

Peningkatan kinerja komputasi dalam pendeteksian diabetic retinopathy melalui data ekspresi gen menggunakan metode parallel two-phase biclustering = Improvement of computational performance in detection of diabetic retinopathy through gene expression data using parallel two phase method biclustering

Gianinna Ardanawati, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20446497&lokasi=lokal>

Abstrak

Diabetic Retinopathy merupakan penyakit yang disebabkan oleh komplikasi mikrovaskuler jangka panjang dari pasien Diabetes Melitus. Salah satu bentuk data biologi molekuler yang tengah berkembang pesat saat ini adalah data ekspresi gen pada microarray. Analisis data ekspresi gen dapat dilakukan dengan berbagai cara termasuk pengelompokan data menggunakan algoritma clustering ataupun biclustering. Salah satu metode untuk menganalisis data ekspresi gen adalah metode Two-Phase Biclustering. Untuk data yang berukuran besar, metode tersebut membutuhkan waktu komputasi yang lama. Penerapan komputasi paralel diperlukan dalam metode Two-Phase Biclustering ini agar waktu komputasi yang dibutuhkan lebih sedikit dan memiliki kinerja yang baik.

Dalam tesis ini akan dibahas implementasi paralel pada metode Two-Phase Biclustering, dimana fase pertama menggunakan algoritma paralel K-Means dan fase kedua menggunakan algoritma biclustering Cheng-Church pada data ekspresi gen Diabetic Retinopathy. Dari hasil penelitian kami diperoleh peningkatan kinerja komputasi berupa speed-up sebesar 3,63x pada multicore paralel R dan 34x pada pemrograman manycore paralel CUDA-GPU.

.....Diabetic Retinopathy is a disease caused by long term microvascular complications on diabetes mellitus patients. Recently, the microarray gene expression data has been developing rapidly in molecular biology. There are many techniques for gene expression data analysis methods using clustering or biclustering algorithms. One of the effective method for analyzing gene expression data is Two Phase Biclustering method. However for large sized data, the method requires long computation time. Implementation of parallel computing is necessary in Two Phase Biclustering method to reduce the computation time and to obtain a good performance.

In this thesis we discuss the implementation of parallel Two Phase biclustering, where in the first phase using parallel K Means algorithm and the second phase using Cheng Church biclustering algorithm on Diabetic Retinopathy gene expression data. In parallel K Means algorithms are applied multicore based parallel R programming and manycore CUDA GPU parallel programming. Parallelization using multicore based parallel R programming achieves an average speed up of 3.63x and using manycore CUDA GPU parallel programming achieves significant performance with an average speed up of 34x.