Graph Analysis of fMRI data

UCLA/Semel Advanced Neuroimaging Summer Program 2015

Sepideh Sadaghiani, PhD
Contents

Introduction
- What is graph analysis?
- Why use graphs in fMRI?

Decisions
- Which fMRI data?
- Which nodes?
- Which edges?
- Which modules?
- Which graph metrics?
What is a graph?

- Mathematical representation of a real-world network with pairwise relations between objects.
Why graphs?

Euler 1736: The bridge puzzle of Königsberg

Necessary condition for the walk crossing each bridge exactly once: Zero or two nodes with odd degree.
Real-world networks

Protein-protein “interactome”

Disease gene network
Goh et al. PNAS 2007

Social network

Brain connectome
Topological Space
Bullmore & Bassett 2010

Physical Space
Graph Theory

- Real-life networks are complex

- Graph theory allows mathematical study of complex networks

- Describe properties of a complex system:
  Quantify topological characteristics of its graph representation

Bullmore & Sporns, 2012
Why use graphs analyses in fMRI?

Quantification of global properties of spatio-temporal network organization

- Early motivation:
  A testable theory of consciousness (Edelman & Tononi 2000)
  Based on global network integration (information theory)

- Structural connectivity ↔ functional connectivity
- Comparisons across individuals (e.g. in disorders)
- Comparison across mental and functional states
Graph construction in fMRI - overview

Wang et al., 2010

- Choice of nodes
- Time series extraction
- Pairwise connectivity (e.g., Pearson’s correlations)
- Thresholding (& optional binarization)
- Adjacency matrix
  - Rows & columns: nodes
  - Entries: edges

Wang et al., 2010
Which datasets?

MENTAL STATE, PREPROCESSING
fMRI Datasets

Connections most commonly derived from resting state, but task data possible in principle

Session length most commonly 5-10min
- Long enough for multiple cycles of infraslow (<0.1Hz) frequencies
- Short enough to minimize mental state change
- Shorter term time-varying dynamics (e.g. sliding window)

Preprocessing: same considerations as any fMRI connectivity study:
- What motion correction?
- Slice time correction?
- Physiological nuisance measures?
- Compartment signal regression (GM, WM, CSF)?
Which nodes?

Anatomical vs. Functional
Atlas vs. Data-Driven
Anatomical atlases

Nodes:
- Internally coherent / homogeneous (connectivity)
- Externally independent

Anatomical atlases
- Automated Anatomical Labeling (AAL) template
- Eickhoff-Zilles (Cytoarchitectonic)
- FreeSurfer (Gyral. Individual surface-based possible)
- Harvard-Oxford
- Talairach & Tournoux

- 😊 Comparability (across subjects and modalities)
- 😞 Highly variable node size. Not functionally coherent.
Functional atlases

- Craddock (local homogeneity of connectivity)
- Power (resting state seed-based & task)
- Stanford Atlas FIND lab (ICA-based)
- ☺ Comparability across subjects.
- ☹ Functionally coherent (☺ but suboptimal for individuals)

Functional subject-specific parcellations

- ICA
- (Seed-based)
- Connectivity homogeneity: Craddock
- ☺ Functionally coherent
- ☹ Time-intensive

List of atlases:  https://en.wikibooks.org/wiki/SPM/Atlases
Which edges?

CONNECTIVITY AND THRESHOLDING
Edges in fMRI

Based on magnitude of temporal covariation
- Pearson’s cross-correlations (by far most common)
- Partial correlations
- Mutual information

→ symmetric adjacency matrices (undirected graphs)

Directionality problematic in fMRI
(but measures of effective connectivity possible)
Other Data Modalities

Structural (e.g. DTI, histological tracing)
- Nodes: cf. fMRI
- Edges: e.g. number of reconstructed fibers

EEG / MEG
- Nodes: sensors or reconstructed sources
- Edges: Correlation in oscillation amplitudes
  Oscillation phase synchrony (coherence of phase locking)
Thresholding

Most metrics require **sparse** graphs

Threshold to remove weak connections

Use proportional thresholds (vs. absolute thresholds)

Use broad range of proportions
Which Modules?

COMMUNITY DETECTION
Modules

Communities of densely interconnected nodes

Optimization algorithms

Wang et al., 2010
Community Detection Algorithms

**Modularity-base algorithms**
Maximize number of within-community edges (compared to random network)
- Newman’s Modularity (Newman, 2006)
- Louvain method (Blondel et al. 2008)

**Infomap algorithm** (Rosvall and Bergstrom, 2008)
Minimize information theoretic descriptions of random walks on the graph

Review: Fortunato, 2010
Which graph metrics?

NODAL AND GLOBAL
Nodal Measures

Degree
Number of edges connected to a node

Nodal Clustering Coefficient (basis for measure of global segregation)
Fraction of all possible edges realized among a node’s neighbors
= Fraction of all possible triangles around a node

Shortest Path Length (basis for measure of global integration)
Number of edges on shortest geodesic path between two nodes

Sporns, 2011
Nodal Measures

Measures of centrality:

**Closeness Centrality**
Inverse of the node’s average Shortest Path Length

**Betweenness Centrality**
Fraction of all shortest paths passing through the node

**Participation Coefficient**
Diversity of intermodular connections

**Within-Module Degree (z-score)**
Degree \(_{\text{intra\ module}}\) z-scored within the node’s module

“Provincial hubs”: high within-module degree & low participation coefficient
“Connector hubs”: high participation coefficient
“Rich club”: densely interconnected connector hubs

Bullmore & Sporns, 2012
## Global Measures

### Measures of Integration

<table>
<thead>
<tr>
<th>Characteristic Path Length</th>
<th>Global Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{n} \sum_{\text{nodes}} \text{average Shortest Path Length to all other nodes}$</td>
<td>$\frac{1}{n} \sum_{\text{nodes}} \text{average inverse Shortest Path Length to all other nodes}$</td>
</tr>
</tbody>
</table>

### Measures of Segregation

<table>
<thead>
<tr>
<th>Clustering Coefficient</th>
<th>Modularity (Newman’s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{n} \sum_{\text{nodes}} \text{nodal Clustering Coefficient}$</td>
<td>$\sum_{\text{Modules}} \text{Fraction of edges falling within the module minus expected fraction in a random network}$</td>
</tr>
</tbody>
</table>

$\rightarrow$ Often used to detect community structure

Rubinov & Sporns, 2010
Global Measures

Small-worldness
Optimal balance between functional segregation and integration

\[
\frac{\text{Clustering Coefficient}_{\text{real}}}{\text{Clustering Coefficient}_{\text{random}}} \quad \frac{\text{Characteristic Path Length}_{\text{real}}}{\text{Characteristic Path Length}_{\text{random}}}
\]

Functionally specialized (segregated) modules *AND* intermodular (integrating) edges

Watts & Strogatz, 1998
Resources

Analysis Software
- MATLAB-based: Brain Connectivity Toolbox (Rubinov & Sporns, 2010)
  https://sites.google.com/site/bctnet/
- Python-based: NetworkX (Hagberg et al., 2008)
  https://networkx.github.io

Visualization
- General: Gephi  http://gephi.github.io
- Anatomical space: Multimodal Connectivity Database
  http://umcd.humanconnectomeproject.org
- Anatomical space: Connectome Visualization Utility
  https://github.com/aestrivex/cvu

Reading
References