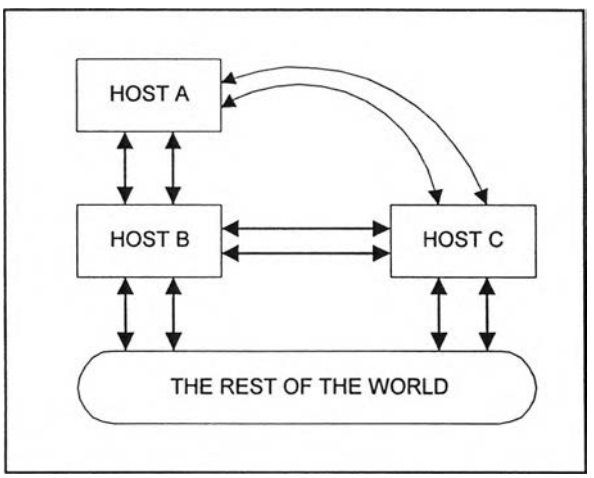
บทที่ 6

ผลการวิจัย

จากการวิจัยการจัดการระบบข่าวภายในเครือข่ายสถาบันอุดมศึกษานี้ สามารถลำดับเหตุการณ์ความก้าวหน้าของการวิจัยได้ ดังนี้

6.1 ศึกษาหลักการของยูสเนตนิวส์และการทำงานของไอเอ็นเอ็น

ช่วงต้นของการวิจัยได้ศึกษาหลักการของยูสเนตนิวส์ การทำงานของไอเอ็นเอ็น ตลอดจนการเข้าใจถึงปัญหาในเรื่องความสามารถในการรับส่งข้อมูลของเครือข่าย ซึ่งจากการศึกษาหลักการของยูสเนตนิวส์ พบว่ายูสเนตนิวส์ไม่มีแหล่งศูนย์กลางของข้อมูล (Smith, 1994) โดยข้อมูลยูสเนตนิวส์จะไหล (Flood) ไปตามตัวบริการข่าวทั่วโลก ดังในรูปที่ 6.1 ขั้นตอนวิธีการไหลยูสเนต (The Usenet Flooding Algorithm)



รูปที่ 6.1 ขั้นตอนวิธีการไหลยูสเนตนิวส์

ส่วนการทำงานของไอเอ็นเอ็น ดังที่ได้กล่าวไว้แล้วในบทที่ 2 ระบบข่าวและการจัดเก็บข้อมูลสถิติยูสเนตนิวส์

6.2 ทำการทดลองติดตั้งไอเอ็นเอ็นบนเครื่องบริการ

หลังจากได้ศึกษาหลักการของยูสเนตนิวส์และการทำงานของไอเอ็นเอ็น แล้วได้รับการอนุเคราะห์จากภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย ให้ทดลองทำการติดตั้งไอเอ็นเอ็น รุ่นที่ 1.5.1 (Salz, August 1992) ที่เครื่องบริการ เลขที่อยู่ไอพี 161.200.93.2 เป็นเครื่อง Sun Solaris 2.5.1 แรม 96 เมกะไบต์ ฮาร์ดดิสก์ 10 กิกะไบต์ ให้เป็นตัวบริการข่าวเพื่อศึกษาลักษณะการทำงานของไอเอ็นเอ็น ในการรับส่งยูสเนตนิวส์

6.3 ทำการป้อนและรับยูสเนตนิวส์จากตัวบริการข่าว

การทดลองทำการป้อนและรับยูสเนตนิวส์จากตัวบริการข่าวนี้ ได้รับการอนุเคราะห์จากสำนักเทคโนโลยีสารสนเทศ จุฬาลงกรณ์มหาวิทยาลัย ให้ทดลองป้อนยูสเนตนิวส์จากตัวบริการข่าวเลขที่อยู่ไอพี 161.200.192.5 มายังตัวบริการข่าวเลขที่อยู่ไอพี 161.200.93.2 ที่ภาควิชาวิศวกรรมคอมพิวเตอร์ จุฬาลงกรณ์มหาวิทยาลัย โดยรับยูสเนตนิวส์จากสำนักเทคโนโลยีสารสนเทศฯ ได้จำนวน 1,036 กลุ่มข่าว

6.4 จากการค้นคว้าการจัดการข้อมูลสถิติยูสเนตนิวส์ที่ผ่านมาพบว่า มีการใช้ซอฟต์แวร์เก็บข้อมูลสถิติยูสเนตนิวส์สำหรับไอเอ็นเอ็นอยู่ 3 ชุด คือโฟลว์สแตตส์ อินโฟลว์สแตต และไอเอ็นเอ็นรีพอร์ต

6.5 ศึกษาขั้นตอนการทำงานของซอฟต์แวร์โฟลว์สแตตส์ (Line, 1996) อินโฟลว์สแตต (Kugler, 1997) และไอเอ็นเอ็นรีพอร์ต (Tassin, 1996)

จากการศึกษาขั้นตอนการทำงานของซอฟต์แวร์ทั้งสาม สามารถแสดงรูปแบบการทำงานได้ดังแสดงในรูปที่ 2.3 รูปแบบการทำงานของโฟลว์สแตตส์ และอินโฟลว์สแตต และรูปที่ 2.4 รูปแบบการทำงานของไอเอ็นเอ็นรีพอร์ต ตามลำดับ

ต่อมาได้ทดลองติดตั้งซอฟต์แวร์โฟลว์สแตตส์ รุ่นที่ 1.5.2 (Line, 1996) เพื่อศึกษาขั้นตอนการทำงานและลักษณะการทำงานบนตัวบริการข่าวเลขที่อยู่ไอพี 161.200.93.2 ส่วนซอฟต์แวร์ไอเอ็นเอ็นรีพอร์ต และอินโฟลว์สแตต ไม่ได้ทดลองทำการติดตั้ง เนื่องจากต้องทำการลงคลังโปรแกรมดังต่อไปนี้บนตัวบริการข่าวก่อน

1. zlib (Grailly, 1995-1998)
2. libpng (Roelofs, 1998)
3. gd library (Boutell, 1996)
4. GD.pm (Stein, 1995-1999)

การติดตั้งคลังโปรแกรมต่าง ๆ ก่อน จึงเป็นเรื่องยุ่งยากในการติดตั้งซอฟต์แวร์ดังกล่าว นอกจากนั้นซอฟต์แวร์อินโฟลว์สแตตจริง ๆ แล้วถูกพัฒนาขึ้นมาให้เหมาะสมกับการติดต่อตัวบริการข่าวที่มีมากกว่า 2 เครื่องขึ้นไป ภายในเครือข่ายที่มีความสามารถในการรับส่งข้อมูลสูงเท่านั้น ไม่มีความเหมาะสมที่จะนำมาใช้กับเครือข่ายที่ติดต่อกับตัวบริการข่าวจากภายนอกเพียงแห่งเดียว

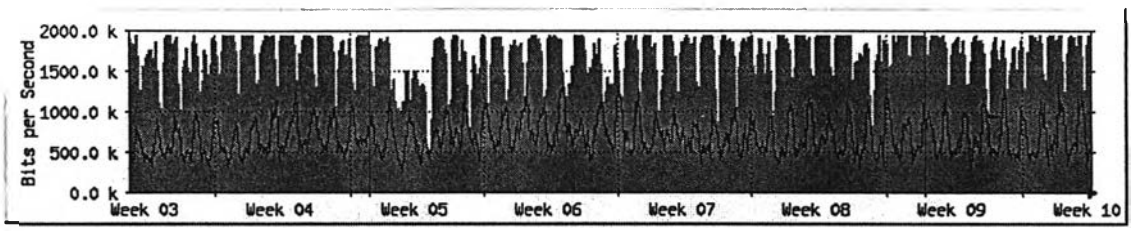
6.6 สรุปผลการศึกษาการจัดการข้อมูลยูสเนตนิวส์ที่ผ่านมาพบว่า

ซอฟต์แวร์ทั้งสามถูกพัฒนาขึ้นมา เพียงเพื่อแสดงรายงานข้อมูลสถิติยูสเนตนิวส์ให้แก่ผู้ใช้บริการทั่วไปดูเท่านั้น จึงไม่สามารถนำมาใช้ในการคัดเลือกกลุ่มข่าวได้ทั้งหมด แต่อย่างไรก็ตาม ซอฟต์แวร์โฟลว์สแตตส์ สามารถนำมาช่วยในการเก็บข้อมูลสถิติจำนวนไบต์ของแต่ละกลุ่มข่าว ซึ่งเป็นปัจจัยหนึ่งสำหรับการวิเคราะห์ข้อมูลเพื่อคัดเลือกกลุ่มข่าวได้ เนื่องจากมีการเก็บรายละเอียดของกลุ่มข่าวที่โดนครอสโพส (cross post) ด้วย

นอกจากนั้นการวิจัยครั้งนี้สามารถดัดแปลงและตัดตอนซอฟต์แวร์ไอเอ็นเอ็นรีพอร์ต ในการดึงข้อมูลจากแฟ้มข้อมูลล็อกของไอเอ็นเอ็น เพื่อใช้เก็บข้อมูลการเรียกใช้งานยูสเนตนิวส์ได้ โดยได้พัฒนาเป็นโปรแกรม usage.pl ซึ่งขั้นตอนการทำงานของโปรแกรมดูได้จากบทที่ 5 การพัฒนาโปรแกรม

6.7 เนื่องจากตัวบริการข่าวของสำนักเทคโนโลยีสารสนเทศ สามารถรับยูสเนตนิวส์ได้เป็นจำนวนน้อย เพียงประมาณ 1,387 กลุ่มข่าว ทำให้ไม่สามารถหาข้อมูลสถิติยูสเนตนิวส์ทั้งหมด 30,533 กลุ่มข่าวได้

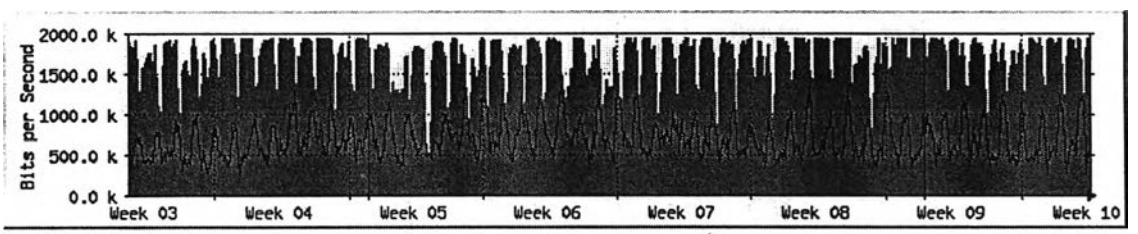
ดังเห็นได้จากรูปที่ 6.2 และรูปที่ 6.3 แสดงอัตราการใช้งานช่องสัญญาณของเครือข่ายจุฬาลงกรณ์มหาวิทยาลัย ซึ่งเป็นการเชื่อมโยงเครือข่ายอินเทอร์เนตจากเครือข่ายยูนิเนต (UNINET) ผ่านบริษัทเทเลโกลป(TeleGlobe) ที่ความเร็ว 2 เมกะบิตต่อวินาทีทั้ง 2 สาย ดังนี้



Max In: 1953.7 kb/s (98.5%) Average In: 1657.9 kb/s (83.6%) Current In: 1949.5 kb/s (98.3%)

Max Out: 1349.5 kb/s (68.0%) Average Out: 690.9 kb/s (34.8%) Current Out: 552.9 kb/s (27.9%)

รูปที่ 6.2 อัตราการใช้งานช่องสัญญาณของเครือข่ายจุฬาลงกรณ์มหาวิทยาลัยสายที่หนึ่ง (Internet Service Office [CU], 2000)



Max In: 1955.0 kb/s (98.5%) Average In: 1666.0 kb/s (84.0%) Current In: 1949.0 kb/s (98.2%)

Max Out: 1333.9 kb/s (67.2%) Average Out: 697.8 kb/s (35.2%) Current Out: 416.7 kb/s (21.0%)

รูปที่ 6.3 อัตราการใช้งานช่องสัญญาณของเครือข่ายจุฬาลงกรณ์มหาวิทยาลัยสายที่สอง (Internet Service Office [CU], 2000)

จากรูปที่ 5.2 แสดงอัตราการใช้งานช่องสัญญาณของเครือข่ายจุฬาลงกรณ์มหาวิทยาลัยสายที่หนึ่ง เลขที่อยู่ไอพี 207.45.217.42 และจากรูปที่ 5.3 แสดงอัตราการใช้งานช่องสัญญาณของเครือข่ายจุฬาลงกรณ์มหาวิทยาลัยสายที่หนึ่ง เลขที่อยู่ไอพี 207.45.214.234 มีอัตราการใช้งานช่องสัญญาณของข้อมูลเข้ามากที่สุดถึง 1,953.7 และ 1,955 กิโลบิตต่อวินาที ตามลำดับ

จากสถิติดังกล่าว เมื่อเทียบกับการรับยูสเนตนิวส์ทั้งหมดประมาณ 30,533 หัวข้อ ตามข้อมูลสถิติของสถาบันสแตนฟอร์ด (Noll, 1999) ยูสเนตนิวส์มีอัตราปริมาณการไหลของยูสเนตนิวส์เพียงอย่างเดียว 2.05 เมกะบิตต่อวินาที ซึ่งแสดงให้เห็นว่า ตัวบริการข่าวของสำนักเทคโนโลยีสารสนเทศไม่สามารถรับจำนวนกลุ่มข่าวทั้งหมดได้

6.8 การแก้ปัญหาในกรณีที่ไม่สามารถเก็บข้อมูลสถิติยูสเนตนิวส์ได้

จากการที่ตัวบริการข่าวของจุฬาลงกรณ์มหาวิทยาลัยไม่สามารถรับยูสเนตนิวส์ทั้งหมดได้ ทำให้ไม่สามารถเก็บรวบรวมข้อมูลสถิติยูสเนตนิวส์ทั้งหมด จึงได้ทำการติดต่อขอข้อมูลสถิติยูสเนตนิวส์จากผู้บริหารยูสเนตนิวส์ที่ให้บริการป้อนยูสเนตนิวส์เต็ม (Full Feed) ในหลายที่ โดยติดต่อผ่านทางจดหมายอิเล็กทรอนิกส์

ซึ่งต่อมาได้รับการตอบรับและความช่วยเหลือจาก Felix Kugler (personal communication, October 8, 1999) ผู้บริหารยูสเนตนิวส์ของ Swiss Academic & Research Network ที่ได้ส่งข้อมูลสถิติยูสเนตนิวส์จากตั้งแต่เดือนเมษายน ปี 1995 จนถึง เดือนตุลาคม ปี 1999 ทางไปรษณีย์มาให้ เป็นขนาดของข้อมูลสถิติยูสเนตนิวส์ประมาณ 3 กิกะไบต์ ซึ่งใช้ซอฟต์แวร์โฟลว์สแตตส์ เป็นเครื่องมือในการจัดเก็บแฟ้มข้อมูลสถิติยูสเนตนิวส์

6.9 การทำแบบสอบถามยูสเนตนิวส์

ได้จัดทำเว็บเพจการสำรวจยูสเนตนิวส์ขึ้น เพื่อทำการเก็บรวบรวมข้อมูลสถิติความต้องการของผู้ใช้บริการภายในเครือข่ายขึ้น โดยมีการพัฒนาโปรแกรมการค้นหาข้อมูล searchngs.pl และโปรแกรมเก็บรวมรวบผลการสำรวจผู้ต้องการใช้ยูสเนตนิวส์ ซึ่งผลของการวิจัยอยู่ระหว่างการดำเนินการ ส่วนรายละเอียดหน้าจอเว็บเพจ ดูได้ที่ภาคผนวก ก.

6.10 ทำการศึกษาหาวิธีการวิเคราะห์ เพื่อทำการคัดเลือกข้อมูลสถิติยูสเนตนิวส์

ในขั้นต้นได้ศึกษาหลักการวิเคราะห์การถดถอย (Regression Analysis) และโปรแกรมกำหนดการเชิงเส้น (Linear Programming) พบว่า ไม่สามารถนำมาช่วยแก้ปัญหาในการวิเคราะห์เพื่อคัดเลือกกลุ่มข่าวได้ เนื่องจากหลักการทางสถิติในการวิเคราะห์การถดถอยและกำหนดการเชิงเส้น มีลักษณะสมการคล้ายกันดังนี้

$$Y = AX_1 + BX_2 + C X_3 + \ldots + ZX_N$$

จากสมการข้างต้น มีตัวแปร X เป็นตัวอิสระ และตัวแปร Y เป็นตัวตาม เพื่อหาความสัมพันธ์ระหว่างตัวแปร Y กับตัวแปร X ที่ 1 จนถึงตัวแปร X ที่ N เมื่อมีการเปลี่ยนแปลงค่าสัมประสิทธิ์ของตัวแปร X เท่านั้น ซึ่งแสดงว่า จะต้องรู้จำนวนตัวแปร X แน่นอน ในขณะที่ปัญหา

ของการวิจัยนี้ ต้องการหาจำนวนตัวแปร X ว่าควรจะมีจำนวนเท่าใด ให้เพียงพอกับค่าของตัวแปร Y มากกว่า

ต่อมาจึงได้ทำการศึกษาเรื่องการสร้างแบบจำลอง (Simulation หรือ Model) เป็นการสร้างแบบจำลองทางคณิตศาสตร์ โดยการนำค่าข้อมูลสถิติของตัวแปรแต่ละตัวมาศึกษา ตัวแปรดังกล่าวมี ความโน้มเอียงของข้อมูลเป็นแบบใด เพื่อให้ได้ค่าสมการรวมของตัวแปรทั้งหมด แต่เนื่องจากต้องทำการจำลอง โดยใช้ข้อมูลทุกตัวมาทำการจำลอง ซึ่งไม่สามารถนำมาใช้กับการแก้ปัญหาของการวิจัยในครั้งนี้ได้ เพราะข้อมูลที่จะนำมาจำลองมีมากเกินไป

ภายหลังได้คำแนะนำจากอาจารย์ชัชวาล วงศ์ศิริประเสริฐ ในการวิเคราะห์ข้อมูลเพื่อทำการคัดเลือกด้วยแนวคิดในการแก้ปัญหาสูงสุด (Optimal Solution) ตามปัญหาแน็บแซค มีหลักการคือ การเลือกสิ่งของที่มีค่ามากใส่ลงในเป้ใบหนึ่งให้ได้มากที่สุดด้วย ซึ่งพบว่า สามารถนำมาปรับใช้ในการแก้ปัญหาการวิเคราะห์ข้อมูลเพื่อทำการคัดเลือกกลุ่มข่าวได้ โดยการพยายามเลือกกลุ่มข่าวที่ผู้ใช้บริการต้องการและมีการใช้งานให้ได้จำนวนกลุ่มข่าวมากที่สุดในขอบเขตของจำนวนความกว้างแถบความถี่ที่มีจำกัด

6.11 การใช้วิธีการแก้ปัญหาแน็บแซค

จากการศึกษาทำการค้นคว้าหาขั้นตอนวิธีการแก้ปัญหาแน็บแซค พบว่า มีขั้นตอนวิธีการแก้ปัญหาแน็บแซคได้หลายวิธี (Plane and McMillan, 1971: 17-20; Sedgewick, 1992: 596-598) เช่น การใช้วิธีการประมาณ (Approximation Method), การใช้วิธีการแจงนับ (Enumeration Method), การใช้วิธีการย้อนรอย (Backtracking Method) และการใช้วิธีการลดทอน (Reduction Method)

ซึ่งต่อมาได้พบการทำให้เกิดผลของขั้นตอนวิธีการแก้ปัญหาแน็บแซคโดยใช้วิธีการลดทอนที่ใช้ภาษาปาสกาลในการพัฒนาโปรแกรมของ Maciej M. Syslo, Nargingh Deo และ Janusz S. Kowalik (Syslo, Deo, et al 1983) ใช้เวลาในการประมวลผลเป็น $O(2^n)$ แต่จากการวิจัยในครั้งนี้ ต้องการทำให้เกิดผลบนเครื่องบริการ จึงได้มีการประยุกต์ใช้ขั้นตอนวิธีการแก้ปัญหาแน็บแซคนี้ด้วยภาษาเพิร์ล

ผลของการทำงานโดยใช้ขั้นตอนวิธีการแก้ปัญหาแน็บแซคมาใช้กับข้อมูลสถิติยูสเนตนิวส์นี้ ขั้นแรกทดลองทำการประมวลผลบนคอมพิวเตอร์ส่วนบุคคล ใช้เวลาประมาณ 4 ชั่วโมง

ครึ่ง ต่อมาเมื่อนำมาทดลองใช้กับเครื่องบริการ เลขที่อยู่ไอพี 161.200.93.2 ซึ่งเป็นเครื่อง Sun Solaris 2.5.1 แรม 96 เมกะไบต์ ฮาร์ดดิสค์ 10 กิกะไบต์ ได้ใช้เวลาการประมวลผล ดังในตารางที่ 6.1 แสดงเวลาที่ใช้การคัดเลือกกลุ่มข่าว แต่เนื่องจากเวลาการประมวลผลไม่สามารถตอบสนองต่อผู้บริหารยูสเนตนิวส์ได้ทันที ดังนั้นในการออกแบบระบบการคัดเลือกกลุ่มข่าว จึงออกแบบระบบให้ส่งจดหมายบอกผู้บริหารยูสเนตนิวส์เมื่อการประมวลผลข้อมูลเสร็จสิ้น

ตารางที่ 6.1 แสดงเวลาที่ใช้การคัดเลือกกลุ่มข่าว

จำนวนกลุ่มข่าว	เวลาที่ใช้ในการประมวลผล (นาที)
3,000 กลุ่มข่าว	3
5,000 กลุ่มข่าว	4
10,000 กลุ่มข่าว	10
15,000 กลุ่มข่าว	23
20,000 กลุ่มข่าว	65
25,000 กลุ่มข่าว	179
30,000 กลุ่มข่าว	185

6.12 การนำเสนอผลลัพธ์ในการคัดเลือกกลุ่มข่าว

การนำเสนอผลลัพธ์ที่ได้จากการคัดเลือกกลุ่มข่าวโดยใช้ขั้นตอนวิธีการแก้ปัญหาแน็บแซคนี้ ไม่ได้ใช้ GD.pm (Stein, 1995-1999) ซึ่งเป็นเพิร์ลโมดูล (Perl Module) เนื่องจากเกิดปัญหาความยุ่งยากในการลงคลังโปรแกรมดังได้กล่าวไว้แล้วข้างต้น

ดังนั้นการวิจัยครั้งนี้จึงได้เลือกใช้ ซอฟต์แวร์ gdchart-perl-1.0 (hempel, 1998) ของ Matt Hempler ในการนำเสนอข้อมูลผลลัพธ์ด้วยกราฟ ซึ่ง gdchart-perl-1.0 นี้ เป็นซอฟต์แวร์ที่ติดตั้งเพิร์ลโมดูล gdchart.pm ไว้สำหรับทำการเขียนกราฟ ซึ่งติดตั้งและใช้งานง่าย

จากการลำดับความก้าวหน้าของการวิจัยโดยรวม แล้วสามารถสรุปผลการวิจัยได้ดังในบทต่อไป