Small-world bias of correlation matrix in natural complex systems

J. Hlinka[1] M. Vejmelka[1] D. Hartman[1] J. Tintěra[2] M.Paluš[1]

[1] Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic
[2] Institute for Clinical and Experimental Medicine, Videnska 1958/9, 140 21 Prague, Czech Republic

This study was supported by the Czech Science Foundation project No. P103/11/J068, No. 13-23940S and by the media agency PHD.

Introduction

Characterization of real-world complex systems such as social networks, Earth climate or brain connectivity networks increasingly involves the study of their topological structure using graph theory [Boccaletti, Bullmore]. Among global network properties, small-world [Watts] property, consisting in existence of relatively short paths together with high clustering of the network, is one of the most discussed and studied [Bassett].
To establish so-called functional connectivity (FC) network of a system, links among its regions are commonly quantified by correlation coefficient, especially when the data are close to Gaussian. Approximately Gaussian distribution has been recently shown both for brain activity data measured by functional magnetic resonance imaging (fMRI) [Hlinka, 2011] and for time series of mean monthly surface air temperatures (SAT) [Hlinka, 2013].
It has been recently shown that the functional connectivity matrices provide upwardly biased estimates of small-world characteristics, including small-world characteristics of connectivity graphs estimated from randomly connected dynamical systems [Hlinka, 2012].

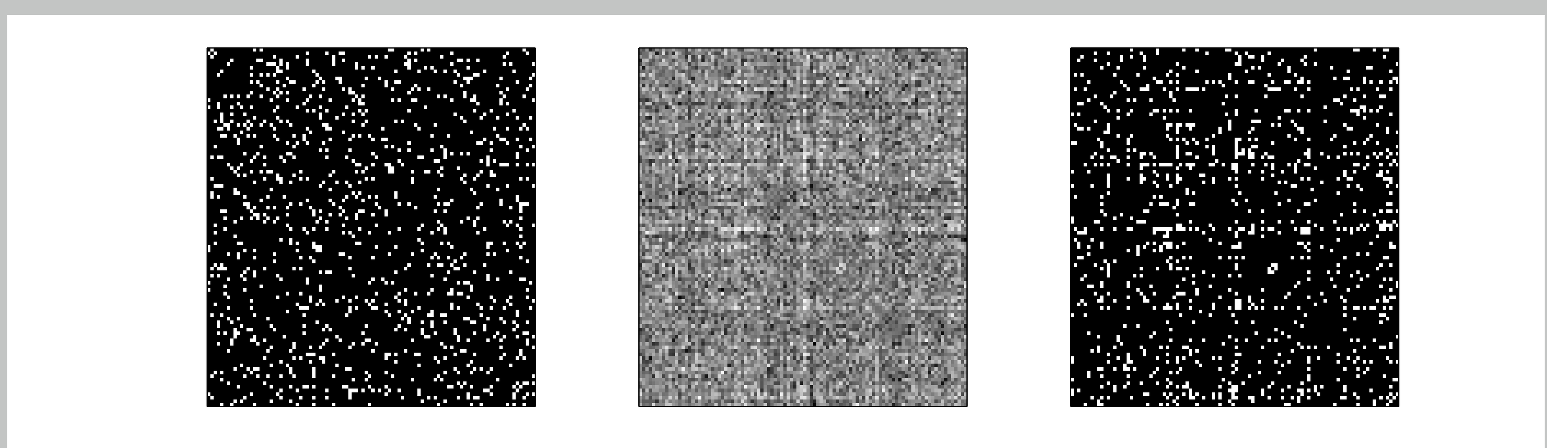

Figure 1: An example of binary functional connectivity matrix (right) generated from random structural connectivity matrix (left) by thresholding the correlation matrix of AR-model generated time series (center, light shades of gray indicate higher correlation values). Network with $N = 100$ nodes shown. Note that the functional connectivity matrix shows a specific structure although the entries of the generating structural connectivity matrix were chosen randomly. See [Hlinka, 2012] for further details.

In this work we investigate the question to what extent may this bias explain the observations of small-world property in Earth climate or brain connectivity networks constructed by functional connectivity approach.

Data: Brain (fMRI)

- $\sim$ 9 minutes, 213 time points of whole brain resting state brain activity
- 26 (12 males, 19-54 years) healthy volunteers
- 3T Siemens Magnetom Trio MRI scanner (GE-EPI, TR/TE=2500/30 ms, voxel size=3x3x3mm), a 3D high-resolution T1-weighted image was used for anatomical reference, slice-timing correction, motion correction, spatial normalization to MNI
- original data $\sim$ 20000 time series, dimensionality reduced to 90 time series by averaging over regions from the Automated Anatomical Labeling atlas
- orthogonalized wrt motion parameters, white matter and CSF signal

Data: Earth Climate (SAT)

- 60 years, 720 time points (monthly averages) of Surface Air Temperature
- for purposes of analysis split into 6 decades
- NECP/NCAR reanalysis dataset [Kistler, 2001]
- original data dense resolution (2.5 °), i.e. over 10000 time series
- dimensionality reduced by VARIMAX-rotated Principal Component Analysis to 67 time series containing altogether above 95% of the variability

Methods

- Correlation matrix of the time series was turned to an unweighted graph by keeping a fixed proportion of the strongest links (20 percent).
- Small-world property was quantified by small-world index [Humphries] $\sigma = \frac{\gamma}{\lambda} \gg 1$, where $\lambda = \frac{L}{L_{rand}} \gtrsim 1$, $\gamma = \frac{C}{C_{rand}} \gg 1$ are the relative average path length and clustering coefficient with respect to a random graph. The average path length and the clustering coefficient are defined as:

$$L = \frac{1}{N \cdot (N-1)} \cdot \sum_{i,j} d_{i,j}, \qquad C = \frac{1}{N} \sum_{i \in V} c_i, \qquad c_i = \frac{\sum_{j,\ell} a_{i,j} a_{j,\ell} a_{\ell,i}}{k_i(k_i - 1)}$$

where $a_{i,j}$ denotes the link between nodes i,j, c_i the local clustering coefficient and $d_{i,j}$ the length of shortest path among nodes i,j.

Methods II: comparison of data and randomly connected process

- Small-world indices were computed in the same way for data and a 'scrambled interaction' time series. This was modeled by fitting an vector autoregressive (VAR) process of order 1 to the BOLD time series:

$$X_t = c + AX_{t-1} + e_t, \tag{1}$$

(where c is a $N \times 1$ vector of constants, A is a $N \times N$ matrix and e_t is a $N \times 1$ vector of error terms) and subsequently randomly scrambling the entries of the autocovariance matrix A.
- To control for the effects of approximation by a VAR process, a realization of the fitted VAR model with scrambling omitted was also analyzed.

Results

- The FC for VAR-fitted/simulated data were similar to the real data FC, unlike the FC patterns observed for the 'scrambled interaction' data (Fig 2).
- **In both brain and climate data, wee have observed small-world indexes of FC robustly higher than 1, i.e. stronger characteristics of small-world than in a corresponding random graph.**
- **However, this was also tru for FC of a realization of a randomly connected vector autoregressive processes ('scrambled interaction').**
- Brain: the mean small-world index was 2.50/2.43/2.22 (real data/VAR-model data/'scrambled interaction' data), see 3(a).
- Climate: the mean small-world index was 1.52/1.49/2.59 (real data/VAR-model data/'scrambled interaction' data), see 3(b).

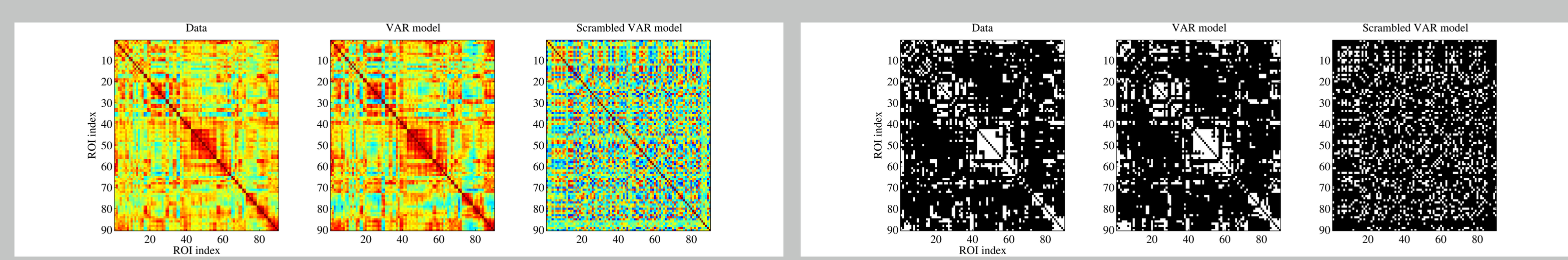

Figure 2: Example FC matrices. Left: raw. Right: thresholded to density 0.2

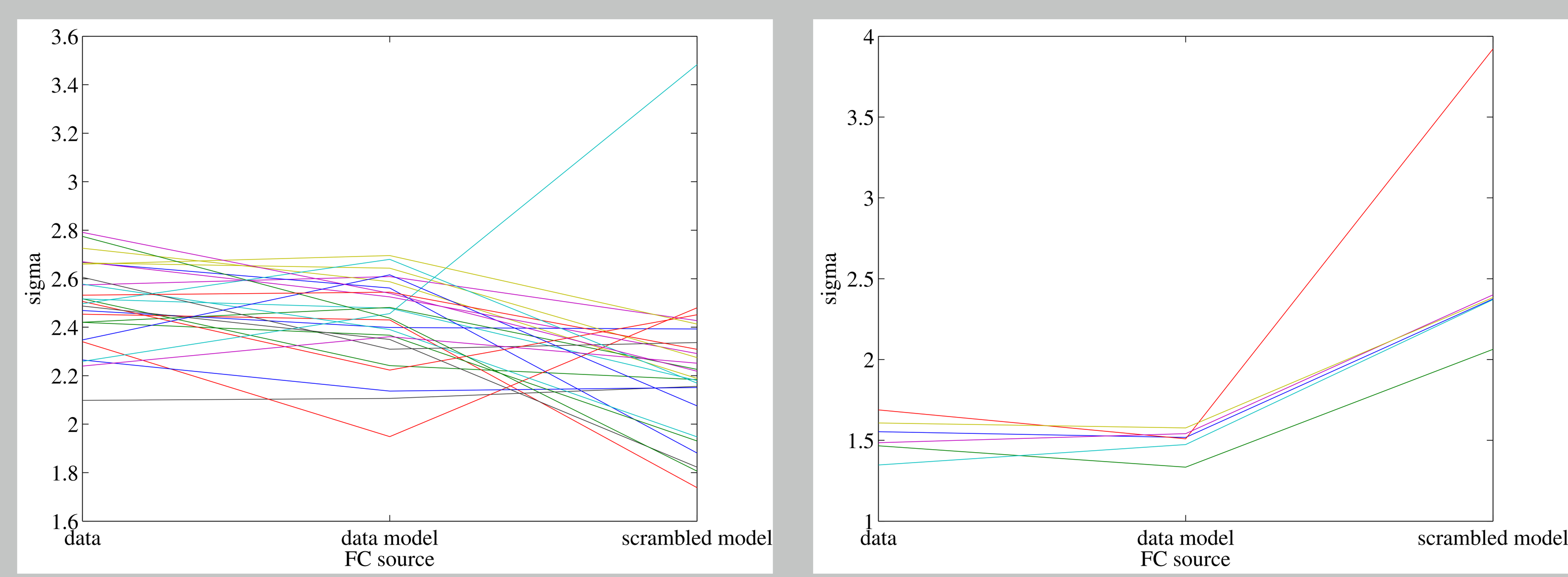

Figure 3: Small-world index for each dataset, the corresponding VAR model realization and realization of VAR model with randomized matrix A. Left: brain (fMRI). Right: climate (SAT)

Discussion and conclusions

- The small-world properties of brain fMRI FC graph are fairly well reproduced by a matching randomly connected multivariate autoregressive process, and therefore partially attributable to the standard FC construction method.
- The small-world properties of the climate SAT network constructed from PCA component time series were even lower than for a randomised VAR model. The role of the specific dimensionality reduction step in this phenomena needs to be assessed.

References

Bassett, D. S. and Bullmore, E. (2006) 'Small-world brain networks' Neuroscientist vol. 12, pp. 512-523
Boccaletti, S. et al. (2006) 'Complex networks: Structure and dynamics', Physics Reports, vol. 424, pp. 175-308
Bullmore, E. and Sporns, O. (2009) 'Complex brain networks: graph theoretical analysis of structural and functional systems' Nature Reviews Neuroscience, vol. 10, pp. 186-198
Hlinka, J. (2011) et al. 'Functional connectivity in resting-state fMRI: Is linear correlation sufficient?' NeuroImage, vol. 54, pp. 2218-2225
Hlinka, J. (2012) et al. 'Small-world topology of functional connectivity in randomly connected dynamical systems.', Chaos, vol 22. no. 3, 033107
Hlinka, J. et al. (2013) 'Non-linear dependence and teleconnections in climate data: sources, relevance, nonstationarity', Climate Dynamics
Humphries, M. D. and Gurney, K. (1998) 'Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence', PLoS One, vol. 3, e0002051.
Kistler, R. et al. (2001) 'The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation', Bulletin of the American Meteorological Society, 82, pp. 247-267
Watts, D. and Strogatz, S. (1998) 'Collective dynamics of 'small-world' networks', Nature, vol. 393, pp. 440-442

http://ndw.cs.cas.cz/ hlinka@cs.cas.cz