บทที่ 1
บทนำ

การค้นคว้าแบบอิสระนี้ต้องการที่จะนำเสนอแนวคิดและกระบวนวิธีในการทำงานร่วมกันระหว่างการบีบอัดและการเข้ารหัส รวมไปถึงการพัฒนาโปรแกรมที่ใช้สำหรับบีบอัดข้อมูลและเข้ารหัสข้อมูลส่วนบุคคล พร้อมทั้งแสดงวัตถุประสงค์และประโยชน์ที่คาดว่าจะได้รับจากการศึกษาค้นคว้า นอกจากนี้ยังได้แสดงขอบเขตของงานวิจัยและวิธีการวิจัยอย่างเป็นขั้นตอน เพื่อนำไปสู่การบรรลุผลสำเร็จในการทำวิจัย และแสดงผลลัพธ์ในเชิงเปรียบเทียบประสิทธิภาพในการทำงานร่วมกันของแต่ละอัลกอริทึมที่ใช้กับไฟล์ประเภทต่างๆ

1.1 ที่มาและความสำคัญของปัญหา

การทำงานในระบบคอมพิวเตอร์ปัจจุบัน ว่าได้ทำให้เกิดข้อมูลอิเล็กทรอนิกส์ขึ้นมามากมาย ซึ่งข้อมูลเหล่านั้นบางส่วนอาจจำเป็นที่จะต้องมีการปกปิดไม่ให้ผู้อื่นที่ไม่ใช่กลุ่มเป้าหมายของผู้ใช้งานล่วงรู้ เช่น เอกสารลับ ข้อมูลสำคัญทางธุรกิจ รูปภาพส่วนตัว เป็นต้น ผู้ใช้ที่ต้องการปกปิดข้อมูลดังกล่าว อาจเป็นผู้ที่ต้องใช้งานคอมพิวเตอร์ร่วมกันกับผู้อื่น หรือผู้ที่ต้องการส่งข้อมูลอิเล็กทรอนิกส์ให้กับผู้รับผ่านทางช่องทางต่างๆ ดังนั้น สิ่งที่จะช่วยทำให้ข้อมูลอิเล็กทรอนิกส์เป็นความลับวิธีหนึ่งคือ วิทยาการเข้ารหัสลับ (Cryptography)

การเข้ารหัสข้อมูลโดยพื้นฐานแล้วก็คือการทำให้ข้อมูลที่เข้ารหัส เปลี่ยนแปลงไปจากข้อมูลตั้งต้นเดิมจนไม่สามารถอ่านได้หากไม่มีกุญแจ (Key) ที่ใช้ในการถอดรหัสนั้นๆ เราเรียกกระบวนการแปรรูปข้อมูลตั้งแต่ต้นว่า "การเข้ารหัสลับ" (Encryption) และเรียกกระบวนการแปลงข้อมูลที่ผ่านการเข้ารหัสแล้ว ให้กลับมาอยู่ในรูปของข้อมูลดั้งเดิมว่า "การถอดรหัสลับ" (Decryption) ซึ่งกระบวนการเหล่านี้จะเกี่ยวข้องกับวิธีการทางคณิตศาสตร์ เรียกว่า อัลกอริทึมในการเข้ารหัสลับ (Encryption Algorithm)

ต่อมาเมื่อมีการศึกษาด้านการเข้ารหัสลับอย่างแพร่หลาย ทำให้กระบวนการเข้ารหัสถูกเปิดเผยมากขึ้น การรักษาความปลอดภัยของข้อมูลด้วยวิธีการเข้ารหัสลับเพียงอย่างเดียวจึงไม่สามารถกล่าวได้ว่าปลอดภัยอย่างสมบูรณ์ การเข้ารหัสลับข้อมูลเพียงอย่างเดียวนั้น ไม่ใช่วิธีการรักษาความปลอดภัยที่ดีที่สุดเสมอไป เพราะหากข้อมูลนั้นถูกถอดรหัสลับได้ หรือผู้ประสงค์ร้ายทราบวิธีในการถอดรหัสลับนั้นๆ ข้อมูลที่เป็นความลับทั้งหมดก็จะถูกเปิดเผย การศึกษาในครั้งนี้จึงขอเสนอวิธีการเพิ่มความปลอดภัยให้กับข้อมูลด้วยการใช้เทคนิคการบีบอัดข้อมูล (Compression) ควบคู่กับการเข้ารหัสลับ โดยแนวคิดของการศึกษานี้เพื่อเป็นการเปรียบเทียบประสิทธิภาพในด้าน

ต่างๆ ได้แก่ เวลาในการดำเนินการ ขนาดของผลลัพธ์ อัตราส่วนในการบีบอัด ประสิทธิภาพในการบีบอัด และประสิทธิภาพด้านความปลอดภัย ของการทำงานร่วมกันระหว่างการบีบอัดข้อมูลและการเข้ารหัสลับ โดยการบีบอัดข้อมูลนั้นได้เลือกอัลกอริทึมที่เป็นแบบไม่สูญเสียข้อมูล (Lossless Data Compression) และอัลกอริทึมที่ใช้ในการเข้ารหัสลับได้เลือกชนิดการเข้ารหัสแบบสมมาตร (Symmetric Cryptography) เพื่อให้ได้ประสิทธิภาพด้านความเร็วที่เหมาะกับการใช้งานส่วนบุคคลมากกว่า

1.2 วัตถุประสงค์ของการศึกษา

1) เพื่อศึกษาเปรียบเทียบประสิทธิภาพการทำงานร่วมกันระหว่างการบีบอัดและการเข้ารหัส
2) เพื่อพัฒนาโปรแกรมเข้ารหัสและการบีบอัดเพื่อใช้งานส่วนบุคคล

1.3 ประโยชน์ที่จะได้รับจากการศึกษา

1) ทราบถึงแนวทางการนำเอาเทคนิคการบีบอัดข้อมูลมาใช้ร่วมกับการเข้ารหัสลับแต่ละแบบ
2) สามารถเปรียบเทียบประสิทธิภาพของการใช้การบีบอัดร่วมกับการเข้ารหัสแต่ละแบบได้
3) ได้โปรแกรมเข้ารหัสและการบีบอัดเพื่อใช้งานส่วนบุคคล

1.4 ขอบเขตและวิธีการศึกษา

1.4.1 ขอบเขตของการศึกษา

ในการศึกษาครั้งนี้ ได้ทำการแบ่งขอบเขตในการศึกษาออกเป็นด้านต่างๆ คือ ขอบเขตด้านอัลกอริทึมที่ใช้ในการเข้ารหัสลับ ด้านอัลกอริทึมที่ใช้ในการบีบอัดข้อมูล ด้านข้อมูลนำเข้าและผลลัพธ์ที่ได้จากโปรแกรม กำหนดเกณฑ์ที่ใช้วัดประสิทธิภาพด้านต่างๆ และการสรุปวิเคราะห์เพื่อนำเสนอแนวทางในการใช้งาน

1) อัลกอริทึมที่ใช้ในการเข้ารหัสลับ เลือกใช้แบบกุญแจสมมาตร ได้แก่
   - DES
   - Triple DES
   - AES
   - Blowfish
   - RC4

2) อัลกอริทึมที่ใช้ในการบีบอัดข้อมูล เลือกใช้แบบที่ไม่เกิดการสูญเสีย ได้แก่

- Huffman Coding
- Deflate

3) การพัฒนาโปรแกรม มีข้อกำหนดในส่วนของข้อมูลนำเข้าและผลลัพธ์ที่ได้ ดังนี้

- ข้อมูลนำเข้า เป็นไฟล์ข้อมูลที่มีขนาดไม่เกิน 5 เมกกะไบต์
- ผลลัพธ์ แสดงผลในรูปแบบของอักขระ หรือแฟ้มข้อความ (Text File)

4) กำหนดเกณฑ์ (Criteria) ที่ใช้วัดประสิทธิภาพด้านต่างๆ ได้แก่

- เวลาที่ใช้ในการดำเนินการ (milliseconds)
- ขนาดของผลลัพธ์ (bytes)
- อัตราส่วนในการบีบอัด

$$\text{Compression Ratio (\%)} = \frac{(\text{ขนาดก่อนบีบอัด} - \text{ขนาดหลังบีบอัด}) \times 100}{\text{ขนาดก่อนบีบอัด}}$$

- ประสิทธิภาพในการบีบอัด

$$\text{Throughput (bytes/second)} = \frac{(\text{ขนาดก่อนบีบอัด} - \text{ขนาดหลังบีบอัด}) \times 1000}{\text{เวลาที่ใช้ดำเนินการ (milliseconds)}}$$

- ประสิทธิภาพด้านความปลอดภัย

$$\text{Brute Force Time (years)} = \frac{2^k}{1000 \text{ MIPS} \times Y}$$

k	หมายถึง ขนาดของกุญแจที่ใช้ (bits)
1000 MIPS	หมายถึง การคำนวณจากคอมพิวเตอร์ที่สามารถทำงานได้ 1000 ล้านคำสั่งต่อวินาที
Y	หมายถึง เวลาใน 1 ปี

5) การสรุปแนวทางในการใช้งานร่วมกันระหว่างการบีบอัดและการเข้ารหัสลับ จะใช้ผลลัพธ์ที่ได้จากการวัดประสิทธิภาพด้านต่างๆ มาวิเคราะห์เพื่อสรุปผลที่ได้ในแต่ละด้าน ของแต่ละอัลกอริทึมว่ามีเหตุมีผลอย่างไรจึงได้ผลลัพธ์ดังกล่าว รวมถึงสรุปแนวทางในการใช้งานร่วมกันให้ได้ประสิทธิภาพที่เหมาะสมกับการใช้งานในลักษณะต่างๆ

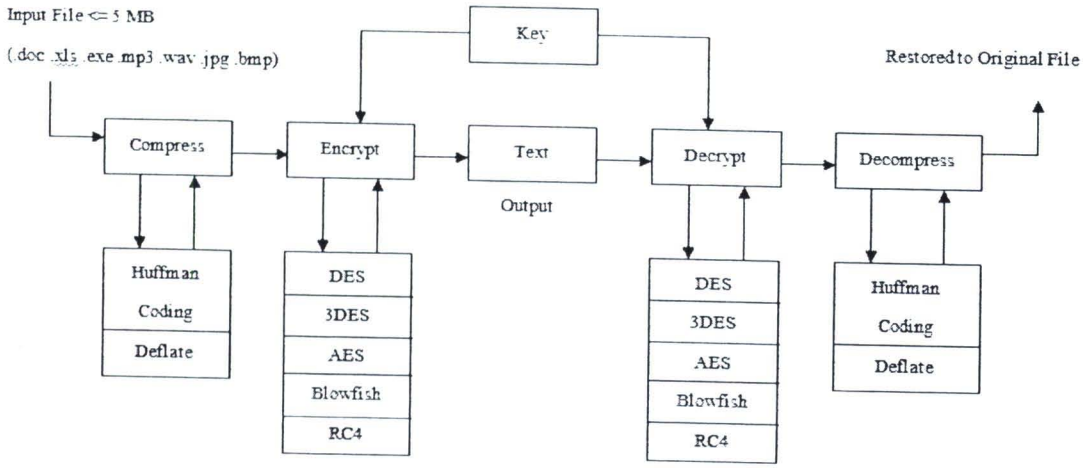
1.4.2 วิธีการศึกษา

1) กำหนดปัญหา และวัตถุประสงค์ในการศึกษาเปรียบเทียบประสิทธิภาพการทำงานร่วมกันระหว่างการบีบอัดและการเข้ารหัสลับ
2) ศึกษางานวิจัยที่เกี่ยวข้องกับการบีบอัดและการเข้ารหัสลับ เพื่อหาแนวทางการนำมาใช้ให้เหมาะกับลักษณะงาน
3) ศึกษางานวิจัยที่เกี่ยวกับการเปรียบเทียบประสิทธิภาพ เพื่อหาแนวทางการกำหนดเกณฑ์วัดประสิทธิภาพในการศึกษา
4) กำหนดขอบเขตอัลกอริทึมที่ใช้ ข้อกำหนดในการพัฒนาโปรแกรม และเกณฑ์การวัดประสิทธิภาพในแต่ละด้าน
    - กำหนดอัลกอริทึมในการเข้ารหัสลับ
    - กำหนดอัลกอริทึมในการบีบอัดข้อมูล
    - กำหนดข้อมูลนำเข้า และผลลัพธ์ของโปรแกรมที่จะพัฒนา
    - กำหนดเกณฑ์การวัดประสิทธิภาพในแต่ละด้านโดยศึกษาจากงานวิจัยที่เกี่ยวข้อง
5) ออกแบบโปรแกรมเข้ารหัสลับร่วมกับการบีบอัด
    - ออกแบบกระบวนการทำงานของโปรแกรมด้วยแผนภาพบล็อก (Block Diagram)
6) พัฒนาโปรแกรมด้วยไมโครซอฟท์วิชวลเบสิค 2008
    - ประกอบไปด้วยส่วนของการเข้ารหัส และส่วนของการถอดรหัส
    - ในแต่ละส่วนจะมีตัวเลือกของอัลกอริทึมที่ใช้ในการเข้ารหัสลับและการบีบอัด
7) ทดสอบโปรแกรมด้วยข้อมูลนำเข้าที่มีขนาดและชนิดแตกต่างกัน
    - ข้อมูลนำเข้า ประกอบไปด้วยชนิดของไฟล์ที่มีนามสกุลเป็น .doc, .xls, .exe, .mp3, .wav, .jpg และ .bmp
    - ขนาดของไฟล์ที่นำเข้ามีขนาดต่างกันแต่ไม่เกิน 5 เมกกะไบต์
8) เปรียบเทียบประสิทธิภาพการทำงานร่วมกันของการบีบอัดและการเข้ารหัสลับแต่ละแบบตามข้อมูลนำเข้าที่มีชนิดและขนาดแตกต่างกัน ตามเกณฑ์ที่กำหนด
    - เปรียบเทียบโดยแยกตามตัวชี้วัดแต่ละด้านตามเกณฑ์ที่กำหนดไว้ในขอบเขต
    - นำมาสรุปรวมเป็นตารางเปรียบเทียบประสิทธิภาพในแต่ละด้านตามลำดับ

- วิเคราะห์เพื่อสรุปผลที่ได้ในแต่ละด้าน ของแต่ละอัลกอริทึมว่ามีเหตุมีผลอย่างไรจึงได้ผลลัพธ์ดังกล่าว
- สรุปแนวทางในการใช้งานร่วมกันให้ได้ประสิทธิภาพที่เหมาะสมกับการใช้งานในลักษณะต่างๆ

9) จัดทำเอกสารสรุปผลการศึกษาและคู่มือการใช้งานโปรแกรม

1.4.3 กระบวนการทำงานของโปรแกรม



1.5 สถานที่ที่ใช้ในการดำเนินการศึกษาและรวบรวมข้อมูล

1) สถานที่

(1) ภาควิชาวิทยาการคอมพิวเตอร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเชียงใหม่

(2) สำนักหอสมุด มหาวิทยาลัยเชียงใหม่

2) อุปกรณ์ที่ใช้ในการดำเนินการวิจัย

(1) คอมพิวเตอร์ส่วนบุคคล จำนวน 1 ชุด

(2) คอมพิวเตอร์โน้ตบุ๊ค จำนวน 1 เครื่อง

3) ซอฟต์แวร์

(1) ระบบปฏิบัติการไมโครซอฟท์วินโดว์เอกซ์พี (Microsoft Windows XP)

(2) โปรแกรมไมโครซอฟท์ออฟฟิซ 2003 (Microsoft Office 2003)

(3) โปรแกรมไมโครซอฟท์วิชวลเบสิค 2008 (Microsoft Visual Basic 2008)