

บทที่ 1

บทนำ

ความเป็นมาของปัญหา

เทคโนโลยีสารสนเทศมีบทบาทที่สำคัญอย่างยิ่งในกิจการต่าง ๆ มีการนำเทคโนโลยีสารสนเทศมาช่วยในการปฏิบัติงานให้เกิดประโยชน์ขึ้นในหน่วยงาน โดยแต่ละหน่วยงานพยายามจัดเก็บข้อมูลที่มีปริมาณมากอย่างถูกต้องเป็นระเบียบ รวมทั้งพยายามหาวิธีการค้นหาข้อมูลที่มีความรวดเร็ว และมีประสิทธิภาพ ด้วยเทคโนโลยีด้านคอมพิวเตอร์ทำให้สามารถจัดเก็บข้อมูลได้ปริมาณมาก การประยุกต์ใช้ข้อมูลที่เก็บอยู่ในฐานข้อมูลขนาดใหญ่จำเป็นต้องมีวิธีการสืบค้นข้อมูลอย่างมีประสิทธิภาพ เพื่อให้สามารถนำข้อมูลมาใช้ให้เกิดประโยชน์สูงสุด ข้อมูลเหล่านี้สามารถนำมาวิเคราะห์เพื่อนำไปสร้างระบบสนับสนุนการตัดสินใจ (Decision Support System) เพื่อประโยชน์ในการบริหารจัดการขององค์กร

ปัจจุบันมีการนำเทคโนโลยีสารสนเทศมาใช้งานในโรงพยาบาลด้านต่าง ๆ เช่น การบริหารจัดการงานด้านเวชระเบียน การวินิจฉัยโรค การรักษาพยาบาล และการจ่ายยาให้ผู้ป่วย เป็นต้น มีการเก็บข้อมูลจำนวนมาก เพื่อนำมาใช้ช่วยตัดสินใจในการวินิจฉัยโรค ประกอบด้วย ระบบสารสนเทศในโรงพยาบาล ข้อมูลสัญญาณชีพ ข้อมูลการตรวจทางห้องปฏิบัติการ และข้อมูลเภสัชกรรม

ปัญหาที่พบในปัจจุบันก็คือเมื่อข้อมูลมีจำนวนมากมายมหาศาลทำให้การสืบค้นข้อมูลใช้เวลายาวนาน จึงจำเป็นต้องศึกษาเทคนิคต่าง ๆ ที่ใช้แก้ปัญหา ในงานวิจัยนี้ผู้วิจัยสนใจใช้เทคนิคแฮชฟังก์ชันโดยนำโครงสร้างเอ็กเทนดิเบิล แฮชซิง (Extendible Hashing) มาใช้แก้ปัญหาเพื่อให้การสืบค้นข้อมูลทำได้อย่างรวดเร็วและเหมาะแก่การนำมาใช้ในการวินิจฉัยโรคในโรงพยาบาล

วัตถุประสงค์ของการวิจัย

1. เพื่อเพิ่มประสิทธิภาพการสืบค้นข้อมูลการวินิจฉัยโรคด้วยแฮชฟังก์ชันโดยใช้โครงสร้างเอ็กเทนดิเบิล แฮชซิง
2. เพื่อเปรียบเทียบประสิทธิภาพการทำงานระหว่างวิธีการสืบค้นข้อมูลที่ใช้ภาษาสอบถามเชิงโครงสร้าง และวิธีการสืบค้นข้อมูลด้วยเอ็กเทนดิเบิล แฮชซิงร่วมกับภาษาสอบถามเชิงโครงสร้าง

สมมติฐานการวิจัย

งานวิจัยนี้ได้เสนอวิธีการสืบค้นข้อมูลการวินิจฉัยโรคด้วยเทคนิคการทำแฮชซิง โดยใช้โครงสร้างข้อมูลเอ็กเทนดิเบิล แฮชซิง ซึ่งข้อมูลจะถูกเก็บในตารางแฮช ภายในตารางจะมีช่องเก็บข้อมูล โดยแต่ละช่องจะเก็บค่าคีย์ของข้อมูลและตำแหน่งที่อยู่ เพื่อใช้ในการอ้างอิงของข้อมูลจริง ลักษณะเด่นของโครงสร้างข้อมูลตารางแฮช คือ สามารถเพิ่ม ลบ ค้นหาข้อมูลได้อย่างรวดเร็ว เมื่อนำมาใช้ในการสืบค้นข้อมูลจะสามารถทำให้การสืบค้นข้อมูลเร็วขึ้น และเพิ่มประสิทธิภาพในการวินิจฉัยโรค

ขอบเขตของการวิจัย

การวิจัยครั้งนี้จะเสนอวิธีการสืบค้นข้อมูลการวินิจฉัยโรคด้วยเทคนิคการทำแฮชซิง โดยใช้โครงสร้างเอ็กเทนดิเบิล แฮชซิง ข้อมูลที่นำมาทำการวิจัยในครั้งนี้ใช้ข้อมูลการรักษาผู้ป่วยจากโรงพยาบาลรัตนบุรี อำเภอรัตนบุรี จังหวัดสุรินทร์ ซึ่งเป็นโรงพยาบาลที่เปิดทำการมานานกว่า ยี่สิบปี มีการเก็บข้อมูลประวัติผู้ป่วยและมีการเก็บข้อมูลการรักษาผู้ป่วยไว้เป็นจำนวนมาก ข้อมูลของการรักษาผู้ป่วยจะใช้ข้อมูลตั้งแต่ พ.ศ.2540 ถึง พ.ศ.2548 มีจำนวน 3,800,000 รายการ เป็นข้อมูลที่มีอยู่และถูกบันทึกด้วยระบบคอมพิวเตอร์ ในการทดลองครั้งนี้จะใช้ตัวอย่างกลุ่มข้อมูล 3 กลุ่ม ได้แก่ กลุ่มข้อมูลขนาดเล็ก มีจำนวน 150,000 รายการ กลุ่มข้อมูลขนาดกลาง มีจำนวน 250,000 รายการ และกลุ่มข้อมูลขนาดใหญ่ มีจำนวน 350,000 รายการ ซึ่งจะทำการทดลองโดยใช้ฮาร์ดแวร์เดียวกัน และใช้ฐานข้อมูลเดียวกัน เป็นการทดลองในระบบปิด ไม่มีการเพิ่มข้อมูลลงไปในฐานข้อมูล การสืบค้นข้อมูลจะใช้อาการของผู้ป่วยเพื่อให้สามารถสืบค้นข้อมูลอาการโรคต่าง ๆ โดยการเปรียบเทียบนี้จะเปรียบเทียบจากลักษณะอาการที่คล้ายคลึงกัน

แล้วนำผลที่ได้มาทำการวินิจฉัยโรค และทำการเปรียบเทียบผลเป็นเวลาของการสืบค้นข้อมูล ระหว่างการวินิจฉัยโรคด้วยเทคนิคการทำแฮชชิง โดยใช้โครงสร้างแฮชเทเบิล แฮชชิง และการสืบค้นข้อมูลการวินิจฉัยโรคด้วยภาษาสอบถามเชิงโครงสร้าง

นิยามศัพท์เฉพาะ

นิยามคำศัพท์เฉพาะที่ใช้ในการศึกษาวิจัยครั้งนี้ มีดังต่อไปนี้

คีย์ข้อมูล (Key) เป็นเหมือนรหัสประจำตัวของข้อมูลแต่ละชุดที่จะเก็บลงในหน่วยความจำ โดยปกติคีย์ข้อมูลจะมีค่าเป็นจำนวนเต็ม เพื่อที่จะสามารถนำไปใช้คำนวณหาตำแหน่งของหน่วยความจำที่เหมาะสมได้ ซึ่งถ้าข้อมูลมีค่าเป็นเลขจำนวนเต็ม ก็จะสามารถนำข้อมูลนั้นมาใช้เป็นคีย์ข้อมูลได้เลย แต่ถ้าข้อมูลเป็นตัวอักษร (Character) หรือเป็นข้อความ (Text) ก็จะต้องมีการแปลงข้อมูลให้เป็นคีย์ข้อมูลที่มีค่าเป็นเลขจำนวนเต็มเสียก่อน

ตำแหน่งที่อยู่ (Address) คือ ตำแหน่งของช่องข้อมูลในตารางแฮช ที่เราสามารถเก็บข้อมูลนั้น ๆ ลงไปได้ ซึ่งส่วนใหญ่ก็จะหมายถึง Index ของอาร์เรย์ที่จะเก็บข้อมูลลงไป

แฮชชิง (Hashing) คือ การแปลงค่าคีย์ข้อมูลให้กลายเป็นตำแหน่งที่อยู่เพื่อที่จะสามารถเก็บค่าคีย์ข้อมูลนั้น ๆ ลงในตารางแฮชได้และสามารถที่จะค้นหาคีย์ข้อมูลนั้นได้ในภายหลังโดยวิธีที่นำมาใช้ในการแปลงค่าก็คือการใช้แฮชฟังก์ชัน

ตารางแฮช (Hash Table) คือ ตารางที่ใช้ในการเก็บคีย์ข้อมูลและตำแหน่งที่อยู่ (Address) ซึ่งภายในตารางจะมีช่องที่ใช้เก็บคีย์ข้อมูลแต่ละตัวเอาไว้ โดยแต่ละช่องเก็บข้อมูลจะมีตำแหน่งที่อยู่กำกับเอาไว้ เพื่อใช้ในการอ้างอิงตำแหน่งที่บันทึกข้อมูลจริง

แฮชฟังก์ชัน (Hash Function) คือ เทคนิควิธี หรือฟังก์ชันทางคณิตศาสตร์ที่นำมาใช้หาตำแหน่งที่อยู่ให้กับข้อมูลในตารางแฮช

แฮชเทเบิล แฮชชิง (Extendible Hashing) เป็นวิธีการทำแฮชชิงรูปแบบหนึ่ง ซึ่งเหมาะสำหรับกรณีที่มีปริมาณข้อมูลเป็นจำนวนมาก สามารถปรับปรุงตารางแฮชได้อย่างยืดหยุ่น (Dynamic) ทำให้ใช้เวลาในการรักษาสถานะในตารางแฮชด้วยเวลาน้อยกว่าเทคนิคแฮชชิงแบบคงที่ (Static) นอกจากนี้ยังสามารถเอื้อประโยชน์ในการจัดการตารางแฮชที่มีขนาดใหญ่เกินกว่าที่จะเก็บในหน่วยความจำหลัก (Main Memory) ที่มีขนาดจำกัดได้หมดในระหว่างที่ประมวลผล

ประโยชน์ที่จะได้รับ

1. มีโปรแกรมเพิ่มประสิทธิภาพการสืบค้นข้อมูลบนฐานข้อมูลขนาดใหญ่ที่มีอยู่เดิม
2. มีระบบที่ช่วยสืบค้นข้อมูลการวิจัยโรคได้อย่างรวดเร็วและมีประสิทธิภาพ
3. สามารถนำเทคนิคที่นำเสนอนี้ไปประยุกต์ใช้กับงานด้านอื่น ๆ ที่มีความต้องการทางด้านการสืบค้นข้อมูลอย่างรวดเร็วและมีปริมาณข้อมูลจำนวนมาก

