

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

M. Syamsuddin Wisnubroto, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20467561&lokasi=lokal>

Abstrak

ABSTRAK

Protein memiliki peranan yang sangat penting dalam kehidupan. Setiap protein berinteraksi dengan protein-protein lain, DNA, dan molekul-molekul lainnya, sehingga terbentuklah jaringan interaksi protein yang berukuran sangat besar. Untuk memudahkan dalam menganalisisnya, diperlukan metode clustering. Algoritma Soft Regularized Markov Clustering (SR-MCL) merupakan pengembangan metode clustering yang mengurangi kelemahan dari Regularized Markov Clustering dan Markov Clustering. Namun, SR-MCL masih memiliki kelemahan yaitu parameter inflasi yang selalu dimasukkan secara manual oleh peneliti. Penelitian ini, SR-MCL digabung dengan Algoritma Firefly yang selanjutnya disebut Firefly Soft Regularized Markov Clustering, dimana posisi setiap firefly menggantikan parameter inflasi. Posisi firefly akan terus diperbaharui dan proses clustering akan terus dilakukan sampai memperoleh global chaos kurang dari threshold. FSR-MCL akan diterapkan secara paralel menggunakan OpenMP, yaitu setiap thread menjalankan SR-MCL dengan posisi setiap firefly yang berbeda. Proses clustering data HIV-1 diperoleh sembilan protein sebagai pusat cluster yang sangat berpengaruh dalam pembentukan dan penyebaran virus, yaitu TAT, REV, ENV, GAG, POL, VPU, VPR, NEF, dan VIF, serta didapat parameter inflasi terbaiknya 8,0 dengan speed up 4,66 kali. Proses clustering data SC5314 diperoleh enam protein sebagai pusat cluster yang merupakan protein penting dalam penyebarannya, yaitu HSP90, CBK1, MED8, NOP1, CEK1, dan CDC4, serta didapat parameter inflasi terbaiknya 5,5 dengan speed up 3,01 kali.

<hr>

ABSTRACT

Protein has a very important role in life. Each protein interacts with other proteins, DNA, and other molecules, resulting in a very large protein-protein interaction. To make it easier to analyze it, clustering method is needed. Soft Regularized Markov Clustering (SRMCL) algorithm is a development of clustering method that reduces the weakness of Regularized Markov Clustering and Markov Clustering. However, SR-MCL still has a weakness that is the parameter of inflation that is always entered manually by researchers. This study, SR-MCL combined with Firefly Algorithm, hereinafter called Firefly Soft Regularized Markov Clustering, where the position of each firefly replace the parameters of inflation. The firefly position will continue to be updated and the clustering process will continue until the global chaos is less than the threshold. FSR-MCL will be applied

in parallel using OpenMP, ie each thread runs SR-MCL with the position of each different firefly. The process of clustering the HIV-1 data obtained by nine proteins as the center of the cluster is very influential in the formation and spread of the virus, namely TAT, REV, ENV, GAG, POL, VPU, VPR, NEF, and VIF, and got the best inflation parameter 8.0 with speed up 4.66 times. SC5314 data clustering process obtained six proteins as the center of the cluster which is an important protein in its spreading, namely HSP90, CBK1, MED8, NOP1, CEK1, and CDC4, and got the best inflation parameter 5.5 with speed up 3.01 times.