ชื่อวิทยานิพนธ์ ดัชนีบิตแมปแบบกระจายสำหรับการสอบถามแบบค่าเท่ากันและแบบความเป็นสมาชิกจากคลังข้อมูล

ผู้เขียน นางสาวจรินทร์ หมั่นฝึกพันธ์

สาขาวิชา วิทยาการคอมพิวเตอร์

ปีการศึกษา 2548

บทคัดย่อ

ในระบบคลังข้อมูล การทำดัชนีแบบบิตแมปมีความสำคัญต่อการเพิ่มประสิทธิภาพโดยไม่ต้องเสียค่าใช้จ่ายในการเพิ่มอุปกรณ์ฮาร์ดแวร์ เพื่อดึงข้อมูลหรือสารสนเทศขึ้นมาใช้งาน ซึ่งส่วนมากมักมีลักษณะการสอบถามข้อมูลแบบซับซ้อนและแบบทันทีทันใด นอกจากนี้ดัชนีบิตแมปยังมีคุณสมบัติในการดำเนินการตรรกะระดับบิตระหว่างบิตแมปเวกเตอร์ก่อนดึงข้อมูลจริง ทำให้ใช้เวลาน้อยลงในการค้นหาข้อมูล

วิทยานิพนธ์นี้ เป็นการนำเสนอเทคนิคดัชนีบิตแมปแบบใหม่ที่เรียกว่า ดัชนีบิตแมปแบบกระจาย ซึ่งใช้พื้นที่น้อยลงกว่าดัชนีบิตแมปแบบเดิมที่เคยมีอยู่ แต่ก็ยังคงรักษาเวลาในการดึงข้อมูล ดัชนีบิตแมปแบบกระจาย จะมีการใช้ 2 บิตแมปเวกเตอร์ในการแทนค่าแต่ละค่า โดยแต่ละบิตแมปเวกเตอร์จะแทนค่าได้หลายค่า จากผลการทดลองเปรียบเทียบกับดัชนีบิตแมปแบบพื้นฐาน แบบช่วง และแบบเข้ารหัส สามารถยืนยันได้ว่า ดัชนีบิตแมปแบบกระจายดีกว่าดัชนีบิตแมปแบบเดิมที่เคยมีมา ในรูปแบบของความสัมพันธ์ระหว่างพื้นที่และเวลา ในกรณีที่เป็นการสอบถามข้อมูลแบบค่าเท่ากันและแบบความเป็นสมาชิก นอกจากนี้ดัชนีบิตแมปแบบกระจายยังสามารถใช้ได้กับแอทริบิวต์ที่มีคาร์ดินอลิตี้สูงได้ด้วย

Thesis Title	Scatter Bitmap Indexing for Equality and Membership Query on Data Warehouse
Author	Miss Jarin Manfuekphan
Major Program	Computer Science
Academic Year	2005

ABSTRACT

In any data warehouse environment, indexing techniques based on bitmap representations are useful for answering complex and ad hoc queries efficiently without adding additional hardware. They significantly improve query processing time by utilizing low-cost Boolean operations on the index level before going to the primary data source.

The objective of this thesis is to propose a new bitmap index called Scatter Bitmap Index that uses less space than existing bitmap indices while maintaining processing time. Scatter Bitmap Index achieves this by representing each attribute value using only two bitmap vectors, but each bitmap vector represents many attribute values. The experimental results compared against Simple, Interval and Encoded Bitmap Index confirm that the performance of Scatter Bitmap Index is much better than existing techniques in terms of space-time trade-off for equality and membership queries.