

Analisis Jaringan Interaksi Protein Coronavirus Menggunakan Multilevel Markov Clustering dengan Maximum Matching = Analyzing of Protein Interaction Network of Coronavirus Using Multilevel Markov Clustering with Maximum Matching

Aminatuzuhriah Rizki, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=20528135&lokasi=lokal>

Abstrak

Protein mempunyai peranan penting sebagai struktur dalam tubuh virus. Protein tersebut saling berinteraksi membentuk jaringan interaksi protein-protein. Jaringan interaksi protein coronavirus mempunyai ribuan protein dan interaksi yang membentuk jaringan interaksi protein yang besar, sehingga diperlukan metode pengelompokan untuk menyederhanakan analisis jaringan tersebut. Pada penelitian ini, algoritma Markov Clustering (MCL) yang dikombinasikan dengan maximum matching dijalankan pada sebuah kerangka kerja multilevel untuk mengelompokkan jaringan interaksi protein coronavirus. Algoritma maximum matching digunakan untuk memperoleh hasil matching yang optimal di fase graph coarsening pada setiap level dan algoritma MCL digunakan untuk mengelompokkan jaringan interaksi protein tersebut. Hasil pengelompokan yang paling optimal diperoleh pada parameter penggelembungan dengan level-1. Hasil tersebut dilihat berdasarkan nilai average N-cut yaitu 0,8729. Semakin kecil nilai average N-cut, maka kualitas hasil pengelompokannya semakin baik. Hasil pengelompokan terbaik pada jaringan interaksi protein coronavirus menggunakan ML-MCL dengan maximum matching dihasilkan 21 klaster terpisah dengan 4911 interaksi. Protein yang mempunyai peran penting dapat dilihat dari pusat klaster pada hasil pengelompokan terbaik dan protein-protein tersebut yaitu ZW10, ZYG11A, ZNF771, ZZEF1, ZNF451, ZNF668, YKT6, WDR11, ZNF318, ZYG11B, ZNF428, ZPR1, ZFR, TK2, ZNF746, UBR5, ZNF609, ZZZ3, ZBTB16, XPNPEP3, dan USP7. Waktu yang dibutuhkan pada hasil pengelompokan terbaik yaitu 376,6494 detik. Pada simulasi ini, tidak terdapat protein yang hilang pada seluruh hasil pengelompokan.

.....Protein has an essential role as a structure in the body of the virus. These proteins interact to form a network of protein-protein interactions. The protein interaction network of coronavirus has thousands of proteins and interactions that form a large protein interaction network, so that a clustering method is needed to simplify the analysis of the network. In this study, the Markov Clustering (MCL) algorithm was built combined with maximum matching in a multilevel framework to cluster the protein interaction network of coronavirus. The maximum matching algorithm is used to obtain the optimal matching result in the graph coarsening phase at each level, and the MCL algorithm is used to cluster the protein interaction network. The most optimal grouping results were obtained at inflation parameter with level-1. These results are seen based on the average N-cut value, which is 0.8729. The smaller the average N-cut value, the better the quality of the clustering results. The best clustering result of the protein interaction network of coronavirus using ML-MCL with maximum matching resulted in 21 separate clusters with 4911 interactions. Proteins that have an important role can be seen from the center of the cluster of the best clustering result, and these proteins are ZW10, ZYG11A, ZNF771, ZZEF1, ZNF451, ZNF668, YKT6, WDR11, ZNF318, ZYG11B, ZNF428, ZPR1, ZFR, TK2, ZNF746, UBR5, ZNF609, ZZZ3, ZBTB16, XPNPEP3, and USP7. The time required for the best clustering results is 376.6494 seconds. In this simulation, there were no missing proteins in all clustering results.