

ชื่อเรื่องการค้นคว้าแบบอิสระ การจัดกลุ่มข้อมูลด้วยอัลกอริทึม เค-มีนส์ แบบ

ไปป์ไลน์

ผู้เขียน

นายเชิงชาย เนรมิตคภพงค์

ปริญญา

วิทยาศาสตรมหาบัณฑิต(วิทยาการคอมพิวเตอร์)

อาจารย์ที่ปรึกษาการค้นคว้าแบบอิสระ

อาจารย์ ดร.สรพรพรรณ กันตะบุตร

บทคัดย่อ

เค-มีนส์ อัลกอริทึม เป็นวิธีที่นิยมใช้ในการจัดกลุ่มข้อมูลในปัจจุบัน ถึงอย่างไรก็ตามเมื่อข้อมูลเข้ามีขนาดใหญ่อัลกอริทึมจะทำงานช้ามาก เพราะเค-มีนส์อัลกอริทึมมีความซับซ้อนของเวลาเป็น $O(KRN)$ เมื่อ K เป็นจำนวนของกลุ่มที่ต้องการ R เป็นจำนวนของการวนทำงาน และ N เป็นจำนวนของจุด เป็นที่รู้กันว่า เค-มีนส์ มีลักษณะใกล้เคียงกับกราฟเส้นตรง(linear rate) เมื่อเข้าใกล้กันเวจเงินซ(convergence) ถ้า R เป็นสัดส่วนกับ N เวลาที่ใช้ในการทำงานของ เค-มีนส์จะเท่ากับ $O(N^2)$ ซึ่ง N ที่มีขนาดใหญ่มาก จะไม่สามารถทำงานได้จริงในขอบเขตของเวลาที่ต้องการ ในผลงานค้นคว้านี้ แสดงให้เห็นว่าระบบคู่ขนานสามารถแก้ปัญหาได้ โดยพูดถึงประโยชน์ของการจัดกลุ่มข้อมูลของเว็คเตอร์ที่ราคาถูกลงและมีราคาฮาร์ดดิสก์ถูก การทดลองของเราแสดงให้เห็นว่าอัลกอริทึมแบบคู่ขนาน มีความเร็วกว่าอัลกอริทึมที่มีอยู่

Independent Study Title Pipelined k-Means Clustering Algorithm

Author Mr. Chengchai Naramittakapong

Degree Master of Science (Computer Science)

Independent Study Advisory Lecturer Dr. Sanpawat Kantabutra

ABSTRACT

K-means algorithm is one of the most popular clustering algorithms in use today. However, when the input size is large, the algorithm runs very slow. This is because it has time complexity of $O(KRN)$ where K is the number of desired clusters, R is the number of iterations, and N is the number of points. It is known that K-means approaches a linear rate when it is close to convergence. If R is proportional to N , the running time of K-means is $O(N^2)$ which is impractical for a large N . In this paper, we offer a parallel solution to the K-means problem by taking advantage of a cluster of inexpensive workstations and a relatively low price of hard disk. Our experiments show that this parallel algorithm achieves a much faster speed than the existing algorithms.

ลิขสิทธิ์มหาวิทยาลัยเชียงใหม่
Copyright © by Chiang Mai University
All rights reserved