Computationally efficient network inference using information theory: fMIDER

Alejandro F Villaverde1, Federico Morán2, Julio R Banga1

1 Bioprocess Engineering Group, IIM (CSIC), Vigo, Spain, 2 Dept of Biochemistry and Molecular Biology, Univ. Complutense Madrid, Spain

Summary:
- Network inference is a challenging problem in many areas of science, particularly in biological systems
- Detection of interactions among nodes commonly carried out with correlation / mutual information measures
- MIDER (Mutual Information Distance & Entropy Reduction) is a general purpose reverse engineering tool
- fMIDER is a FORTRAN implementation of MIDER that is more efficient (faster), open source, can be run in parallel environments, and does not require any commercial software

1. Mutual Information Distance:
1.1. Adaptive estimation of I(X,Y) from data
1.2. Calculate distances between variables for several time delays
1.3. Apply MDS to the distance matrix → 2D map of variables

2. Entropy Reduction:
• Goal: determine if the variation in a variable Y can be explained by the variations in other variables. If (Y,X) are independent, H(Y|X) = H(Y); if not, H(Y|X) < H(Y)
• By iterating through cycles of adding a variable that reduces H(Y|X), the entropy reduction step yields the set of variables that control the variation in Y
• RESULT: links with width proportional to interaction strength (% entropy reduction)
• Directions are assigned using Transfer Entropy

fMIDER

Note that MIDER can be run sequentially or in parallel. Parallelization is automatic, i.e. transparent to the user, who only needs to choose the number of threads

CONCLUSIONS

- detects indirect interactions & assigns directions
- provides a visual representation of the network
- does not require previous knowledge/expertise
- publicly available as a Matlab toolbox (GPL)
- FORTRAN code (f90), faster than Matlab
- No commercial software required
- Minimum speedups = ×5 can be up to ×30 for some problems

REFERENCE