

Penerapan komputasi paralel pada algoritma firefly-markov clustering untuk menganalisis jaringan interaksi protein = Application of parallel computing on firefly markov clustering algorithm for analyzing protein interaction network

Ida Mujtahidah, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20467559&lokasi=lokal>

Abstrak

ABSTRAK

Protein memiliki peranan penting dalam mengontrol fungsi enzim, mengatur produksi dan aktivitas di dalam sel. Untuk melakukan hal ini, protein berinteraksi dengan protein-protein lainnya, DNA dan molekul-molekul lainnya. Jaringan interaksi protein-protein ini memiliki ukuran yang sangat besar. Sehingga untuk memudahkan dalam menganalisis jaringan ini diperlukan metode clustering. Pada penelitian ini, algoritma Markov clustering digabungkan dengan algoritma firefly yang disebut sebagai algoritma firefly-Markov clustering FMCL . Pada algoritma FMCL, posisi kunang-kunang akan menggantikan nilai parameter inflate. Kemudian proses clustering akan berlangsung menggunakan algoritma Markov clustering MCL . Selanjutnya posisi kunang-kunang akan terus diperbarui dan proses clustering akan terus dilakukan sampai diperoleh hasil clustering terbaik. Komputasi paralel pada algoritma FMCL menggunakan OpenMP. Setiap thread menjalankan proses pada Markov clustering menggunakan parameter inflate yang ditentukan oleh posisi kunang-kunangnya. Hasil yang diperoleh pada data jaringan interaksi protein HIV ada 4 cluster. Dari 4 cluster ini terdapat satu cluster besar yang saling terhubung dengan 6 pusat cluster lainnya yaitu NEF, GAG, GAG-POL, VPR, VIF dan VPU. Pada 3 cluster lainnya yang menjadi pusat cluster adalah TAT, REV dan ENV. Sedangkan dengan menggunakan data jaringan interaksi protein pada Human Herpesvirus tipe 4 HHV-4 diperoleh 14 cluster. Protein yang menjadi pusat cluster adalah EBNA-LP, BKRF1, BPLF1, LMP1, SUMO2, BBLF2-BBLF3, EBNA3B, BRLF1, BGLF4, BYRF1. Selain itu, juga dapat dilihat bahwa ada beberapa cluster yang hanya merupakan interaksi antara dua protein yaitu BBRF1 dengan NFkB2, EBNA3A dengan CHEK2, LMP2A dengan ITCH, dan EBNA3C dengan EP300. Speed up algoritma FMCL yang dijalankan menggunakan OpenMP pada data HIV dan HHV-4 adalah 4.73x dan 3.21x lebih cepat dibandingkan dengan algoritma FMCL yang dijalankan secara sekuensial.

<hr />

ABSTRACT

Protein plays an important role in controlling enzyme function, regulating production and activity in cell. To do this function, proteins will interact with other protein, DNA and other molecules. Protein interaction network have a very large size. Then to simplify analyzing this network is required clustering method. In this study, Markov clustering algorithm combined with a firefly algorithm called firefly Markov clustering algorithm FMCL . In FMCL algorithm, firefly position will be replace the value of inflate parameter. Then clustering process will take place using Markov clustering algorithm MCL . Futhermore, the firefly position will be updated and clustering process will be continue until its get the best clustering. Parallel computing on FMCL algorithm using OpenMP. Each thread will run the process on Markov clustering by using inflate parameter specified by the position of firefly. The clustering result from protein interaction network on HIV is 4 clusters. From this cluster, there is a large cluster connected with 6 other cluster centers, they are NEF,

GAG, GAG POL, VPR, VIF and VPU. In the 3 other clusters that become the center of the cluster are TAT, REV and ENV. While by using protein interaction network on HHV 4 obtained 14 clusters. The proteins that become the center of the clusters are EBNA LP, BKRF1, BPLF1, LMP1, SUMO2, BBLF2 BBLF3, EBNA3B, BRLF1, BGLF4 and BYRF1. In addition, it can be seen that there are several clusters that are just interaction between two proteins, BBRF1 with NFKB2, EBNA3A with CHECK2, LMP2A with ITCH and EBNA3C with EP300. The speed up of FMCL algorithm by using OpenMP HIV and HHV 4 data is 4.73x and 3.21x faster than the sequentially executed.