Step-by-step guide to execute the multiNG method

Introduction

The primary input files consist of $L \, n \times n$ matrices. Each of them provides the strength of pairwise connections in an $n \times n$ weighted network. Such matrices can be generated starting from a pertinent data set, with the help of a tool to be chosen by the user. For the purpose of carrying phylogenetic analysis, the n network nodes correspond to organisms to be placed in the phylogenetic tree. Each matrix corresponds to an identity matrix, obtained by evaluating pairwise protein alignment among all n organisms. The used data consists of aminoacid sequences of L proteins for all selected n organisms. The Supplementary Material contains further details of the use alignment procedure. To run the multiNG procedure, follow the sequence of steps:

- 1. Run the program redecritica.for in order to localize suitable values of the parameter σ_c at which the phylogenetic analysis will be carried out. The parameter σ_c is a threshold value to select edges that will be placed in the unweighted network. Repeat this procedure for each of the L proteins. Input data to run the code should be informed in file redecritica.dat
- 2. After selecting a suitable value of \sigma_cr, which usually correspond to values of \sigma just before large peaks that can be observed in the output of redecritica.for, run the program matrix_adj.f90 to generate the L adjacency matrices of each network. Note that you may chose different values of \sigma_cr for different matrices corresponding to each protein. Input data to run the code should be informed in file matrix_adj.dat
- 3. Assemble the $(Ln)\times(Ln)$ adjacency matrix of the multiplex network with L layers using the program multiplex.for. Input data to run the code should be informed in file multiplex.dat.
- 4. Run the program caminhos.for to obtain the $(Ln)\times(Ln)$ multiplex neighborhood matrix, the elements of which correspond to the shortest path between two nodes. Input data to run the code should be informed in file caminhos.dat.
- 5. Use the resulting matrix as input to main program MultiNG.for, which leads to the final results of the procedure, which are available in output files. Input data to run the code should be informed in file MultiNG.dat.
- 6. All codes are written in Fortran (either Fortran 77 or Fortran 90). The executable file was obtained using the gfortran compiler.