

ชื่อวิทยานิพนธ์	วิธีการทำดัชนีบิตแมปแบบคู่กันสำหรับการสอบถามแบบค่าเท่ากัน จากคลังข้อมูล
ผู้เขียน	นางสาวนิวรรณ วัฒนกิจรุ่งโรจน์
สาขาวิชา	วิทยาการคอมพิวเตอร์
ปีการศึกษา	2550

บทคัดย่อ

คลังข้อมูลเป็นแหล่งที่เก็บข้อมูลที่มีขนาดใหญ่มหาศาล เพื่อใช้ในการสนับสนุนการตัดสินใจ ดังนั้นประสิทธิภาพในการสอบถามจึงเป็นสิ่งสำคัญ การทำดัชนีบิตแมปเป็นวิธีการที่จะช่วยให้การสอบถามมีประสิทธิภาพยิ่งขึ้น โดยไม่จำเป็นต้องเสียค่าใช้จ่ายในการเพิ่มอุปกรณ์ใด ๆ เนื่องจากการทำดัชนีบิตแมปใช้การดำเนินการระดับบิตอันเป็นการสนับสนุนการทำงานของฮาร์ดแวร์ก่อนเข้าถึงข้อมูลจริง ทำให้การสอบถามได้คำตอบในเวลาอันรวดเร็ว

วิทยานิพนธ์นี้ เป็นการนำเสนอวิธีในการทำดัชนีบิตแมปแบบใหม่ที่ได้คิดค้นขึ้น มีชื่อว่า ดัชนีบิตแมปแบบคู่กัน (Dual Bitmap Index) ซึ่งมีประสิทธิภาพในเรื่องของพื้นที่ที่ใช้ในการสร้างดัชนีมากกว่าดัชนีบิตแมปที่เคยมีมา แต่ยังคงมีประสิทธิภาพในเรื่องของเวลาในการตอบคำถามแบบค่าเท่ากันอยู่ ดัชนีบิตแมปแบบคู่กันใช้หลักการที่ว่า แต่ละค่าของแอทริบิวต์ถูกแทนด้วย 2 บิตแมปเวกเตอร์ และทำให้ทุกบิตแมปเวกเตอร์ถูกใช้อย่างคุ้มค่าที่สุด จากผลวิเคราะห์และการทดลองเปรียบเทียบระหว่างดัชนีบิตแมปแบบคู่กันที่คิดค้นขึ้นและดัชนีบิตแมปที่เคยมีมา ได้แก่ ดัชนีบิตแมปแบบพื้นฐาน ดัชนีบิตแมปแบบช่วง ดัชนีบิตแมปแบบกระจาย และดัชนีบิตแมปแบบเข้ารหัส พบว่า ดัชนีบิตแมปแบบคู่กันมีประสิทธิภาพในแง่ Space-Time Trade-off (การแลกเปลี่ยนระหว่างประสิทธิภาพของพื้นที่กับเวลา) สำหรับการสอบถามแบบค่าเท่ากันมากกว่าดัชนีบิตแมปที่เคยมีมา

Thesis Title	Dual Bitmap Indexing Technique for Equality Queries on Data Warehouse
Author	Miss Niwan Wattanakitrungrroj
Major Program	Computer Science
Academic Year	2007

ABSTRACT

Indexing techniques based on bitmap representations are known to improve query response time in a data warehouse environment. Specifically, bitmap indexing is a good method for improving complex and ad hoc queries without adding additional hardware. This method improves query processing using large numbers of fast Boolean operations and multiple index scans. By performing simple predicate conditions on the index level before going to the primary data source, the average query response time is improved.

In this thesis, Dual Bitmap Index is introduced, which uses less space than existing bitmap indices while maintaining the same improvements in query processing speed. The Dual Bitmap Index represents each attribute value using only two bitmap vectors, with each bitmap vector representing many attribute values. This research shows a comparative study of five bitmap indexing techniques, and shows that the Dual Bitmap Index is more efficient than the existing techniques (i.e., Simple Bitmap Index, Interval Bitmap Index, Scatter Bitmap Index and Encoded Bitmap Index) for equality from a space-time trade-off perspective.