที่น่าพอใจ ซึ่งสามารถจำแนกความแตกต่างระหว่างภาพอนาจารและภาพที่ไม่ลามกอนาจารได้
 - 2.2 Text Detection จะทำการตรวจสอบ HTML Code เพื่อทำการหา Tag และเนื้อความต่างๆ ที่จำเป็นในการคัดกรองข้อมูลเว็บนำมาตรวจสอบ ซึ่งอาจมีความผิดพลาดได้เนื่องจากเว็บไซต์ที่ไม่ใช่เว็บไซต์ที่ลามกอนาจารบางเว็บไซต์ มีคุณสมบัติของเนื้อหาบางอย่างคล้ายคลึงกับเว็บไซต์ลามกอนาจาร

6.2 ปัญหาอุปสรรคและแนวทางในการแก้ไข

1. การใช้งานขึ้นอยู่กับความเร็วของอินเทอร์เน็ต และความเร็วของเครื่องคอมพิวเตอร์ที่ใช้
2. การคัดกรองรูป ถ้าเว็บเพจไหนมีรูปมากๆ และมีขนาดใหญ่จะทำให้การทำงานช้าลงไปมาก และอาจเกิดความผิดพลาดได้ หากรูปนั้นมีพื้นที่ที่คล้ายสีผิวมากและอยู่รวมเป็นกลุ่มกัน เช่น
 - โปรแกรมจะตรวจสอบสีผิวในรูปภาพได้เฉพาะช่วงของสีผิวที่ทำการกำหนดไว้เท่านั้น
 - โปรแกรมอาจเกิดความผิดพลาดเนื่องจากภาพที่ไม่ใช่ภาพลามกอนาจาร แต่มีสีอยู่ในช่วง ของสีผิว เช่น ภาพใบหน้าคน, ภาพผลไม้, ภาพวิวทะเลทราย
3. การหาสถิติต่างๆยังมีข้อผิดพลาดหรือคลาดเคลื่อนไป เพราะอันเนื่องมาจาก
 - 3.1 ในเว็บเพจแต่ละหน้ามีโครงสร้างที่หลากหลายทำให้จำแนกได้ยากกว่าเป็นเว็บไซต์ลามกอนาจารหรือไม่ ทำให้โปรแกรมอาจเกิดความผิดพลาดได้ เนื่องจากเนื้อหาในเว็บไซต์ที่ไม่ใช่เว็บไซต์ลามกอนาจารบางเว็บไซต์ อาจมีคุณสมบัติใกล้เคียงกับเว็บไซต์ลามกอนาจาร
 - 3.2 กรณีการพิจารณาคำศัพท์ว่าเป็นคำที่ไม่เหมาะสมหรือไม่ จะเกิดข้อผิดพลาดก็ต่อเมื่อ
 - คำๆ นั้นตรงกับชื่อเฉพาะต่างๆ เช่น ชื่อเมือง ชื่อคน
 - คำๆ นั้นตรงกับเนื้อหาที่อยู่ในเว็บ Wikipedia หรือประเภท Dictionary
 - คำๆ นั้นตรงกับเนื้อหาที่อยู่ในเว็บเกี่ยวกับเพศศึกษา ซึ่งเราสามารถแก้ปัญหาในส่วนนี้ได้โดยมีการเพิ่มชื่อชื่อ URL ที่น่าเชื่อถือ (Trusted URL) ลงในโปรแกรม และไม่กรองเว็บไซต์ที่ประกอบไปด้วย Domain Name ที่ได้รับความน่าเชื่อถือ อาทิ เช่น .edu, .ac.th, .co.th, .go.th, .gov เป็นต้น

6.3 แนวทางการพัฒนาต่อ

1. ในอนาคตอาจมีการพัฒนาให้มีประสิทธิภาพมากยิ่งขึ้น โดยสามารถประยุกต์ใช้งานร่วมกับโครงการระบบตรวจหาข้อมูลอันตรายในเวิร์ลไวด์เว็บ (World Wide Web Malefic Code Inspection System) ซึ่งเป็นระบบที่ทำการค้นหารายชื่อเว็บไซต์ที่มีความเสี่ยงซึ่งเปรียบเสมือนแหล่งเก็บข้อมูล หรือฐานข้อมูลเกี่ยวกับเว็บไซต์อันตรายต่างๆ
2. ปรับปรุงการทำงานของโปรแกรมให้มีการทำงานที่รวดเร็วยิ่งขึ้น และลดการใช้ทรัพยากรบนระบบได้มากขึ้น
3. ปรับปรุงค่าที่เก็บอยู่ใน Registry ให้มีการเข้ารหัสข้อมูล เพื่อตรวจสอบไม่ให้ผู้ใช้ที่ไม่ได้รับอนุญาต มาทำการแก้ไขข้อมูลใน Registry อาทิ เช่น รหัสผ่าน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. พัฒนาให้สามารถตรวจสอบได้ว่า Log files มีการถูกแก้ไขหรือไม่
5. พัฒนาให้สามารถรองรับการทำงานได้ในหลายๆ โปรโตคอล เช่น https และ ftp



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บรรณานุกรม

- [1] RFC Documentation [online]. Available:
<http://www.freesoft.org/CIE/RFC/2068/index.htm>
- [2] MSDN Documentation [online]. Available:
<http://msdn2.microsoft.com/en-us/default.aspx>
- [3] The Code Project - Free Source Code and Tutorials [online]. Available:
<http://www.codeproject.com>
- [4] The number one the number one developer site [online]. Available:
<http://www.codeguru.com/>
- [5] Rafael C. Gonzalez / Richard E. Woods, Digital Image Processing, Prentice Hall PTR
- [6] Andrew Krowczk / Vinod Kumar / Nauman Laghari / Ajit Mungse / Christian Nagel / Tim Parker / Srinivasa Sivakumar , Professional .NET Network Programming, WROX Press
- [7] W.H Ho, P.A. Watters “Identifying and Blocking Pornographic Content”
- [8] ASP.Net examples (example source code) Organized by topic [online]. Available:
<http://www.java2s.com>
- [9] C# Corner- C-Sharp C#.NET CSharp VB.NET ASP.NET Visual Studio .NET [online]. Available:<http://www.c-sharpcorner.com/>
- [10] Christophe Garcia, Georgios Tziritas “Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis”, IEEE Trans. Multimedia, vol.1, no.3, PP. 264-277, September 1999.
- [11] Rein-Lien Hsu, Mohamed Abdel-Mottaleb, Anil K. Jain “Face Detection in Color Images”, IEEE Transaction on pattern analysis and machine intelligence, Vol. 24 No.5 May 2002
- [12] Qing-Fang Zheng , Wei Zeng , Gao Wen , Wei-Qiang Wang “Shape-based Adult Image Detection”, Proceedings of the Third International Conference on Image and Graphics (ICIG’04)
- [13] Feng Jiao, Wen Gao, Lijuan Duan, Guoqin Cui “Detecting Adult Image Using Multiple Features” ,The Institute of Computing Technology, Chinese Academy of Science, Beijing 100080



ภาคผนวก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



ภาคผนวก ก.

สถิติของเว็บไซต์ทั้งหมดที่นำมาทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ก.

สถิติของเว็บไซต์ทั้งหมดที่นำมาทดสอบ

สถิติของเว็บไซต์ทั้งหมดที่นำมาทดสอบ

1) เว็บไซต์ไม่ลามกอนาจาร ที่นำมาทดสอบผ่านโปรแกรมทั้งหมด 100 เว็บไซต์ คัดกรองได้อย่างถูกต้อง 97 เว็บไซต์ (97%) และผิดพลาด 3 เว็บไซต์ (3%) มีสถิติค่าต่างๆ ดังนี้

ตารางที่ ก.1 เวลาที่ใช้ ในการเปิดเว็บไซต์ (ไม่อนาจาร) เมื่อคัดกรองได้อย่างถูกต้อง

รายชื่อเว็บไซต์ (ไม่ลามกอนาจาร) ที่คัดกรองได้อย่างถูกต้อง 97 เว็บไซต์	เวลาที่ใช้ เมื่อไม่ผ่าน โปรแกรม (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด High (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด Medium (วินาที)
http://en.wikipedia.org/wiki/Wiki	14.62	30.85	23.84
http://hopeanuoli.cjb.net/	8.85	14.55	10
http://load-info.us/	3.75	5.44	6
http://offthemat.net/	3.23	4.5	3.64
http://postjung.com/	4.48	41.5	12.89
http://teenee.com/	24.98	69.2	24.33
http://top.greek-sites.eu	27.28	68.87	36.62
http://top-online.us/	4.28	5.97	5.07
http://www.212cafe.com/	4.96	54.98	15.17
http://www.365jukebox.com/	7.19	57.99	55.12
http://www.abitur-werne.de/blog	15.74	21.01	12.93
http://www.aios.org/	18.7	95.63	33.74
http://www.a-jokes.com/	6.67	6.51	6.7
http://www.amazon.com/	30.11	48.43	33.03
http://www.appjob.com/	29.92	325.24	38.53
http://www.apple.com/	8.01	24.71	12.13
http://www.asiasoft.co.th/	2.55	2.68	2.94
http://www.bbc.co.uk/thai/	10.11	12.6	12.67
http://www.bbznnet.com/	3.82	26.75	2.92
http://www.bloggang.com/	4.6	60.87	9.93
http://www.cnn.com/	19.83	37.76	26.62
http://www.cool50.eu/	33.8	68.99	38.62
http://www.craftercity.com/	17.32	20.68	17.62
http://www.dailynews.co.th/	6.99	8.28	8.23
http://www.defendingarnhem.com/	27.62	102.12	24.37
http://www.dek-d.com/	9.32	107.85	16.46
http://www.doctorsan.com/	6.96	49.39	23.05
http://www.downloadget.com/	19.29	14.57	21.69
http://www.elovedates.com/	32.79	120.36	53.1
http://www.exteen.com/	7.25	12.1	9.33
http://www.filevalley.com/	6.24	11.73	7.61
http://www.freeflashonlinegames.com/	19.31	42.28	27.94
http://www.freehitwebcounters.com/	43.5	100.75	31.45
http://www.freehitwebsitecounter.com/	44.94	136.41	62.38
http://www.free-palm-soft.com/	62.66	264.22	42.97
http://www.free-pocket-pc-games.com/	25.82	114.46	36.75

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ได้รับอนุญาต

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายชื่อเว็บไซต์ (ไม่ลามกอนาจาร) ที่คัดกรองได้อย่างถูกต้อง (ต่อ)	เวลาที่ใช้ เมื่อไม่ผ่าน โปรแกรม (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรม High (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรม Medium (วินาที)
http://www.freeproxyserverlist.com/	6.99	28.74	10.56
http://www.gmember.com/	2.03	2.43	2.56
http://www.google.com/	2.12	2.67	2.13
http://www.gr50.com/	26.47	187.14	24.02
http://www.greekmeeting.com/	10.19	30.48	14.55
http://www.hartl.info/	14.37	16.67	13.16
http://www.hunsa.com/2005/	7.88	35.56	14.67
http://www.ini3.co.th/	4.37	7.67	4.56
http://www.jabchai.com/	18.43	165.67	33.37
http://www.kapook.com/	12.36	28.82	21.05
http://www.kirstiealleyfan.com/	5	7.36	8
http://www.komchadluek.net/	15.05	142.33	31.19
http://www.lyricsmusicongs.com/	13.32	46.83	12.1
http://www.lyricspedia.com/	13.28	18.16	12.85
http://www.manager.co.th/	10.43	12.82	8.25
http://www.marketathome.com/	20.17	223.27	48.89
http://www.matichon.co.th/	7.04	7.15	7.06
http://www.mcot.net/	12.11	49.69	26.47
http://www.meemodel.com/home/	9.35	45.73	15.11
http://www.meesook.com/	9.94	84.85	26.25
http://www.moohin.com/	11.69	38	22.05
http://www.mthai.com/	11.1	45.95	23.27
http://www.musicatm.com/	6.46	57.37	21.55
http://www.narak.com/	5.25	42.2	8.65
http://www.pantip.com/	11.47	74.67	22.01
http://www.pantipmarket.com/	3.48	33.69	15.24
http://www.passaparolach.com/	44.29	150.98	42.94
http://www.pixiart.com/	6.87	22.94	21.12
http://www.play-bit.com/	17.62	53.97	17.42
http://www.playpark.com/	1.46	3.24	2.42
http://www.popcornfor2.com/	11.1	49.76	16.71
http://www.pramool.com/	4.03	13.98	6.31
http://www.rheuma.de/	13.6	12.85	18.19
http://www.sangabrielmission.org/	13.58	25	21.6
http://www.sanook.com/	10.98	30.45	24.93
http://www.scandicon.net/	2.16	19.12	2.89
http://www.secretsofredgate.com/	12.5	13.87	10.62
http://www.sf.in.th/intro/	3.92	5.6	2.44
http://www.short-love-poem.com/	28.71	34.5	30.19
http://www.siamcomic.com/	14.82	35.05	16.73
http://www.siamphone.com/	2.3	2.4	2.44
http://www.siamsport.co.th/	11.96	19.23	12
http://www.siamza.com/home/	12.48	58.85	20.45
http://www.siamzone.com/	5.81	61.93	9.42
http://www.softwarewin.com/	23.67	38.94	24.43
http://www.storythai.com/	9.75	29.06	10.4
http://www.swisswatchstation.com/	13.58	23.22	15.65
http://www.tarad.com/	10.6	34.23	12.08
http://www.thaibq.com/	4.44	25.1	5.72

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายชื่อเว็บไซต์ (ไม่ลามกอนาจาร) ที่คัดกรองได้อย่างถูกต้อง (ต่อ)	เวลาที่ใช้ เมื่อไม่ผ่าน โปรแกรม (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด High (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด Medium (วินาที)
http://www.thailand.com/	7.75	21.12	8.2
http://www.ThaiSecondhand.com/	24.5	129.29	66.33
http://www.thaiware.com/	8.16	108.46	18.67
http://www.timber-merchants-importers.com/	12.81	34.46	26.62
http://www.toolshed-ps.com/	10.23	22.09	11.46
http://www.truehits.net/	6.1	11.12	8.44
http://www.ttonline.net/	13.62	85.19	15.23
http://www.uploadtoday.com/	2.48	5.48	7.5
http://www.vadino.com/	16.85	45.98	21.72
http://www.vividroom.com.au/	23.97	28.74	27.37
http://www.void.be/	4.25	4.28	3.55
http://www.w4lqh.com/	3.42	5.3	3.35
http://www.yenta4.com/	24.41	86.42	44.44
เฉลี่ย	13.52	50.79	19.04

ตารางที่ ก.2 เวลาที่ใช้ ในการเปิดเว็บไซต์ (ไม่อนาจาร) เมื่อคัดกรองผิดพลาด

รายชื่อเว็บไซต์ (ไม่ลามกอนาจาร) ที่คัดกรองผิดพลาด 3 เว็บ	เวลาที่ใช้ เมื่อไม่ผ่าน โปรแกรม (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด High (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด Medium (วินาที)
http://www.boxsat.com/	11.94	8.92	9.75
http://www.dellawear.de/	30.49	5	5.35
http://www.t-scripts.com/	3.61	2.78	2.56
เฉลี่ย	15.34	5.56	5.88

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.3 จำนวน Tag ต่างๆ ของเว็บไซต์ (ไม่онаจาร) ที่คัดกรองได้อย่างถูกต้อง

รายชื่อเว็บไซต์ (ไม่онаจาร) คัดกรองได้อย่างถูกต้อง 97 เว็บ	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://en.wikipedia.org/wiki/Wiki	15	3	0	9	318	15	303	125	193	0
http://hopeanuoli.cjb.net/	0	0	0	3	6	0	6	5	1	0
http://load-info.us/	1	2	0	0	202	1	201	3	199	0
http://offthemat.net/	2	10	0	0	6	1	5	2	4	0
http://postjung.com/	90	3	0	12	158	40	118	77	81	0
http://teenee.com/	377	1	0	8	340	17	323	326	4	8
http://top.greek-sites.eu	69	7	0	1	122	40	82	71	51	0
http://top-online.us/	2	2	0	1	202	1	201	3	199	0
http://www.212cafe.com/	65	3	0	10	104	16	88	46	58	4
http://www.365jukebox.com/	106	9	0	6	445	41	404	177	268	14
http://www.abitur-nerne.de/blog	2	2	0	16	152	2	150	94	58	0
http://www.aios.org/	119	1	0	1	53	11	42	13	40	0
http://www.a-jokes.com/	6	2	0	7	27	5	22	12	15	0
http://www.amazon.com/	62	2	0	5	283	50	233	3	280	0
http://www.appjob.com/	392	5	0	5	402	121	281	129	273	0
http://www.apple.com/	29	6	0	11	34	23	11	3	31	0
http://www.asiasoft.co.th/	34	1	0	5	12	3	9	12	0	0
http://www.bbc.co.uk/thai/	62	8	0	4	58	17	41	33	25	0
http://www.bbznet.com/	36	3	0	2	18	5	13	12	6	0
http://www.bloggang.com/	17	2	0	14	195	5	190	70	125	0
http://www.cnn.com/	259	5	0	33	259	70	189	32	227	0
http://www.cool50.eu/	113	6	0	1	238	62	176	138	100	0
http://www.craftercity.com/	70	7	0	0	150	58	92	2	148	0
http://www.dailynews.co.th/	232	8	0	16	167	62	105	4	163	0
http://www.defendingarnhem.com/	21	3	0	1	38	16	22	7	31	0
http://www.dek-d.com/	421	1	0	13	453	72	381	146	307	0
http://www.doctorsan.com/	256	7	0	13	241	43	198	83	158	4
http://www.downloadget.com/	18	6	0	9	47	7	40	24	23	0
http://www.elovedates.com/	104	8	0	20	168	101	67	128	40	0
http://www.exteen.com/	31	1	0	0	63	1	62	55	8	2
http://www.filevalley.com/	34	4	0	15	136	13	123	22	114	0
http://www.freeflashonlinegames.com/	50	4	0	24	74	12	62	4	70	0
http://www.freehitwebcounters.com/	70	2	0	28	90	17	73	15	75	0
http://www.freehitwebsitecounter.com/	60	2	0	27	88	10	78	13	75	0
http://www.free-palm-soft.com/	31	4	0	45	126	19	107	2	124	0
http://www.free-pocket-pc-games.com/	25	4	0	42	58	12	46	2	56	0
http://www.freeproxyserverlist.com/	13	4	0	42	15	0	15	1	14	0
http://www.gmember.com/	211	5	0	12	100	26	74	59	41	45
http://www.google.com/	0	1	0	0	1	0	1	1	0	0
http://www.gr50.com/	7	6	0	0	17	6	11	17	0	0
http://www.greekmeeting.com/	17	6	0	1	25	16	9	25	0	0
http://www.hartl.info/	9	20	0	5	40	9	31	25	15	0
http://www.hunsa.com/2005/	287	3	0	8	376	45	331	335	41	0
http://www.ini3.co.th/	212	3	0	9	62	18	44	39	23	0
http://www.jabchai.com/	196	4	0	9	234	94	140	180	54	11
http://www.kapook.com/	294	3	0	13	31	111	200	292	19	4
http://www.kirstiealleyfan.com/	0	4	0	0	0	0	0	0	0	0
http://www.komchadluek.net/	399	4	0	8	186	86	100	71	115	55
http://www.lyricsmusicongs.com/	11	4	0	44	63	0	63	1	62	0

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายชื่อเว็บไซต์(ไม่มาจากร) คัดกรองได้อย่างถูกต้อง (ต่อ)	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://www.manager.co.th/	871	1	0	15	600	172	428	142	458	9
http://www.marketathome.com/	228	5	0	8	287	50	237	107	180	46
http://www.matichon.co.th/	160	4	0	5	132	51	81	10	122	23
http://www.mcot.net/	231	3	0	9	132	62	70	79	53	11
http://www.meemodel.com/home/	80	3	0	2	263	9	254	261	2	4
http://www.meesook.com/	119	2	0	19	386	29	357	271	115	2
http://www.moohin.com/	364	0	0	7	301	45	256	235	66	2
http://www.mthai.com/	200	6	0	14	442	88	354	393	49	4
http://www.musicatm.com/	133	2	0	15	81	20	61	31	50	6
http://www.narak.com/	228	6	0	5	243	65	178	102	141	4
http://www.pantip.com/	99	1	0	19	296	81	215	146	150	2
http://www.pantipmarket.com/	124	1	0	6	341	40	301	331	10	2
http://www.passaparolach.com/	56	2	0	9	172	31	141	169	3	0
http://www.pixiart.com/	13	6	0	115	291	7	284	54	237	0
http://www.play-bit.com/	43	4	0	47	109	28	81	1	108	0
http://www.playpark.com/	14	1	0	4	8	3	5	4	4	0
http://www.popcornfor2.com/	78	1	0	4	71	11	60	26	45	4
http://www.pramool.com/	199	3	0	4	136	29	107	67	69	0
http://www.rheuma.de/	20	7	0	3	19	12	7	3	16	0
http://www.sangabrielmission.org/	24	0	0	16	38	20	18	1	37	0
http://www.sanook.com/	250	5	0	24	375	40	335	343	32	0
http://www.scandicon.net/	21	1	0	3	5	1	4	0	5	0
http://www.secretsofredgate.com/	6	3	0	3	20	2	18	1	19	0
http://www.sf.in.th/intro/	14	1	0	4	8	3	5	4	4	0
http://www.short-love-poem.com/	25	4	0	21	75	15	60	60	15	0
http://www.siamcomic.com/	540	1	0	7	337	63	274	57	280	4
http://www.siamphone.com/	7	8	0	7	31	4	27	26	5	0
http://www.siamsporl.co.th/	323	7	0	24	265	70	195	21	244	34
http://www.siamza.com/home/	591	3	0	33	279	50	229	212	67	0
http://www.siamzone.com/	180	3	0	14	284	81	203	5	279	3
http://www.softwarewin.com/	89	13	0	26	161	23	138	22	139	0
http://www.storythai.com/	86	6	0	16	111	11	100	96	15	0
http://www.swisswatchstation.com/	34	24	0	31	51	11	40	20	31	0
http://www.tarad.com/	353	9	0	14	367	108	259	205	162	0
http://www.thaibg.com/	62	5	0	6	98	17	81	7	91	3
http://www.thaiload.com/	21	2	0	2	138	0	138	14	124	8
http://www.ThaiSecondhand.com/	112	3	0	4	331	53	278	286	45	3
http://www.thaiware.com/	246	11	0	12	134	38	96	53	81	6
http://www.timber-merchants-importers.com/	52	3	0	1	75	8	67	38	37	0
http://www.toolshed-ps.com/	22	6	0	9	31	14	17	3	28	0
http://www.truehits.net/	96	2	0	3	148	18	130	141	7	6
http://www.ttonline.net/	349	18	0	6	243	27	216	116	127	10
http://www.uploadtoday.com/	33	3	0	9	12	4	8	5	7	0
http://www.vadino.com/	61	5	0	27	110	57	53	40	70	0
http://www.vividroom.com.au/	21	6	0	3	41	11	30	2	39	0
http://www.void.be/	2	3	0	0	13	0	13	8	5	0
http://www.w4lgh.com/	1	4	0	0	3	1	2	0	3	0
http://www.yenta4.com/	274	7	0	17	395	91	304	207	188	3
เฉลี่ย	123.54	4.47	0	12.41	155.55	31.37	127.03	75.16	83.14	3.53

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.4 จำนวน Tag ต่างๆ ของเว็บไซต์ (ไม่อนาจาร) ที่คัดกรองผิดพลาด

รายชื่อเว็บไซต์ (ไม่อนาจาร) คัดกรองผิดพลาด 3 เว็บ	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://www.boxsat.com/	58	7	4	11	189	45	144	140	49	0
http://www.dellawear.de/	14	8	1	1	9	7	2	2	7	0
http://www.t-scripts.com/	2	2	1	0	21	2	19	21	0	0
เฉลี่ย	24.66	5.66	2	4	73	18	55	54.33	18.66	0



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2) เว็บไซต์ลามกอนาจาร ที่นำมาทดสอบทั้งหมด 100 เว็บไซต์ คัดกรองได้อย่างถูกต้อง 86 เว็บไซต์ (86%) และผิดพลาด 14 เว็บไซต์ (14%) มีสถิติค่าต่างๆ ดังนี้

ตารางที่ ก.5 เวลาที่ใช้ ในการเปิดเว็บไซต์ (อนาจาร) เมื่อคัดกรองได้อย่างถูกต้อง

รายชื่อเว็บไซต์ (ลามกอนาจาร) ที่คัดกรองได้อย่างถูกต้อง 86 เว็บไซต์	เวลาที่ใช้ เมื่อไม่ ผ่านโปรแกรม (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมใหม่ High (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมใหม่ Medium (วินาที)
http://clothing-mania.com/	52.85	6.6	3.16
http://ohgood.com/	7.35	13.31	14.51
http://oh-teen.com/	5.2	10.86	10.25
http://spunkysheets.com/	5.75	8.47	9.51
http://straight.theater.moviemonster.com	70.19	33.65	28.43
http://www.247pantyhose.com/	12.82	15.56	12.28
http://www.adult-series.com/	20.45	7.98	5.57
http://www.ah-me.com/	14.48	13.18	15.16
http://www.apieceofass.com/	6.19	3.59	3.78
http://www.aroundtheworldsex.com/	12.48	5.15	5.91
http://www.asstraffic.com/	8.04	4.25	3.94
http://www.babesglamour.com/	55.14	10.34	8.64
http://www.bestxxxseries.com/	20.62	24.28	26.46
http://www.bizarre-rituals.com/	14.17	17.1	17.3
http://www.blogchalking.com/life/psm/forced-big-cock-sex/	15.37	3.67	4.83
http://www.bootae.com/	48.25	11.55	8.24
http://www.candylist.com/	58.62	19.4	22.87
http://www.chubbyland.com/	13.5	5.61	6.12
http://www.cleancafe.com/	19.69	13.19	13.86
http://www.crazynakedchick.com/	13.94	13.08	9.98
http://www.crazy-series.com/	72.04	7.32	5.71
http://www.croseseries.com/	16.19	8	5.5
http://www.dariusporn.com/	33.93	12.24	14.6
http://www.el-ladies.com/	73.72	6.53	5.94
http://www.eurothumb.com/	8.4	4.54	7.9
http://www.fmgalleries.com/	13.09	6.72	8.7
http://www.footster.net/	7.28	5.64	3.87
http://www.freebigmovies.com/	28.64	15.42	12.69
http://www.freepicseries.com/	15.1	10.17	11.29
http://www.freepornpictures.org/	2.93	2.78	3.7
http://www.frogsex.com/	10.3	9.21	8.21
http://www.fuckcat.com/	5.01	8.76	9.92
http://www.fullvidz.com/	5.89	18.24	16.51
http://www.greenthumbstgp.com/	4.4	4.29	4.48
http://www.hankshoneys.com/	13.41	5.63	9.9
http://www.hardsex-movies.com/	60.99	8.62	6.4
http://www.horny-grannysex.com/	14.19	11.97	12.31
http://www.hotfreeesex4all.com/	63.97	11.12	9.87
http://www.hotfunhouse.com/	14.49	18.31	17.56
http://www.huggernet.com/	17.2	6.24	5.3
http://www.imagefest.com/	7.43	9.21	8.97
http://www.infospiral.com/category/mature/	9.86	9.38	6.19

เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์ไว้เพื่อการศึกษาเท่านั้น ไม่อนุญาตให้เผยแพร่โดยไม่ได้รับอนุญาตจากทางราชการ

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายชื่อเว็บไซต์ (ตามกอนาจาร) ที่คัดกรองได้อย่างถูกต้อง (ต่อ)	เวลาที่ใช้ เมื่อไม่ ผ่านโปรแกรม	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด
	(วินาที)	High (วินาที)	Medium (วินาที)
http://www.ixixi.com/	12	17.16	22.08
http://www.jennysbookmarks.com/	7.95	7.34	7.37
http://www.justseries.com/	2.69	2.6	3.59
http://www.legs-fetish.com/	22.12	3.67	4.53
http://www.legs-fetish.com/thumbnailpost/	23.49	15.27	13.5
http://www.leg-zone.com/	5.92	5.6	6.35
http://www.mature.dk/	25.04	3.79	3.49
http://www.meoland.com/	43.04	35	31.35
http://www.nslit.com/	26.62	18.16	14.5
http://www.old69.com/	12.68	20.34	17.4
http://www.onefullmovie.com/	3.45	2.98	2.6
http://www.pandalist.com/	10.37	2.19	4.3
http://www.pinkvisualtgp.com/	24.95	4.54	5.19
http://www.pornfoundation.com/	5.1	12.35	13.92
http://www.pornpixxxs.com/	25.78	12.18	10.68
http://www.premiumxxxcore.com/	4.78	2.5	2.25
http://www.pued.com/	32.63	13.3	11.19
http://www.pushin pussy.com/	17.94	12.5	12.78
http://www.pussydivine.com/	8.94	3.2	2.46
http://www.realxworld.com/	11.3	15.51	10.91
http://www.series-xxx.com/	1.33	2.55	2.15
http://www.sicklist.com/	17.1	22.4	25.54
http://www.smutcastle.com/	3.55	4.02	2.62
http://www.spankstation.com/	3.5	12.18	11.17
http://www.stocking-mania.com/	4.76	3.26	3.27
http://www.stocking-movies.com/	14.82	10.4	6.53
http://www.stockings-world.com/	21.81	7	10.2
http://www.stocking-tease.com/	208.12	22.89	24.54
http://www.teeniesxxx.com/	52.12	9.75	8.8
http://www.tgplibrary.com/	4.23	5.37	2.75
http://www.thumbzilla.com/	3.8	3.53	3.31
http://www.tiava.com/	14.2	13.02	14
http://www.tommys-bookmarks.com/	9.68	4.34	4.41
http://www.transmqp.com/	9.1	7.56	8.5
http://www.twilightsex.com/	23.62	15.34	12.6
http://www.uniformslutzone.com/	7.73	3.11	3.54
http://www.vibe.com/	52.9	17.67	18.37
http://www.videoshome.com/	18.63	23.57	20.2
http://www.voyeurzine.com/	9.62	10.32	9.45
http://www.worldsex.com/	23.73	13.32	11.14
http://www.x-3.net/	11.82	11.25	18.12
http://www.xxxvogue.net/	69.32	35.48	33.25
http://www.yusex.com/	4.62	3.84	8.53
เฉลี่ย	22.24	10.74	10.46

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ ก.6 เวลาที่ใช้ ในการเปิดเว็บไซต์ (อนาจาร) เมื่อคัดกรองผิดพลาด

รายชื่อเว็บไซต์ (ตามอนาจาร) ที่คัดกรองผิดพลาด 14 เว็บไซต์	เวลาที่ใช้ เมื่อไม่ ผ่านโปรแกรม (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด High (วินาที)	เวลาที่ใช้ เมื่อผ่าน โปรแกรมโหมด Medlum (วินาที)
http://stockingfever.com/	17.57	10.4	10.31
http://www.between-legs.com/	5.81	32.93	12.5
http://www.kongolist.com/	29.88	52.05	33.28
http://www.libraryofthumbs.com/	6.18	6.4	6.29
http://www.puppykibble.com/	41.2	153.54	41.37
http://www.redhothoneys.com/	33.17	52.24	27.44
http://www.serieshardcore.com/	16.73	53.79	18.1
http://www.stockingpost.com/	4.48	7.67	6.42
http://www.thumbclub.com/	5.55	17.93	6.4
http://www.uniformpost.com/	54.32	105.25	23.74
http://www.vixena.it/	42.85	71.56	45.67
http://www.wolflist.com/	31.73	273.37	49.26
http://www.x-mint.com/	68.19	130.56	77.16
http://www.xxbeauties.com/	26.5	106.56	31.6
เฉลี่ย	27.44	76.73	27.82

ตารางที่ ก.7 จำนวน Tag ต่างๆ ของเว็บไซต์ (อนาจาร) ที่คัดกรองได้อย่างถูกต้อง

รายชื่อเว็บไซต์ (อนาจาร) คัดกรองได้อย่างถูกต้อง 86 เว็บไซต์	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://clothing-mania.com/	175	0	1	1	503	100	403	32	471	0
http://ohgood.com/	3	0	0	6	1206	1	1205	16	1190	0
http://oh-teen.com/	878	2	5	5	917	0	917	44	873	0
http://spunkysheets.com/	164	4	11	1	161	96	65	109	52	0
http://straight.theater.moviemonster.com/	53	6	0	8	211	52	159	8	203	4
http://www.247pantyhose.com/	17	2	1	1	362	16	346	58	304	0
http://www.adult-series.com/	169	9	5	3	180	39	141	151	29	0
http://www.ah-me.com/	244	5	3	5	424	49	365	34	390	0
http://www.apieceoffass.com/	4	2	4	28	49	4	45	46	3	0
http://www.aroundtheworldsex.com/	10	6	9	6	248	7	241	32	216	0
http://www.asstraffic.com/	35	0	3	2	34	14	20	18	16	0
http://www.babesglamour.com/	136	3	4	3	202	107	95	197	5	0
http://www.bestxxxseries.com/	1	3	3	7	1574	0	1574	90	1484	0
http://www.bizarre-rituals.com/	10	4	10	11	0	394	4	390	112	282
http://www.blogchalking.com/life/psm/forced-big-cock-sex/	20	8	6	0	187	0	187	184	3	0
http://www.bootae.com/	127	13	9	5	261	122	139	9	252	0
http://www.candylist.com/	14	6	3	2	350	12	338	334	16	0
http://www.chubbyland.com/	10	3	3	3	273	4	269	194	79	0
http://www.cleancafe.com/	8	4	18	4	728	6	722	30	698	0
http://www.crazynakedchick.com/	251	4	2	1	394	18	376	166	228	0
http://www.crazy-series.com/	96	8	3	1	237	95	142	237	0	0
http://www.croseries.com/	63	4	4	2	151	34	117	46	105	0
http://www.dariusporn.com/	99	11	9	1	637	98	539	17	620	0
http://www.e-ladies.com/	77	3	4	3	85	66	19	81	4	0
http://www.eurothumb.com/	194	22	5	7	1201	177	1024	29	1172	0

ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายชื่อเว็บไซต์(อนาจาร) คัดกรองได้อย่างถูกต้อง (ต่อ)	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://www.fmgalleries.com/	55	3	9	4	209	10	199	88	121	0
http://www.footster.net/	6	5	3	8	10	4	6	8	2	0
http://www.freebigmovies.com/	417	12	14	4	557	32	525	357	200	0
http://www.freepicseries.com/	11	4	3	6	545	5	540	60	485	0
http://www.freepornpictures.org/	81	8	12	2	531	37	494	100	431	0
http://www.frogsex.com/	1	1	0	5	928	0	928	81	847	0
http://www.fuckcat.com/	10	3	3	0	852	0	852	9	843	0
http://www.fullvidz.com/	45	3	5	2	670	6	664	628	42	0
http://www.greenthumbstgp.com/	8	15	23	5	56	4	52	56	0	0
http://www.hankshoneys.com/	15	10	9	5	6	5	1	6	0	0
http://www.hardsex-movies.com/	164	8	3	10	315	159	156	128	187	0
http://www.hornygrannysex.com/	587	4	5	4	660	13	647	86	574	0
http://www.hotfreesex4all.com/	268	8	5	5	396	218	178	66	330	0
http://www.hotfunhouse.com/	10	2	6	2	411	5	406	9	402	0
http://www.huggernet.com/	99	3	3	6	230	60	170	167	63	0
http://www.imagefest.com/	4	2	0	28	48	4	44	46	2	0
http://www.infospiral.com/category/mature/	257	2	3	1	101	0	101	90	11	0
http://www.ixixi.com/	21	11	6	4	2029	21	2008	8	2021	0
http://www.jennysbookmarks.com/	5	3	5	2	740	0	740	13	727	0
http://www.justseries.com/	1	9	9	0	2	0	2	1	1	0
http://www.legs-fetish.com/	5	3	4	4	26	2	24	26	0	0
http://www.legs-fetish.com/thumbnailpost/	23	4	5	1	753	15	738	620	133	0
http://www.leg-zone.com/	7	15	13	4	4	1	3	2	2	0
http://www.mature.dk/	4	2	7	15	18	1	17	1	17	0
http://www.meoland.com/	30	8	12	7	1530	27	1503	1389	141	0
http://www.nslst.com/	22	3	7	8	660	19	641	642	18	0
http://www.old69.com/	0	3	4	4	1212	0	1212	1123	89	0
http://www.onefullmovie.com/	21	2	2	5	64	9	55	57	7	0
http://www.pandalist.com/	27	2	2	2	89	24	65	89	0	0
http://www.pinkvisualtgp.com/	65	3	10	2	118	55	63	62	56	0
http://www.pornfoundation.com/	16	13	24	3	941	0	941	1	940	0
http://www.pornpixxxs.com/	59	4	5	10	669	57	612	25	644	0
http://www.premiumxxxcore.com/	9	8	14	9	393	6	387	37	356	0
http://www.pued.com/	22	5	13	3	971	2	969	123	848	0
http://www.pushin pussy.com/	5	7	19	3	346	5	341	222	124	0
http://www.pussydivine.com/	4	6	8	7	7	2	5	7	0	0
http://www.realxworld.com/	1	3	4	1	766	0	766	43	723	0
http://www.series-xxx.com/	13	3	7	2	6	0	6	5	1	0
http://www.sicklist.com/	8	10	3	3	1861	4	1857	5	1856	0
http://www.smutcastle.com/	0	3	5	2	2	0	2	1	1	0
http://www.spankstation.com/	1	12	12	2	941	0	941	7	934	0
http://www.stocking-mania.com/	8	2	4	1	5	4	1	4	1	0
http://www.stocking-movies.com/	60	3	6	4	241	57	184	240	1	0
http://www.stockings-world.com/	68	13	4	7	139	47	92	127	12	0
http://www.stocking-tease.com/	234	4	5	7	526	126	400	524	2	0
http://www.teeniesxxx.com/	232	3	6	4	245	145	100	27	218	0
http://www.tqplibrary.com/	4	1	5	0	7	0	7	6	1	0
http://www.thumbzilla.com/	2	3	3	1	3	0	3	0	3	0
http://www.tiava.com/	2	3	4	1	1298	2	1296	17	1281	0
http://www.tommys-bookmarks.com/	4	5	7	2	58	2	56	30	28	0

ไม่วารณมีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รายชื่อเว็บไซต์(อนาจาร) คัดกรองได้อย่างถูกต้อง (ต่อ)	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://www.transmqp.com/	215	8	5	7	292	166	126	82	210	0
http://www.twilightsex.com/	296	3	6	5	475	62	413	284	191	0
http://www.unifor:mslutzone.com/	0	7	2	1	49	0	49	49	0	0
http://www.vibe.com/	35	3	1	14	188	23	165	43	145	0
http://www.videoshome.com/	128	4	3	3	884	8	876	353	531	0
http://www.voyeurzine.com/	3	10	3	3	602	1	601	599	3	0
http://www.worldsex.com/	67	4	8	15	605	64	541	488	117	0
http://www.x-3.net/	6	4	4	2	734	4	730	15	719	0
http://www.xxxvogue.net/	256	2	1	5	1668	102	1566	14	1654	0
http://www.yusex.com/	8	2	2	2	29	7	22	28	1	0
เฉลี่ย	81.09	5.30	6	4.70	462.30	36.97	429.89	140.89	327.31	3.36

ตารางที่ ก.8 จำนวน Tag ต่างๆ ของเว็บไซต์ (อนาจาร) ที่คัดกรองผิดพลาด

รายชื่อเว็บไซต์(อนาจาร) คัดกรองผิดพลาด 14 เว็บ	IMG	META	KEY- WORD	SCRIPT	ALL- LINK	IMG- LINK	TXT- LINK	LINK- NON	LINK- RELA	PARAM
http://stockingfever.com/	5	0	0	3	242	1	241	88	154	0
http://www.between-legs.com/	15	3	0	4	158	5	153	156	2	0
http://www.kongolist.com/	9	1	0	33	34	2	32	5	29	0
http://www.libraryofthumbs.com/	1	3	0	0	2	0	2	2	0	0
http://www.puppykibble.com/	170	1	0	2	192	141	51	141	51	0
http://www.redhothoneys.com/	47	1	0	3	73	27	46	6	67	0
http://www.serieshardcore.com/	87	0	0	4	86	57	29	84	2	0
http://www.stockingpost.com/	4	0	0	1	2	1	1	1	1	0
http://www.thumbclub.com/	1	0	0	1	359	1	358	0	359	0
http://www.uniformpost.com/	97	0	0	6	133	92	41	14	119	0
http://www.vixena.it/	52	6	0	3	135	37	98	95	40	0
http://www.wolflist.com/	2773	1	0	1	2819	0	2819	1	2818	0
http://www.x-mint.com/	121	7	0	4	327	120	207	0	327	0
http://www.xxbeauties.com/	119	1	0	1	345	111	234	136	209	0
เฉลี่ย	250.07	1.71	0	4.71	350.5	42.5	308	52.07	298.42	0

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ข. คู่มือการติดตั้งซอฟต์แวร์

คู่มือการติดตั้งซอฟต์แวร์

1. ทำการติดตั้งโปรแกรม Web browser Fire Fox เวอร์ชัน 1.5.06 ขึ้นไปเพื่อให้สามารถใช้งานโปรแกรมนี้ผ่าน Fire Fox ได้
2. ทำการติดตั้งโปรแกรม Web browser Internet Explorer เวอร์ชัน 6 ขึ้นไป ขึ้นไปเพื่อให้สามารถใช้งานโปรแกรมนี้ผ่าน Internet Explorer ได้
3. ทำการดับเบิลคลิกตัวลงโปรแกรม setup.exe หรือ WebContentFiltering.msi ดังรูปที่ 1 เพื่อทำการติดตั้งโปรแกรมนี้



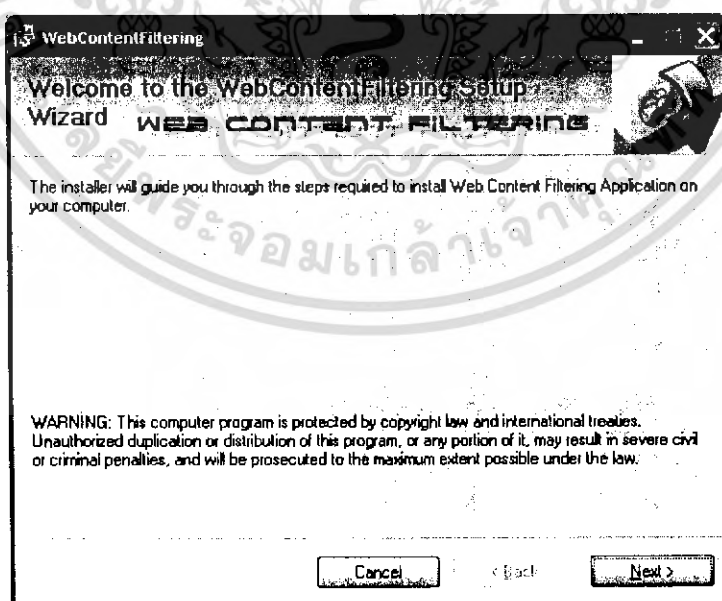
setup.exe



WebContentFiltering.msi

รูปที่ 1 ไอคอนที่จะทำการติดตั้งโปรแกรม

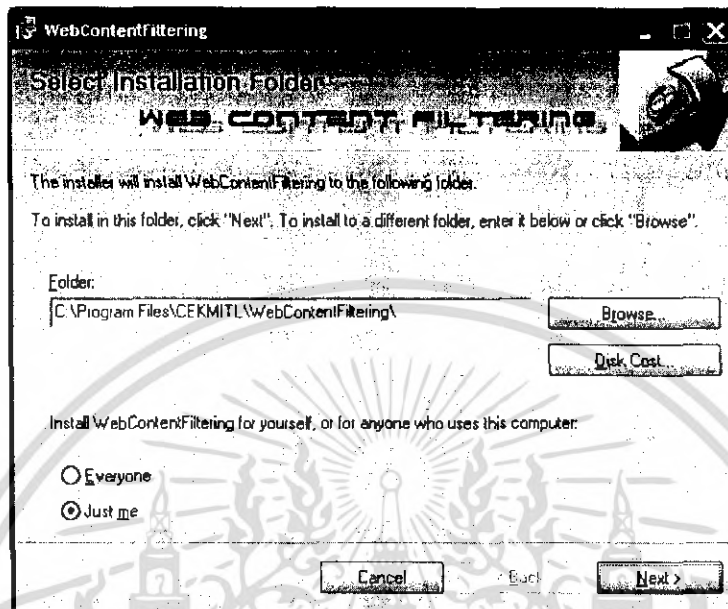
4. หลังจากดับเบิลคลิกตัวลงโปรแกรม แล้วจะปรากฏหน้าจอ ดังรูปที่ 2



รูปที่ 2 หน้าจอเริ่มต้นเมื่อทำการติดตั้งโปรแกรม

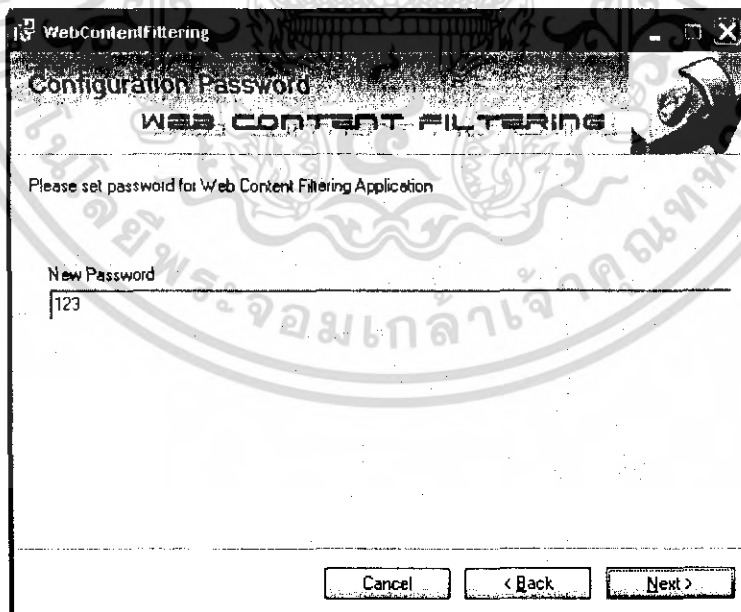
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5. จากนั้นจะเป็นการเลือก Path ของโปรแกรมที่ผู้ใช้งานจะทำการติดตั้ง และจะติดตั้งลงใน User ประเภทใดบ้าง ดังรูปที่ 4



รูปที่ 4 หน้าจอเลือกPathที่จะทำการติดตั้งโปรแกรม

6. จากนั้นจะโปรแกรมจะถามการตั้งรหัส Password เพื่อป้องกันการปรับแต่ง โปรแกรมโดยไม่ได้ รับผิดชอบ โดยที่ค่าเริ่มต้นจะเป็นรหัส 123



รูปที่ 5 หน้าจอตั้ง Password โดยค่าที่ตั้งมาให้คือ123

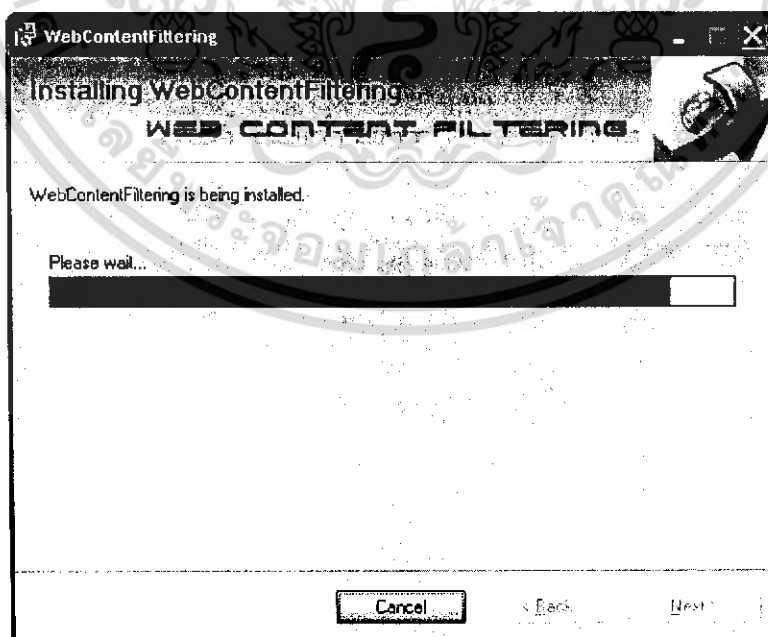
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

7. ขั้นตอนต่อไปจะเป็นการยืนยันเพื่อจะทำการติดตั้งโปรแกรม หลังจากนั้นให้กดปุ่ม “Next” ซึ่งแสดงดังรูปที่ 6



รูปที่ 6 การยืนยันเพื่อจะติดตั้งโปรแกรม

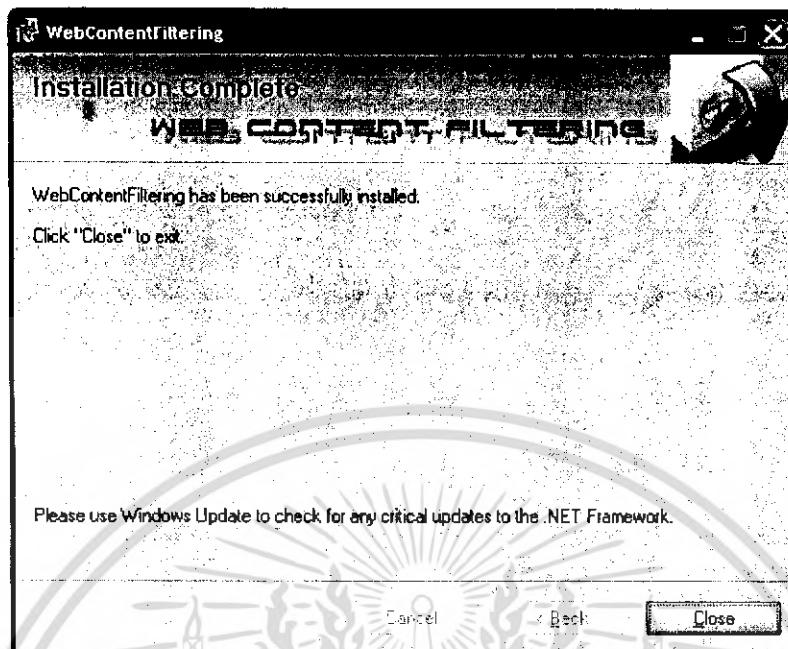
8. โปรแกรมจะทำการติดตั้งจนเสร็จสมบูรณ์



รูปที่ 7 สถานะที่อยู่ระหว่างการติดตั้งโปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้คัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

9. เมื่อได้ทำการติดตั้งเสร็จเรียบร้อยแล้วคลิกปุ่ม “Close” เพื่อสิ้นสุดการติดตั้งโปรแกรม



รูปที่ 8 การสิ้นสุดการติดตั้ง โปรแกรม

10. เมื่อทำการติดตั้งโปรแกรมเสร็จเรียบร้อยแล้วจะมีไอคอนขึ้นที่หน้า Desktop และ All Program



รูปที่ 9 ไอคอนของ โปรแกรม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ภาคผนวก ค.

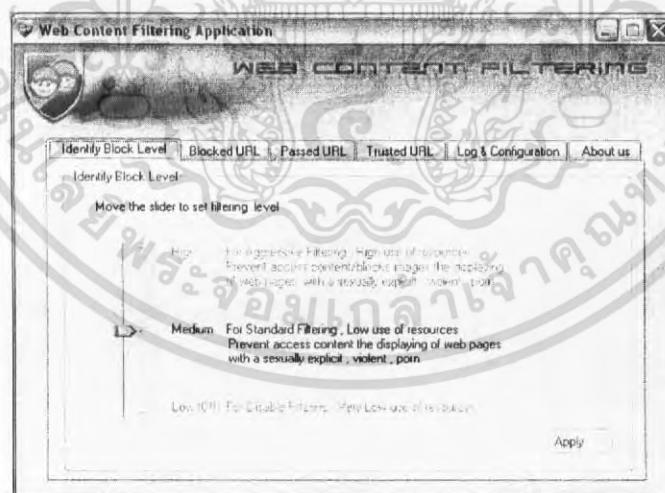
คู่มือการใช้งานซอฟต์แวร์

คู่มือการใช้งานซอฟต์แวร์

1. เมื่อติดตั้งเสร็จเรียบร้อยแล้วให้ดับเบิลคลิกที่ไอคอนที่อยู่บนหน้าจอ Desktop หรือ All Program จะปรากฏหน้าต่างที่จะทำการควบคุม , แสดงและปรับแต่งค่าต่างๆ ได้ โดยที่จะมีแท็บให้เลือกดังนี้

1.1 Identify Block Level จะเป็นการปรับแต่งระดับการคัดกรองข้อมูลโดยที่

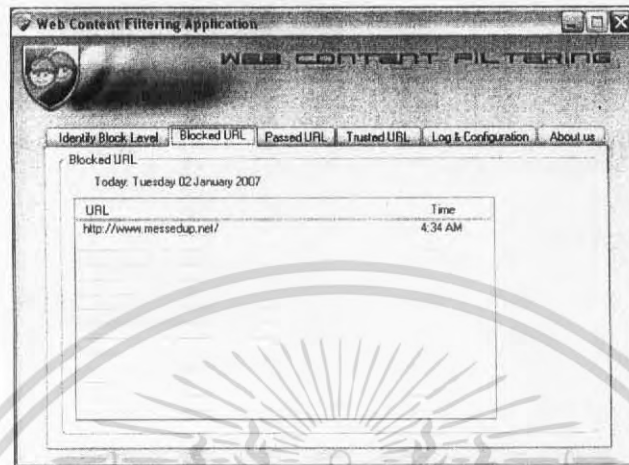
- ระดับ High จะเป็นการคัดกรองที่ละเอียดที่สุด โดยที่จะคัดกรองทั้งรูปและเนื้อหา (Content) ที่อยู่ในโค้ด HTML ที่ได้รับมา ถ้าเนื้อหาไม่เหมาะสมจะทำการบล็อกทั้งหน้าเว็บเพจนั้นๆ และถ้าหากพบรูปลามกอนาจารจะไม่แสดงรูปนั้นๆ บนหน้าเว็บเพจ ซึ่งในระดับนี้จะใช้ทรัพยากรบนเครื่องเปลืองมากที่สุด
- ระดับ Medium เป็นการคัดกรองระดับมาตรฐานซึ่งจะคัดกรองเฉพาะเนื้อหา (Content) ที่อยู่ในโค้ด HTML ที่ได้รับมา ถ้าเนื้อหาไม่เหมาะสมจะทำการบล็อกทั้งหน้าเว็บเพจนั้นๆ ซึ่งในระดับนี้จะใช้ทรัพยากรของเครื่องน้อย และเป็นระดับที่แนะนำในการใช้งาน
- ระดับ Low(off) เป็นระดับที่ไม่มีการคัดกรองข้อมูลบนหน้าเว็บ



รูปที่ 1 หน้าจอการปรับระดับการคัดกรองข้อมูลเว็บไซต์

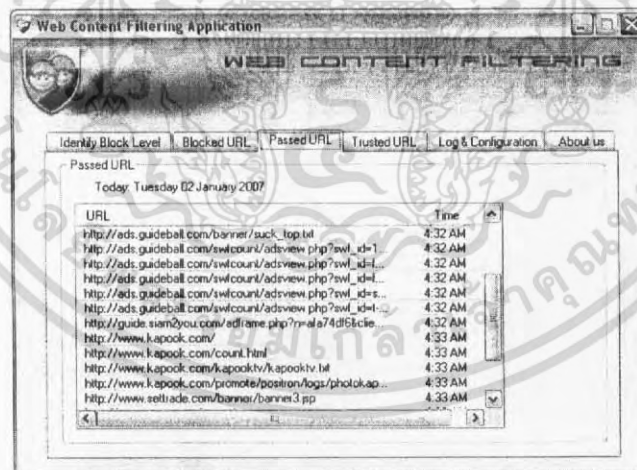
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

1.2 Blocked URL จะแสดงรายชื่อเว็บไซต์ URL ที่ถูกลบออกจากโปรแกรม พร้อมทั้งแสดงเวลาที่ใช้



รูปที่ 2 หน้าจอเก็บรายชื่อเว็บไซต์ URL ที่ถูกลบออก

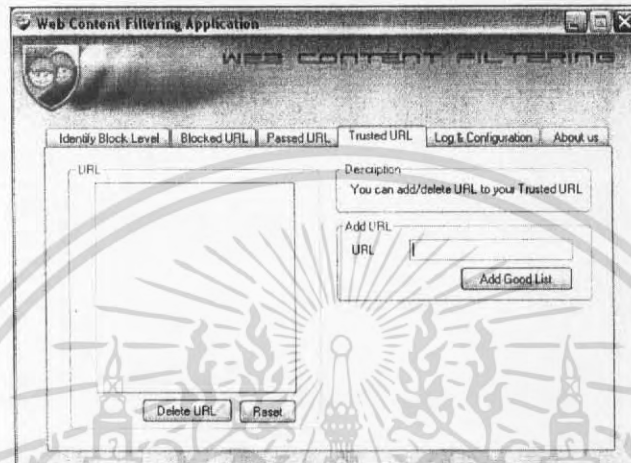
1.3 Passed URL จะแสดงรายชื่อเว็บไซต์ URL ที่เป็นเว็บไซต์ปกติ ไม่ลามกอนาจาร (ไม่ถูกลบออก) พร้อมทั้งแสดงเวลาที่ใช้



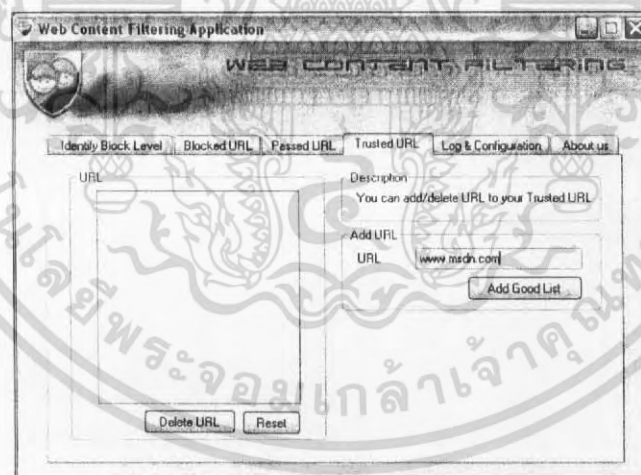
รูปที่ 3 หน้าจอเก็บรายชื่อเว็บไซต์ URL ที่ผ่านการคัดกรอง (ไม่ถูกลบออก)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- 1.4 Trusted URL เป็นหน้าต่างแสดงรายชื่อ เว็บไซต์ URL ที่นำเชื่อถือ โดยโปรแกรมจะไม่ทำการคัดกรองข้อมูลเว็บในเว็บไซต์เหล่านี้เลย ซึ่งผู้ใช้สามารถทำการเพิ่มเติมและลบชื่อเว็บไซต์ URL ที่นำเชื่อถือเหล่านี้ได้

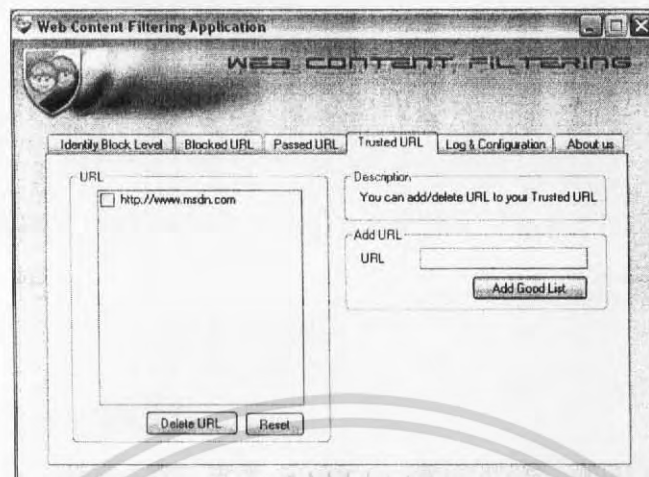


รูปที่ 4 หน้าจอ Trusted URL

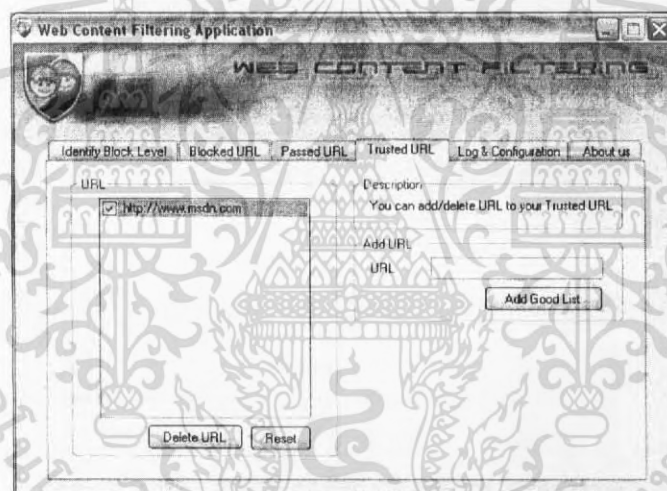


รูปที่ 5 การเพิ่มรายชื่อเว็บไซต์ URL เพื่อไม่นำ URL นั้นๆ มาคัดกรอง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

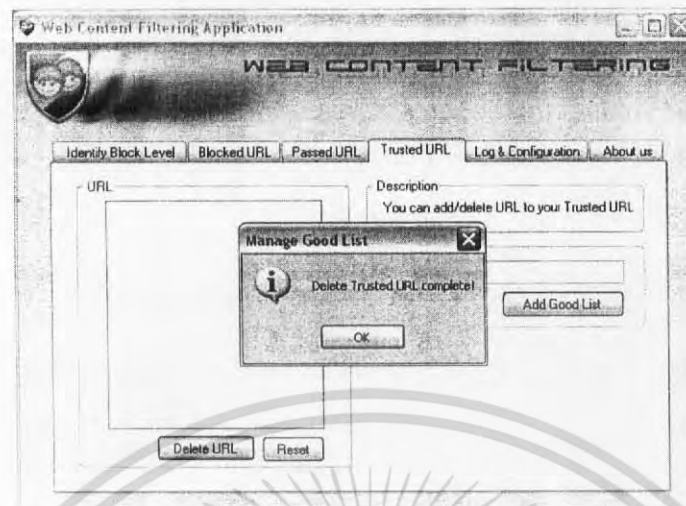


รูปที่ 6 รายชื่อเว็บไซต์ URL ที่ได้ถูกเพิ่มไปเรียบร้อยแล้ว



รูปที่ 7 การเลือกรายชื่อเว็บไซต์ URL เพื่อที่จะทำการ ลบ URL นั้น

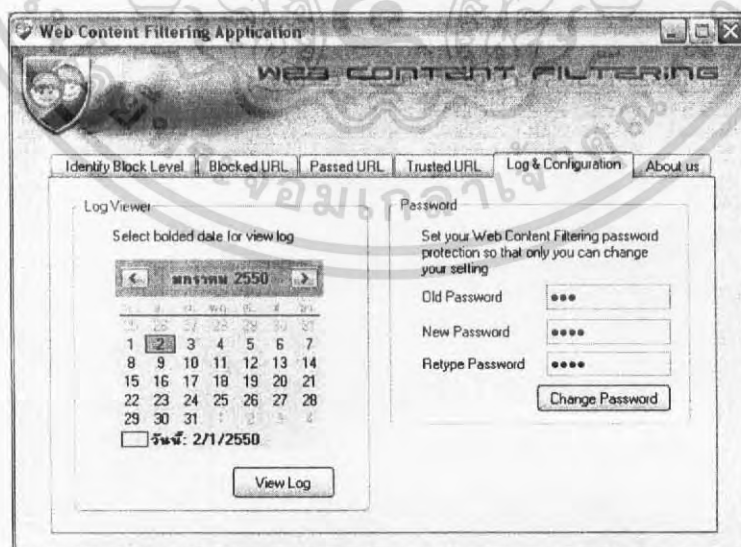
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 8 หน้าจอหลังจากลบรายชื่อเว็บไซต์ URL นั้นแล้ว

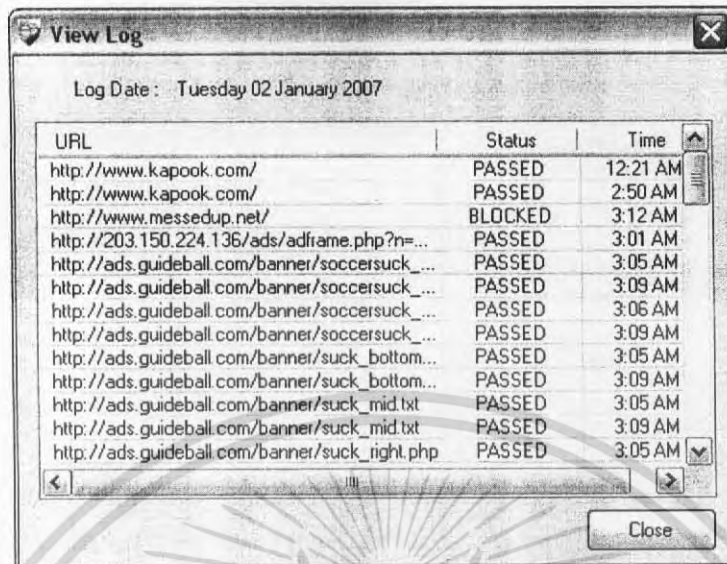
1.5 Log & Configuration

- ทางด้าน Log Viewer (บันทึกการใช้งาน)จะเป็นการดู Log เพื่อดูว่าได้เข้าเว็บไซต์ไหนไปบ้าง โดยที่โปรแกรมจะเก็บ Log ทุกๆ 23.59 น. ของทุกวัน และทุกครั้งที่เปิด โปรแกรม ซึ่งวันที่เก็บ Log ได้จะแสดงเป็นเลขหน้าส่วนวันที่ไม่ได้เก็บ Log ไว้จะเป็นตัวบางธรรมดา
- ทางด้าน Password เป็นการปรับเปลี่ยน Password โดยที่ต้องใส่ค่า Password เก่าก่อนและถึงจะทำการใส่ Password ได้



รูปที่ 9 หน้าจอ Log & Configuration เพื่อทำการดู Log และเปลี่ยน Password

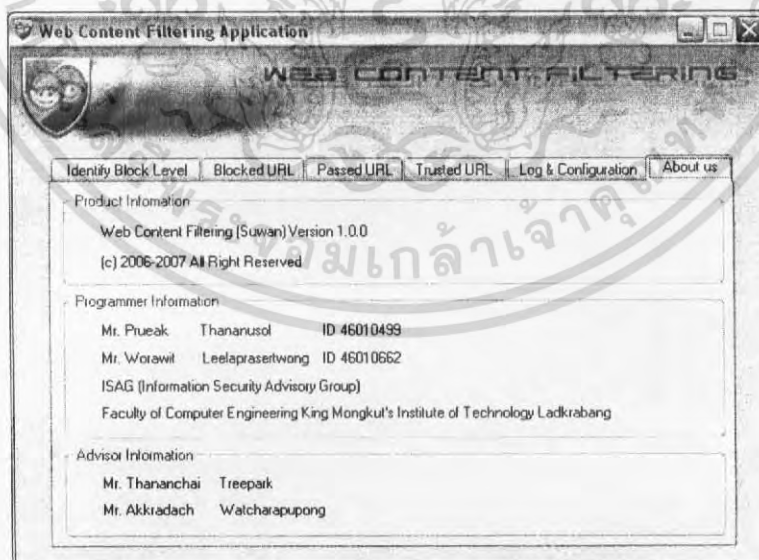
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 10 หน้าจอเมื่อกด “View Log”

1.6 About us

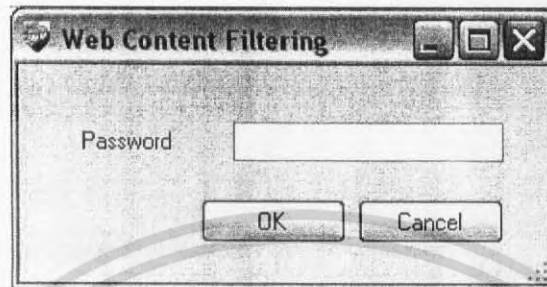
- Product Information เป็นการแสดงข้อมูลของโปรแกรมนั่นคือ Web Content Filtering มีCode name ว่า “Suwan” (เป็นผู้เก็บบัญชีของหน้าเว็บคือเป็นบัญชีความดีซึ่งหมายถึงเป็นการก่ลดร่องสิ่งที่ดีเข้ามาเปรียบเหมือนเว็บไซต์ไม่ลามกอนาจาร) และ Version 1.0.0
- Programmer Information เป็นการแสดงรายชื่อผู้พัฒนา โปรแกรม
- Advisor Information เป็นการแสดงอาจารย์ที่ปรึกษาเกี่ยวกับ โปรแกรมนี้



รูปที่ 11 หน้าจอ About us

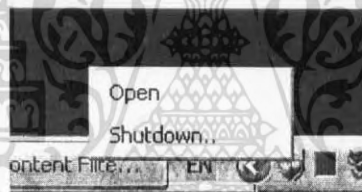
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หมายเหตุ ทุกครั้งที่มีการปรับแต่งโปรแกรมผู้ใช้ต้องมีการกรอก Password ก่อนที่จะทำการปรับแต่งค่า และปิดโปรแกรม



รูปที่ 12 หน้าจอกรอก Password

2. ผู้ใช้งานทำการปิดโปรแกรมได้โดยคลิกขวาที่ไอคอนมุมล่างซ้ายและเลือก Shutdown..



รูปที่ 13 หน้าจอการคลิกขวาที่ไอคอนมุมล่างขวา

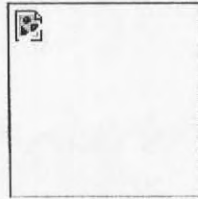
3. หน้า Web Browser ที่ถูกบล็อกจะแสดงหน้าจอดังนี้



รูปที่ 14 หน้าเว็บเพจที่ถูก Block Content (Filter Level Medium หรือ High)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4. หน้า Web Browser ที่ถูกรองรูปจะไม่แสดงรูปดังหน้าจอนี้



รูปที่ 15 หน้าเว็บที่ถูกรองรูป แต่ Content ไม่โดน Block (Filter Level High)



รูปที่ 16 หน้าเว็บปกติ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้