บทที่ 6
บทสรุป

งานวิจัยนี้เกิดขึ้นจากต้องการในการเพิ่มประสิทธิภาพในการประมวลผลยูอาร์แอลในการทำงานของเว็บพร็อกซี่โดยการย่อหรือลดขนาดยูอาร์แอลด้วยการเข้ารหัส เว็บพร็อกซี่โดยทั่วไปนิยมใช้การเข้ารหัสแบบ MD5 ซึ่งเป็นที่น่าสนใจว่าในความเป็นจริงแล้ว วิธี MD5 นั้นมีความเหมาะสมสำหรับแอปพลิเคชั่นเว็บทุกกรณีหรือไม่ จากผลการวิจัยทำให้เราทราบว่า การพิจารณาวิธีการเข้ารหัสที่เหมาะสมนั้น จะต้องพิจารณาความต้องการของแอปพลิเคชั่นด้วย

การวิจัยนี้ได้ทำการเปรียบเทียบคุณลักษณะทางด้าน ความเร็วในการเข้ารหัส ความยาวของรหัสที่ได้ และปริมาณการชนกันของรหัส ของวิธีการเข้ารหัสต่างๆ ที่เลือกขึ้นมา 12 วิธี ดังนี้ คือ MD2, MD4, MD5, SHA-1, CRC-16, CRC-CCITT, CRC-32, Division Method, Folding Method, Digit Analysis Method, Midsquare Method และ Huffman Coding การพัฒนาโปรแกรมซึ่งใช้ในการเข้ารหัสนั้นใช้ภาษาซีในการพัฒนาทั้งหมด และข้อมูลที่ใช้ในการทดลองนำมากจากสำนักเทคโนโลยีสารสนเทศ จุฬาลงกรณ์มหาวิทยาลัย ระหว่างวันที่ 22-26 พฤศจิกายน 2542 ข้อมูลที่ได้อยู่ในรูปแบบของสควิดล็อกไฟล์ สำหรับการทำการทดลองจะทำการทดลองซ้ำกับล็อกไฟล์เดิมจำนวน 5 ครั้ง จากนั้นนำผลที่ได้มาวิเคราะห์และสรุปผลการทดลองดังแสดงในบทที่ 4 และ 5 ตามลำดับ และสำหรับในบทนี้จะเป็นการสรุปงานวิจัยทั้งหมด ซึ่งมีรายละเอียดดังนี้

6.1 สรุปผลการวิจัย

จากการวิเคราะห์ผลการทดลองในบทที่ 5 ซึ่งได้จากการทดสอบการเข้ารหัสยูอาร์แอลจากข้อมูลการเรียกขอ ของสำนักเทคโนโลยีสารสนเทศ จุฬาลงกรณ์มหาวิทยาลัย ด้วยอัลกอริทึมในการเข้ารหัสแบบต่างๆ สามารถสรุปลักษณะโดยทั่วๆไปของอัลกอริทึมที่ใช้ในการเข้ารหัสที่นำมาทดสอบได้ดังนี้

- อัลกอริทึมที่ใช้เวลาในการเข้ารหัสน้อย มักเป็นอัลกอริทึมไม่ซับซ้อน แต่ก็มีแนวโน้มที่เกิดการชนกันของรหัสสูง
- อัลกอริทึมที่มีความยาวของรหัสมากกว่า มีแนวโน้มที่จะมีปริมาณการชนกันของรหัสน้อยกว่า
- อัลกอริทึมที่พบปริมาณการชนกันของรหัสน้อย มีแนวโน้มที่ใช้ใช้เวลาในการเข้ารหัสมากกว่าอัลกอริทึมที่มีปริมาณการชนกันของรหัสมากกว่า
- อัลกอริทึมที่มีขนาดรหัสสั้นกว่ามีประสิทธิภาพในการลดขนาดยูอาร์แอลสูงกว่า

จากลักษณะทั่วๆไปที่ได้จากการวิเคราะห์ผลการทดลองข้างต้นสามารถสรุปได้ว่า ไม่มีอัลกอริทึมใดที่เหมาะสมกับงานทุกๆงาน เนื่องจากแต่ละอัลกอริทึมมีข้อดี ข้อเสีย ที่แตกต่างกันออกไป อีกทั้งแอพปลิเคชั่นที่ใช้งานการเข้ารหัสก็มีความต้องการที่แตกต่างกัน งานวิจัยนี้จึงเพียงเสนอแนะแนวทางในการเลือกอัลกอริทึมให้เหมาะสมสำหรับแอพปลิเคชั่นเท่านั้น ซึ่งจากการศึกษาพบว่าแอพปลิเคชั่นส่วนใหญ่ที่ใช้การเข้ารหัสยูอาร์แอล มีลักษณะความต้องการดังนี้

- ***แอปพลิเคชั่นที่ต้องการความเร็วในการเข้ารหัสและรหัสที่สั้นเป็นหลัก*** เช่น การตัดสินในการเลือกตำแหน่งที่จะแคชข้อมูลใน Super Proxy [20] เป็นต้น ควรเลือกอัลกอริทึมในการเข้ารหัสที่ไม่ซับซ้อน เพื่อให้สามารถทำงานได้อย่างรวดเร็ว และเลือกอัลกอริทึมที่มีขนาดรหัสสั้น เนื่องจากสิ้นเปลืองทรัพยากรระบบในการประมวลผลต่ำกว่า แต่ต้องการให้มีการชนกันของรหัสเกิดขึ้น (เพื่อใช้เป็นค่าในการตัดสินใจเลือกตำแหน่งของแคช) ซึ่งอัลกอริทึมที่เหมาะสมคือ Digit Analysis method

- ***แอปพลิเคชั่นที่ต้องการความเร็วในการเข้ารหัสและปริมาณการชนกันของรหัสน้อยหรือไม่มีเลยเป็นหลัก*** เช่น การวิเคราะห์ข้อมูลเรียกขอหรือล็อกไฟล์ (log file) หรือการสอบถามข้อมูลว่ามีอยู่ในแคชหรือไม่ เป็นต้น ควรเลือกอัลกอริทึม MD4 เนื่องจากไม่มีการชนกันของรหัสเลย และเวลาที่ใช้น้อยกว่าอัลกอริทึม MD อื่นๆ

- ***แอปพลิเคชั่นที่ต้องการรหัสสั้นและปริมาณการชนกันของข้อมูลเป็นหลัก*** เช่น ใช้ในการสอบถามถึงข้อมูลที่มีอยู่ระหว่างเว็บพร็อกซี่ หรือใช้การเข้ารหัสในโปรโตคอลที่สอบถามข้อมูลระหว่างเว็บพร็อกซี่ด้วยกัน แอปพลิเคชั่นประเภทนี้ ยอมให้เกิดการชนกันของรหัสได้บ้าง แต่รหัสที่สั้นจะช่วยเพิ่มประสิทธิภาพในการใช้ช่องทางการสื่อสารมากขึ้น อัลกอริทึมที่น่าจะใช้งานในแอปพลิเคชั่นลักษณะนี้ คือ อัลกอริทึม CRC-32, Folding method 4 ไบต์ และ 8 ไบต์

จากผลการวิจัยทำให้ทราบถึงแนวทางในการเลือกใช้อัลกอริทึมที่มีความเหมาะสมในการเข้ารหัสยูอาร์แอล และชี้ให้เห็นว่าไม่มีอัลกอริทึมในการเข้ารหัสใดเพียงอัลกอริทึมเดียวที่ดีที่สุดในการเข้ารหัสยูอาร์แอล ยังอาจมีอัลกอริทึมอื่นนอกจาก MD5 ที่มีประสิทธิภาพในการเข้ารหัสยูอาร์แอล และการเลือกใช้นั้นจะต้องขึ้นอยู่กับความต้องการของแอปพลิเคชั่น

อย่างไรก็ตามผลการวิจัยพบว่าอัลกอริทึมที่มีประสิทธิภาพโดยรวมที่ดีที่สุดในจำนวนอัลกอริทึมที่ใช้ในงานวิจัยทั้งหมด คือ CRC-32 เนื่องจากใช้เวลาเข้ารหัสน้อย ประสิทธิภาพในการลดขนาดยูอาร์แอลสูง และเกิดการชนกันของรหัสน้อยมาก

6.2 ปัญหาและข้อจำกัดที่พบในการวิจัย

- การจับเวลาในการเข้ารหัส ไม่สามารถจับเวลาที่ความละเอียดต่ำกว่า 1 ไมโครวินาทีได้ เนื่องจากเป็นข้อจำกัดของตัวแปรภาษา
- ในการคัดเลือกยูอาร์แอลที่แคชได้นั้น ทำได้ไม่ครอบคลุมทุกกรณี เนื่องจากการระบุว่ายูอาร์แอลใดเป็นยูอาร์แอลที่ไม่สามารถแคชได้ทำได้ค่อนข้างยาก ในการวิจัยนี้จะใช้ส่วนขยายและเครื่องหมายคำถามในยูอาร์แอลเป็นตัวระบุว่ายูอาร์แอลนั้นสามารถแคชได้หรือไม่ ซึ่งในความเป็นจริงแล้ว อาจมีส่วนขยายอื่นๆที่ไม่สามารถแคชได้นอกเหนือจากที่ได้แสดงในการวิจัยนี้
- ข้อมูลยูอาร์แอลที่ได้ใช้ในการวิจัย ได้มาจากสควิดล็อกไฟล์ ซึ่งในแต่ละวันแบ่งออกเป็น 3 ไฟล์ และทั้ง 3 ไฟล์มีข้อมูลบางส่วนที่ซ้ำซ้อนกันอยู่ รวมทั้งมีการเหลื่อมเวลาในการเก็บข้อมูลแต่ละวันด้วย อันเนื่องมาจากขนาดของฮาร์ดดิสก์ที่จำกัด ทำให้จำนวนยูอาร์แอลที่ใช้ในแต่ละวันนั้นคลาดเคลื่อนจากความเป็นจริงไปเล็กน้อย
- ในการทำการทดลองแต่ละครั้งใช้เวลาในการทดลองที่ค่อนข้างนาน อีกทั้งบางครั้งรหัสที่ได้มีขนาดไฟล์ใหญ่มาก ทำให้เกิดปัญหาในการจัดเก็บ และเมื่อต้องการทำการทดลองใหม่จะต้องลบผลลัพธ์เดิมทิ้ง ทำให้ไม่สะดวกในการเปรียบเทียบผล

6.3 ข้อเสนอแนะ

- สำหรับการทำวิจัยซ้ำ หรือต้องการวิจัยต่อเนื่อง จะต้องคำนึงถึงพื้นที่ในการเก็บผลลัพธ์หรือรหัสที่ได้ด้วย เนื่องจากบางครั้งผลลัพธ์ที่ได้มีขนาดใหญ่
- จากผลการทดลองพบว่าอัลกอริทึม CRC-32, Folding method ที่มีความยาวรหัสมากกว่า 2 ไบต์ พบว่านอกจากจะใช้เวลาในการเข้ารหัสต่ำ ความยาวของรหัสเพียง 32 – 128 บิต และมีอัตราการชนกันของรหัสต่ำอีกด้วย ถ้าหากมีการปรับปรุงโดยการรวมอัลกอริทึมทั้งสองเข้าด้วยกัน อาจทำให้เกิดอัลกอริทึมที่ไม่เกิดการชนกันของข้อมูลเลยและสามารถทำงานได้เร็วกว่าอัลกอริทึมที่ไม่มีการชนกันของรหัสที่ใช้อยู่ในปัจจุบัน