



ผลการวิเคราะห์ความซ้ำซ้อนของข้อมูล บนหน้าเว็บไซต์ในระบบเครือข่ายอินเทอร์เน็ต*

วสุธาน ตันบุญเฮง**

บทคัดย่อ

การวิจัยนี้มีวัตถุประสงค์เพื่อศึกษาปริมาณความซ้ำซ้อนของข้อมูลหน้าเว็บไซต์ที่ถูกส่งผ่านระบบเครือข่ายอินเทอร์เน็ต โดยกลุ่มตัวอย่างที่ใช้ในการวิจัยครั้งนี้คัดเลือกจากเว็บไซต์ที่มีการให้บริการแตกต่างกันจำนวนสาม ประเภทคือ (ก) เว็บไซต์ที่ให้บริการข่าวสารข้อมูล (ข) Web Board ซึ่งเป็นเว็บไซต์ที่เปิดโอกาสให้ผู้ใช้สามารถแสดงความคิดเห็นได้ และ (ค) Web Search Engine ซึ่งเป็นเว็บไซต์ที่ให้บริการค้นหาข้อมูล ข้อมูลส่วนที่เป็น Script ซึ่งเขียนด้วยภาษา HTML (Hyper Text Markup Language) จะถูกนำมาเปรียบเทียบกันระหว่างแต่ละหน้าของเว็บไซต์เพื่อวิเคราะห์หาอัตราส่วนของข้อมูลที่มีการส่งซ้ำซ้อนต่อปริมาณข้อมูลที่ถูกส่งทั้งหมด ผลการวิจัยพบว่า เว็บไซต์ประเภทที่ให้บริการข่าวสารข้อมูล (ก) มีอัตราส่วนข้อมูลที่มีการส่งซ้ำซ้อนมากที่สุดคือสูงถึง 66% หรือกล่าวอีกนัยหนึ่งคือทุกครั้งที่ใช้คลิกเพื่อเปลี่ยนหน้าเว็บไซต์เพื่ออ่านข่าว จะมีข้อมูลมากถึง 66% ที่จะต้องถูกส่งซ้ำ ส่วนเว็บไซต์ประเภทที่เปิดโอกาสให้ผู้ใช้สามารถแสดงความคิดเห็นได้ (ข) มีอัตราส่วนข้อมูลที่มีการส่งซ้ำซ้อนน้อยที่สุดคือ 23% สำหรับเว็บไซต์ประเภทที่ให้บริการค้นหาข้อมูลนั้น (ค) มีอัตราส่วนข้อมูลที่มีการส่งซ้ำซ้อนอยู่ที่ 38% ซึ่งใกล้เคียงกับเว็บไซต์ประเภท (ข)

คำสำคัญ : ข้อมูลซ้ำซ้อน เว็บไซต์ อินเทอร์เน็ต

* ได้รับทุนวิจัยจากสถาบันการจัดการปัญญาภิวัฒน์

** อาจารย์ประจำ คณะวิศวกรรมศาสตร์และเทคโนโลยี สาขาวิศวกรรมคอมพิวเตอร์ สถาบันการจัดการปัญญาภิวัฒน์



Analysis Redundant Data Over Internet Website*

Vasutan Tunbunheng**

Abstract

The present was aimed to analysis redundant data over internet which is using to display webpage. The sample websites are classified into three categories: (a) website for broadcasting news or information data over internet, (b) web board website which allows users to post their comments or share information to each others, and (c) web search engine website which provides searching service. The script data, which is HTML (Hyper Text Markup Language), is analyzed and compared to the script of another page of the same website to find ratio of redundant data. The result shows that the news website (a) has highest ratio of redundant data. In the other words, when a user click on the website to seek to another page, 66% of the script data must be send again. The web board website (b) has lowest ratio of redundant data, that is 23%. For the web search engine website (c), the ratio of redundant data is 38%, which is almost the same as the web board website (b).

Key words : Redundant Data; Website; Internet

* This research was funded from Panyapiwat Institute of Management (PIM)

** Faculty member, Computer Engineering, Faculty of Engineering and Technology, Panyapiwat Institute of Management (PIM)



บทนำ

ระบบเครือข่ายอินเทอร์เน็ตเป็นช่องทางหนึ่งที่ถูกใช้เพื่อติดต่อสื่อสาร และ แลกเปลี่ยนข้อมูลระหว่างสถานที่ซึ่งอยู่ห่างไกลกันได้อย่างรวดเร็ว ผู้ใช้สามารถเรียกข้อมูลจากที่อื่นมาดูได้ด้วยการใช้โปรแกรม Web Browser เพื่อเชื่อมต่อไปยังเว็บไซต์ต่างๆ โดยการระบุ URL (Uniform Resource Locator) ของเว็บไซต์ที่ต้องการ จากนั้นเว็บไซต์ต้นทางจะจัดเตรียมข้อมูลที่จะส่ง โดยแบ่งออกเป็นข้อมูลย่อยจำนวนมากในรูปแบบที่เรียกว่า Package ซึ่งมีการระบุปลายทางว่าจะถูกส่งไปที่ไหนและลำดับของข้อมูล เมื่อข้อมูลถูกส่งเข้าไประบบเครือข่ายอินเทอร์เน็ต ข้อมูลแต่ละ Package อาจใช้เส้นทางในการเดินทางต่างกันขึ้นอยู่กับว่าเส้นทางไหนมีความหนาแน่นในการส่งข้อมูลน้อยที่สุดในขณะนั้น ข้อมูลที่เดินทางถึงปลายทางอาจไม่เรียงตามลำดับจะต้องมีการเรียงลำดับข้อมูลให้ถูกต้องก่อน แล้วจึงแยกเอาเฉพาะข้อมูลที่ต้องการออกมาจากข้อมูล Package นำมารวมกันเป็นข้อมูลเดิมที่สมบูรณ์ เพื่อนำมาใช้แสดงผลที่ปลายทาง

ข้อมูลที่ถูกส่งประกอบด้วยสองส่วนคือ ส่วนที่เป็นเนื้อหา เช่น ตัวอักษร ข้อความ Link เพื่อใช้อ้างอิงเชื่อมโยงถึงรูปภาพ เสียงเพลง วิดิทัศน์ เอกสารต่างๆ และ ข้อมูลส่วนที่ใช้ควบคุมให้โปรแกรม Web Browser แสดงการจัดข้อมูลเนื้อหาในรูปแบบที่ต้องการเพื่ออำนวยความสะดวก หรือเพื่อความสวยงาม ซึ่งข้อมูลส่วนนี้มีลักษณะเป็นภาษา Script อยู่ปะปนกับข้อมูลเนื้อหาแต่มีการใช้สัญลักษณ์พิเศษกำกับเพื่อให้แยกออกจากกันได้

ชัดเจน เช่น การขึ้นบรรทัดใหม่ในภาษา HTML (Hyper Text Markup Language) ทำได้โดยการแทรกคำสั่ง "
" ที่ท้ายข้อความ เป็นต้น

เมื่อเทคโนโลยีมีการพัฒนาเติบโตขึ้น ปริมาณความต้องการใช้ช่องทางข้อมูลนี้เพื่อบริโภคข่าวสารและความหลากหลายของการใช้งานก็เพิ่มขึ้นตามไปด้วย จึงเกิดมีงานวิจัยจำนวนมากเพื่อพัฒนาให้ระบบเครือข่ายอินเทอร์เน็ตสามารถรับข้อมูลได้ในปริมาณมากขึ้นและส่งข้อมูลโดยใช้เวลาน้อยลง เช่น การปรับปรุงประสิทธิภาพอุปกรณ์ Hardware ทำให้สามารถส่งข้อมูลได้เร็วขึ้น การบริหารความแออัดโดยเลือกใช้เส้นทางที่ข้อมูลจะไปถึงปลายทางได้เร็วที่สุด การลดขนาดข้อมูลที่ส่งให้น้อยลงด้วยวิธีการบีบอัดข้อมูลรูปภาพ เสียงเพลง และ วิดิทัศน์ ทำให้สามารถรองรับการใช้งานมากขึ้นได้

การลดข้อมูลที่ส่งในระบบเครือข่ายอินเทอร์เน็ตอีกประการที่สามารถทำได้คือ การลดขนาดข้อมูลโดยการส่งเฉพาะข้อมูลที่จำเป็นเท่านั้น ซึ่งเว็บไซต์จำนวนมากถูกสร้างขึ้นมาโดยไม่ได้คำนึงถึงจุดนี้ เนื่องจากถูกพัฒนาขึ้นมาในลักษณะ Dynamic Website [1] ข้อมูลเนื้อหาที่ถูกเก็บรวบรวมไว้ในฐานข้อมูลจะถูกนำมาแสดงโดยมีรูปแบบการจัดหน้าที่ตายตัว เมื่อผู้ใช้เปลี่ยนไปดูหน้าอื่นจะมีเฉพาะข้อมูลส่วนเนื้อหาเท่านั้นที่เปลี่ยนไปในขณะที่ข้อมูลคำสั่ง Script สำหรับใช้ในการจัดหน้าจะถูกส่งซ้ำ ส่งผลให้ข้อมูลบนระบบเครือข่ายอินเทอร์เน็ตเพิ่มขึ้นโดยไม่จำเป็น

ปัจจุบันมีเทคโนโลยีที่ช่วยเพิ่มลูกเล่นให้โปรแกรม Web Browser แสดงข้อมูลได้น่าสนใจ



มากขึ้น เช่น การแทรกคำสั่งเรียกโปรแกรมที่เขียนโดยใช้ภาษา Java [2] หรือ ActiveX [3] ลงไปเพื่อแสดงการเคลื่อนไหวและเสียงตอบสนองผู้ใช้สามารถนำมาดัดแปลงเพื่อใช้ช่วยจัดหน้าสำหรับแสดงข้อมูลเนื้อหาในบริเวณที่กำหนดได้ การใช้โปรแกรมภาษา Java จะทำงานอยู่ภายใต้โปรแกรม Java Virtual Machine แยกต่างหากอีกชั้นเพื่อความปลอดภัยของเครื่องผู้รับ แต่มีข้อเสียคือความเร็วของการประมวลผลช้ากว่าการประมวลผลที่เครื่องสามารถทำได้จริง ในขณะที่การใช้โปรแกรมประเภท ActiveX ควบคุมการทำงานของโปรแกรม Web Browser โปรแกรมจะถูกนำมาประมวลผลโดยตรงที่เครื่องปลายทางซึ่งเร็วกว่าแต่อาจมีผลเสียต่อระบบความปลอดภัยของเครื่องผู้ใช้ จึงมีการพัฒนาโปรแกรม Flash [4] ให้รวมเข้ากับโปรแกรม Web Browser ขึ้นมาโดยรวมเอาทั้งข้อดีทั้งด้านความปลอดภัยและความเร็วในการประมวลผล เหมาะสำหรับการสร้างสื่อบันเทิงที่มีความสามารถตอบโต้กับผู้ใช้ได้

ถึงแม้ว่าเว็บไซต์จะมีหลายหน้าให้ผู้ใช้ได้เลือกชม แต่สิ่งที่เหมือนกันคือ ข้อมูลแต่ละหมวดมีรูปแบบการจัดหน้าตามลำดับที่ตายตัว และ มีการเปลี่ยนเพียงข้อมูลส่วนเนื้อหาให้สอดคล้องกับหมวดหมู่ของข้อมูลที่เลือกเท่านั้น หากแสดงหน้าเว็บไซต์โดยดึงเฉพาะข้อมูลที่เป็นส่วนเนื้อหาเท่านั้น จะสามารถช่วยลดปริมาณและความหนาแน่นของข้อมูลบนระบบเครือข่ายอินเทอร์เน็ตทำให้สามารถรองรับการใช้งานได้มากขึ้น หากค่าความหนาแน่นของข้อมูลมากเกินไประบบจะสามารถ

รองรับได้ ข้อมูลส่วนเกินจะต้องหยุดรอจนกว่าระบบจะว่างพอสำหรับการส่งข้อมูลชุดถัดไป ทำให้เกิดความล่าช้าในการแลกเปลี่ยนข่าวสารข้อมูล

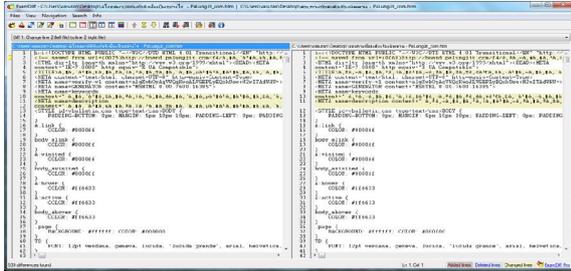
วัตถุประสงค์ของการวิจัย

เพื่อศึกษาอัตราส่วนระหว่างข้อมูลที่มีการส่งซ้ำซ้อนและข้อมูลส่วนที่เป็นเนื้อหาซึ่งสัมพันธ์กับการแสดงผลในแต่ละหน้าของเว็บไซต์ โดยตั้งอยู่บนสมมติฐานว่าแต่ละหน้าของเว็บไซต์เดียวกันจะมีรูปแบบการจัดวางและแสดงเนื้อหาที่อยู่ในหน้าคล้ายกัน ดังนั้นข้อมูลส่วนที่ใช้สำหรับควบคุมการจัดรูปแบบเพื่อแสดงเนื้อหาจะต้องถูกส่งซ้ำทุกครั้งที่ผู้ใช้เปลี่ยนไปดูหน้าอื่นของเว็บไซต์เดิม

วิธีดำเนินการวิจัย

ข้อมูลที่ใช้สำหรับแสดงเว็บไซต์แต่ละหน้าสามารถถูกจัดเก็บเพื่อนำมาวิเคราะห์ได้โดยการเลือกใช้คำสั่งบันทึก (File => Save As...) ของโปรแกรม Internet Explorer (version 8) เพื่อเก็บข้อมูลลงในเครื่องคอมพิวเตอร์ในลักษณะ “Webpage, Complete (*.htm;*.html)” ข้อมูลที่ได้มาประกอบด้วยเอกสารข้อมูลนามสกุล HTML จำนวน 1 file ซึ่งบรรจุข้อมูลทั้งส่วนการจัดหน้าและส่วนเนื้อหารวมกัน และเพิ่มข้อมูลชื่อเดียวกับเอกสารดังกล่าวอีก 1 directory สำหรับรวบรวมข้อมูลรูปภาพ jpg gif กับเอกสารนามสกุล js css htm เพื่อใช้สร้างส่วนประกอบต่างๆของหน้าเว็บไซต์หลังจากบันทึกข้อมูลเอกสารเป็นนามสกุล HTML ของเว็บไซต์หน้าต่าง ๆ ออกมาแล้ว จะนำข้อมูลเฉพาะที่อยู่ภายในเอกสารมาเปรียบเทียบกัน ระหว่างหน้าที่มีลักษณะ

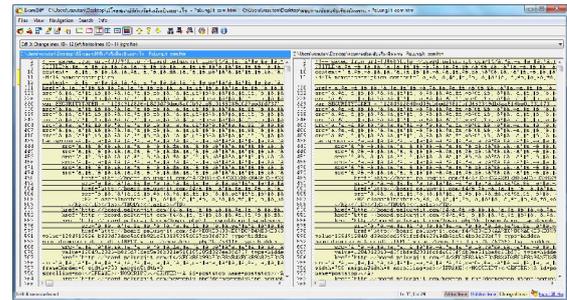
การจัดใกล้เคียงกันซึ่งแตกต่างเฉพาะส่วนเนื้อหาโดยการเปรียบเทียบจะใช้โปรแกรม Exam Diff [5] เพื่อหาบรรทัดที่แตกต่างกัน



ภาพที่ 1 แสดงภาพหน้าจอของโปรแกรม Exam Diff

หน้าจอของโปรแกรม Exam Diff สามารถแสดงได้ดังภาพที่ 1 ซึ่งใช้ในการเปรียบเทียบเอกสารข้อมูล โดยการเปรียบเทียบเอกสารภาษาไทยจะต้องเลือกใช้ Option "Treat both files as text files" เอกสารที่นำมาเปรียบเทียบจะถูกแสดงในพื้นที่ด้านซ้ายและขวา โดยมีการระบายสีบรรทัดที่มีข้อความแตกต่างกัน เมื่อกดปุ่ม "Show Only Differences (Ctrl - D)" บน Menu Bar (ปุ่มที่ 12 เริ่มนับจากทางด้านซ้ายมือ) จะมีผลทำให้โปรแกรมแสดงเฉพาะบรรทัดที่ข้อมูลแตกต่างกันดังรูปที่ 2

เมื่อได้เฉพาะบรรทัดข้อมูลที่แตกต่างระหว่างเอกสารซึ่งเป็นข้อมูลในส่วนของเนื้อหา จะนำเอามันก็กลายเป็นเอกสารใหม่ทั้งข้อมูลด้านซ้ายและขวา แล้วจึงเปรียบเทียบขนาดของ File เอกสารใหม่กับเอกสารเก่าว่ามีขนาดต่างกันเท่าไร ซึ่งเป็นปริมาณข้อมูลส่วนที่ใช้ในการจัดหน้า เช่น การเปรียบเทียบขนาดของข้อมูลเอกสารที่บันทึกมาจาก URL "http://board.palungjit.com/f4/นิโรธสมบัติที่แท้จริงนั้นเป็นอย่างไร -258116.html" และ "http://



ภาพที่ 2 แสดงเฉพาะบรรทัดข้อมูลที่แตกต่างของเอกสารที่นำมาเปรียบเทียบ

board.palungjit.com/f4/สอบถามข้อสงสัยเรื่องนิพพาน-257957.html" ซึ่งเป็นกระตุ้จาก Web Board แห่งเดียวกันและมีขนาดข้อมูลคือ 468 KB และ 293 KB ตามลำดับ เมื่อนำมาแยกเอาเฉพาะข้อมูลส่วนเนื้อหาที่แตกต่างแล้วจะเหลือขนาดเพียง 355 KB และ 179 KB ตามลำดับ ซึ่งมีขนาดลดลงจากเดิม 133 KB และ 114 KB ตามลำดับ สาเหตุที่ปริมาณข้อมูลส่วนจัดหน้าที่คำนวณได้แตกต่างกันเนื่องมาจากจำนวนความคิดเห็นที่แสดงในกระตุ้ต่างกัน ข้อมูลที่ใช้จัดหน้าในส่วนความคิดเห็นที่ไม่เท่ากันนั้นรวมอยู่ในส่วนต่างด้วย โดยในกระตุ้แรกมีอยู่ทั้งหมด 20 ความคิดเห็น ส่วนกระตุ้ที่สองมีอยู่ทั้งหมด 11 ความคิดเห็น ดังนั้นหากส่งตัดข้อมูลส่วนที่ใช้สำหรับจัดหน้าซึ่งซ้ำซ้อนออกไปปริมาณข้อมูลที่ส่งผ่านระบบเครือข่ายอินเทอร์เน็ตจะลดลงประมาณ 123 KB ต่อการแสดงผลหน้ากระตุ้แต่ละครั้ง เป็นต้น

เว็บไซต์ที่นำมาวิเคราะห์ถูกคัดเลือกจากเว็บไซต์ที่มีการให้บริการแตกต่างกันจำนวนสามประเภทคือ (ก) เว็บไซต์ที่ให้บริการข่าวสารข้อมูล (ข) Web Board ซึ่งเป็นเว็บไซต์ที่เปิดโอกาสให้ผู้ใช้สามารถแสดงความคิดเห็นได้ และ (ค) Web Search



Engine ซึ่งเป็นเว็บไซต์ที่ให้บริการค้นหาข้อมูล โดย URL ของเว็บไซต์ที่ถูกคัดเลือกมาใช้ในการวิเคราะห์จะถูกแสดงในตารางที่ 1 และมีการระบุอักษรย่อไว้ในช่องด้านซ้ายมือสุดเพื่อใช้อ้างอิงข้อมูลที่แสดงในแผนภูมิ รูปที่ 3 – 5

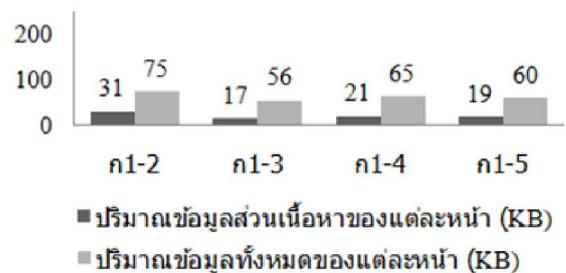
ตารางที่ 1 แสดง URL เว็บไซต์ที่ถูกนำมาใช้ในการทดลอง

อักษรย่อ	URL
ก	http://news.mthai.com/category/ general-news
ข	http://board.palungjit.com/f4/อภิญา- สมาธิ/
ค1	http://www.wikipedia.org/
ค2	http://dict.longdo.com/

การวิเคราะห์เว็บไซต์แต่ละแหล่งจะเลือกหน้าที่แตกต่างกันเพื่อนำมาเปรียบเทียบกันจำนวน 5 หน้า โดยมีวิธีการเลือกดังต่อไปนี้ (ก) หาหน้าที่จะนำมาวิเคราะห์โดยการเลือกข่าวที่อยู่ 5 อันดับแรกที่ปรากฏอยู่ในหน้านี้เพื่อเปิดเข้าไปดูรายละเอียดของข่าว (ผู้อ่านสามารถร่วมแสดงความคิดเห็นได้) (ข) หาหน้าที่จะนำมาวิเคราะห์โดยการเลือก 5 กระทู้แรกที่ถูกจัดอยู่ในหมวด "แนะนำเรื่องเด่น" (ค1) หาหน้าที่จะนำมาวิเคราะห์โดยค้นหาคำว่า "Web Search Engine" เมื่อได้หน้าแรกแล้วจึงเชื่อมโยงไปหน้าอื่นโดยการเลือก Link แรกที่ยังไม่เคยอ่านจากส่วน "See Also" ด้านล่างของหน้าที่ผ่านมาต่อไปอีก 4 หน้า (ค2) หาหน้าที่จะนำมาวิเคราะห์โดยค้นหาคำว่า "Web"; "Search"; "Engine"; "Dictionary", และ "ลองดู" ซึ่งแต่ละหน้า ผลลัพธ์จะแสดงความหมายของแต่ละคำ

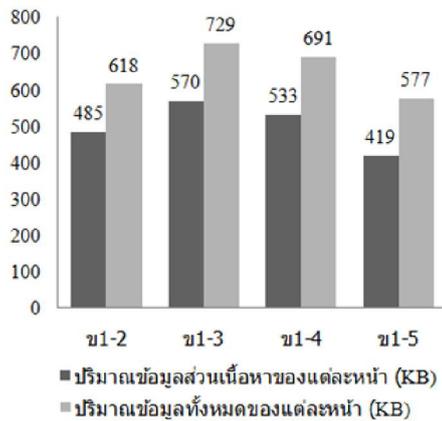
ผลการทดลองและวิจารณ์ผล

ผลการทดลองที่แสดงในแผนภูมิรูปที่ 3 - 5 ได้จากการเปรียบเทียบข้อมูลของหน้าที่มีการจัดรูปแบบแสดงเนื้อหาคล้ายคลึงกันจำนวน 5 หน้า จากแต่ละเว็บไซต์ที่แสดงในตารางที่ 1 โดยเอาหน้าแรกเป็นตัวตั้งเพื่อเปรียบเทียบกับหน้าที่เหลือจากรูปภาพแผนภูมิจะเห็นว่าเว็บไซต์ทุกประเภทมีข้อมูลส่วนที่ใช้จัดหน้าเนื้อหาซ้ำกันอยู่ภายใน แต่ละรูปจะแสดงกราฟแท่งจำนวนสี่คู่และมีข้อความกำกับอยู่ด้านล่างเพื่ออ้างอิงการเปรียบเทียบข้อมูลระหว่างหน้า เช่น ข้อความ "ก1-2" ในแผนภูมิภาพที่ 3 แสดงผลลัพธ์ของการเปรียบเทียบข้อมูลระหว่างหน้าแรกกับหน้าที่สอง เป็นต้น



ภาพที่ 3 แสดงแผนภูมิปริมาณข้อมูลส่วนที่ต่างจากหน้าแรกซึ่งเป็นข้อมูลส่วนเนื้อหาเปรียบเทียบกับข้อมูลรวมจากทั้งสองหน้าของเว็บไซต์ประเภท (ก)

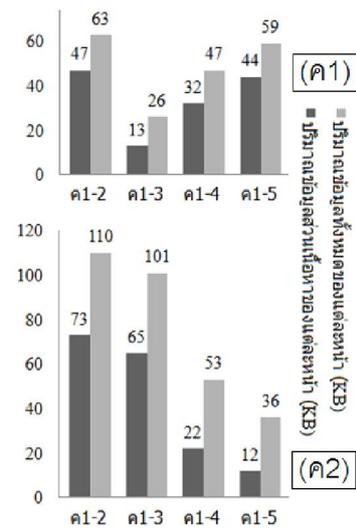
เว็บไซต์ประเภท (ก) ซึ่งแสดงเนื้อหาเกี่ยวกับหัวข้อข่าวประจำวัน แม้จะมีบริการให้ผู้อ่านสามารถแสดงความคิดเห็นได้ แต่จะถูกเปลี่ยนใหม่ทุกวัน ทำให้จำนวนความคิดเห็นมีน้อยเนื้อหาส่วนใหญ่เขียนสรุปเกี่ยวกับข่าวสั้นๆ จึงมีข้อมูลส่วนที่เป็นเนื้อหาน้อย ส่วนสาเหตุที่ข้อมูลส่วนการจัดหน้ามีมากกว่าเป็นเท่าตัวเนื่องจากในแต่ละหน้ามีการจัดแบ่งพื้นที่ลงโฆษณาอยู่จำนวนมาก



ภาพที่ 4 แสดงแผนภูมิปริมาณข้อมูลส่วนที่ต่างจากหน้าแรกซึ่งเป็นข้อมูลส่วนเนื้อหาเปรียบเทียบกับข้อมูลรวมจากทั้งสองหน้าของเว็บไซต์ประเภท (ข)

ในทางกลับกันเว็บไซต์ประเภท (ข) ที่ดึงข้อมูลจากกระตุ้แนะนำวิเคราะห์ เป็นกระตุ้ที่ถูกนำมาแสดงให้เด่นกว่ากระตุ้ปกติทั่วไปปรากฏอยู่เป็นเวลานานประมาณ 1 เดือน ทำให้มีผู้ตอบสนองความคิดเห็นจำนวนมาก อีกทั้งแต่ละความคิดเห็นยังมีกรอธิบายละเอียดเพราะเป็นเรื่องถกประเด็นเกี่ยวกับความเชื่อทางศาสนา จึงมีข้อมูลส่วนที่เป็นเนื้อหามาก สำหรับเว็บไซต์ประเภท (ค) ซึ่งแสดงผลการค้นหาข้อมูลจาก 2 แหล่ง ผลลัพธ์การค้นหาจากแหล่งค้นหาข้อมูล (ค1) มีการส่งข้อมูลการจัดหน้าและเนื้อหาตามปกติในขณะที่ข้อมูลจากแหล่ง (ค2) ซึ่งใช้การแสดงผลเนื้อหาในรูปแบบตารางทำให้มีข้อมูลส่วนที่ใช้ในการจัดหน้ามากเป็นพิเศษ

จากข้อมูลที่ได้จากการทดลองในขั้นต้นสามารถสรุปปริมาณข้อมูลรวมเพื่อเปรียบเทียบอัตราส่วนระหว่างส่วนเนื้อที่และส่วนการจัดหน้าโดยใช้แผนภูมิวงกลมดังภาพที่ 6 เว็บไซต์ประเภท (ก) มีอัตราส่วนข้อมูลการจัดหน้ามากที่สุด ซึ่งเป็นข้อมูลที่

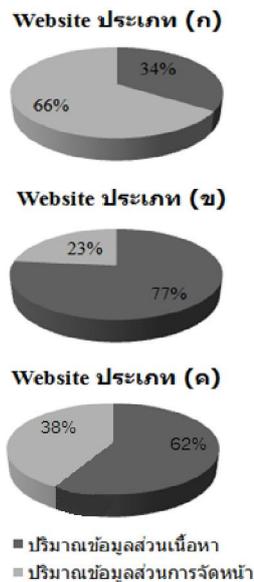


ภาพที่ 5 แสดงแผนภูมิปริมาณข้อมูลส่วนที่ต่างจากหน้าแรกซึ่งเป็นข้อมูลส่วนเนื้อหาเปรียบเทียบกับข้อมูลรวมจากทั้งสองหน้าของเว็บไซต์ประเภท (ค)

จะต้องส่งซ้ำทุกครั้งที่ใช้เปลี่ยนไปดูหน้าอื่นคือ 66% ส่วนเว็บไซต์ประเภท (ข) มีอัตราส่วนส่วนข้อมูลการจัดหน้าน้อยที่สุดคือ 23% แต่เมื่อเปรียบเทียบข้อมูลการจัดหน้าในเชิงปริมาณแล้ว เว็บไซต์ประเภท (ข) มีปริมาณข้อมูลส่วนการจัดหน้ามากที่สุด ในขณะที่เว็บไซต์ประเภท (ก) มีปริมาณข้อมูลส่วนการจัดหน้าน้อยที่สุด สำหรับเว็บไซต์ประเภท (ค) นั้นมีแนวโน้มอัตราส่วนส่วนข้อมูลการจัดหน้า น้อยเช่นเดียวกับเว็บไซต์ประเภท (ข) คือ 38% ของปริมาณที่ส่งผ่านระบบเครือข่ายอินเทอร์เน็ตทั้งหมด

สรุปผลการวิจัย

เนื่องจากข้อมูลหน้าเว็บไซต์ที่ส่งผ่านระบบเครือข่ายอินเทอร์เน็ตเข้ามายังเครื่องผู้ใช้ประกอบด้วยทั้งข้อมูลส่วนที่ใช้ในการจัดหน้าและข้อมูลที่เป็นส่วนเนื้อหาซึ่งอาจมากหรือน้อยตามแต่ลักษณะการใช้งานของเว็บไซต์แต่ละประเภท จากผลการทดลองโดยการนำข้อมูลจากเว็บไซต์ต่างๆ จำนวนสามประเภท



ภาพที่ 6 แสดงแผนภูมิเพื่อเปรียบเทียบอัตราส่วนปริมาณข้อมูลของเว็บไซต์ประเภทต่างๆ

(ทั้งหมด 4 เว็บไซต์) โดยสุ่มตัวอย่างมาวิเคราะห์เว็บไซต์ละ 5 หน้า ปริมาณข้อมูลที่ต้องส่งผ่านระบบเครือข่ายอินเทอร์เน็ตทั้งหมดคือ 3,366 KB (ไม่รวมปริมาณข้อมูลเอกสารอื่นๆที่ใช้อ้างอิงเชื่อมโยง เช่น ข้อมูลรูปภาพ เสียงเพลง ภาพยนตร์) ข้อมูลที่ถูกส่งประกอบด้วยสองส่วนคือ ข้อมูลส่วนที่เป็นเนื้อหา และ ข้อมูลส่วนที่ใช้ควบคุมให้โปรแกรม Web Browser จัดรูปแบบการแสดงผลซึ่งจะถูกส่งซ้ำซ้อนเพื่อใช้สำหรับการควบคุมแสดงผลแต่ละหน้า

การวิเคราะห์ทำโดยนำข้อมูลแต่ละหน้ามาเปรียบเทียบกัน โดยใช้โปรแกรม Exam Diff เพื่อหาปริมาณข้อมูลที่ถูกส่งผ่านระบบเครือข่ายอินเทอร์เน็ตซ้ำซ้อน พบว่ามีปริมาณมากถึง 963 KB และ ปริมาณข้อมูลที่เป็นส่วนเนื้อหาคือ 2,403 KB คิดเป็นอัตราส่วนประมาณ 29% ของปริมาณข้อมูลที่ถูกส่งทั้งหมด

อภิปรายผลและข้อเสนอแนะ

จากผลการวิจัยพบว่า การแสดงหน้าเว็บไซต์ซึ่งพัฒนาโดยใช้ภาษา HTML Script จะมีการส่งข้อมูลส่วนที่ใช้สำหรับควบคุมรูปแบบการจัดหน้าเพื่อแสดงเนื้อหาซ้ำซ้อนเสมอ ดังนั้นหากพัฒนาโปรแกรม "Web Browser" ให้สามารถเก็บข้อมูลส่วนที่ใช้สำหรับควบคุมรูปแบบการจัดหน้าเพื่อแสดงเนื้อหาเอาไว้ในเครื่องคอมพิวเตอร์ของผู้ใช้แล้วส่งเพียงข้อมูลเฉพาะส่วนที่เป็นเนื้อหาจริงๆเท่านั้น จะช่วยลดปริมาณข้อมูลที่ต้องส่งผ่านระบบเครือข่ายอินเทอร์เน็ตลงได้ หรือกล่าวอีกนัยหนึ่งคือ ระบบเครือข่ายอินเทอร์เน็ตจะสามารถรองรับข้อมูลได้มากขึ้นเมื่อความหนาแน่นของข้อมูลที่ส่งผ่านลดลงแล้ว ทำให้สามารถส่งข้อมูลไปยังปลายทางได้รวดเร็วยิ่งขึ้น

เอกสารอ้างอิง

- Adobe Systems. (n.d.). Adobe Flash. Retrieved December 30, 2011 from [http:// en.wikipedia.org/wiki/Macromedia_Flash](http://en.wikipedia.org/wiki/Macromedia_Flash)
- Veglis, A. (n.d). PHP and SQL Made Simple, review of Build Your Own Database Driven Website Using PHP & My SQL: A Practical Step-by-Step Guide, 3rd ed.," by Kevin Yank. 6(8) Retrieved August, 2005, from IEEE Distributed Systems.
- Microsoft. (n.d.). ActiveX. Retrieved December 30, 2011 from <http://en.wikipedia.org/wiki/ActiveX>.
- Sun Microsystems. (n.d.). Java Virtual Machine. Retrieved December 30, 2011 from http://en.wikipedia.org/wiki/Java_Runtime_Environment,
- Presto Soft. (n.d.). Exam Diff. Retrieved December 30, 2011 from http://www.prestosoft.com/edp_examdiff.asp