Hierarchy in Structural Brain Networks

Fani Deligianni[1]
fani.deligianni@imperial.ac.uk
Emma Robinson[1]
ecr05@doc.ic.ac.uk
David J. Sharp[2]
david.sharp@imperial.ac.uk
A. David Edwards[3]
david.edwards@imperial.ac.uk
Daniel Rueckert[1]
d.rueckert@imperial.ac.uk
Daniel C. Alexander[4]
D.Alexander@cs.ucl.ac.uk

[1] Department of Computing, Imperial College London
[2] Clinical Neuroscience Department, Imperial College London
[3] Institute of Clinical Sciences, Imperial College London
[4] Department of Computing, University College London

Abstract

Whole-brain structural connectivity matrices extracted from Diffusion Weighted Images (DWI) provide a systematic way of representing anatomical brain networks. They are equivalent to weighted graphs that encode both the topology of the network as well as the strength of connection between each pair of region of interest (ROIs). Here, we exploit their hierarchical organization to infer probability of connection between pairs of ROIs. Firstly, we extract hierarchical graphs that best fit the data and we sample across them with a Markov Chain Monte Carlo (MCMC) algorithm to produce a consensus probability map of whether or not there is a connection. We apply our technique in a gender classification paradigm and we explore its effectiveness under different parcellation scenarios. Our results demonstrate that the proposed methodology improves classification when connectivity matrices are based on parcellations that do not confound their hierarchical structure.

1 Introduction

Anatomical brain connectivity refers to the existence of axonal connections between two brain areas. With the advent of Diffusion Weighted Imaging (DWI) neuronal connections can be extracted *in vivo* and characterized non-invasively. Within tissue with an oriented structure, such as white matter, the diffusion of water is hindered in the direction perpendicular to the fiber tracts. There is an inherent complexity in exploiting this directional information of each voxel and reproducing the neuronal pathways. These limitations originate from the fact that DWI is a macroscopic technique utilised to infer microscopic tissue properties. Currently, there are several techniques to reconstruct fiber tracts [3]. Among the most successful is probabilistic tractography, which utilises a probabilistic framework to propagate local probability density functions on parameters in the diffusion model [3]. However, in

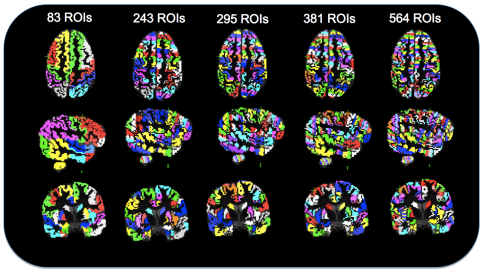


Figure 1: The original 83-ROIs parcellation as well as the four sub-parcellations based on an indicative/approximate number of 500, 400, 300 and 200 voxels per region, respectively.

probabilistic tractography a number of connections emerge that may or may not represent real fibers.

Here, we exploit the hierarchical organization of brain networks to redefine whole-brain connectivity matrices. We propose inferring hierarchical structures from the observed anatomical connectivity with a technique that has been developed recently and has been tested in both biological and social networks [2]. This technique uses statistical inference combined with a MCMC sampling algorithm to derive hierarchical models, also called dendrograms, with probability proportional to the likelihood that they generate the observed network. This model allows the assignment of a probability for each connection that reflects the confidence in its existence based on the whole-brain network topology and the assumption that it is hierarchically organized. We use this methodology to analyze whole-brain connectivity matrices derived from a number of different sub-parcellations (scales). We applied our approach in the paradigm of gender classification. Leave-one-out cross-validation is used to compare the performance of classification with and without the application of the hierarchical algorithm.

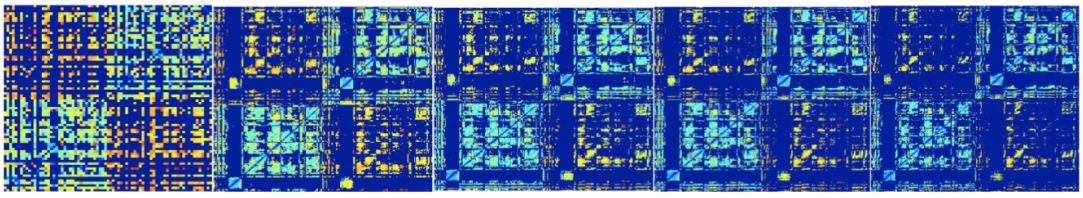
2 Methods

2.1 Pre-processing and Extraction of Brain Networks

FSL was the main tool for pre-processing of DWI . This involved eddy current correction and brain extraction. Bias correction was applied to T1 and B0 images to improve the robustness of the non-rigid registration tools. In order to extract anatomical brain networks from DWI, ROIs are defined based on the fusion of 83-ROIs atlas based segmentation and soft-tissue segmentation. This facilitated the extraction of ROIs that are anatomically sensible and they are located in gray-matter. Segmentations were transformed to diffusion space with non-rigid registration. This procedure has been described in details in previous work [4]. Subsequently, connections between regions are identified using a standard probabilistic algorithm available as part of FSL [1]. However, we estimate the local diffusion anisotropy by determining the diffusive transfer between voxels using the orientation distribution function (ODF) [4].

2.2 Sub-Parcellation of ROIs

To sub-parcellate the 83 segmented ROIs we transformed the segmentations from diffusion space into standard (MNI) space using affine registration. We counted the voxels within each



(a) 83 ROIs (b) 243 ROIs (500 voxels) (c) 295 ROIs (400 voxels) (d) 381 ROIs (300 voxels) (e) 564 ROIs (200 voxels)

Figure 2: Structural connectivity matrices that correspond to the parcellations in Figure 1.

region and across subjects and we predefined the number of sub-areas based on an indicative (approximate) number of voxels per area. Each area was sub-divided into a number of regions, according to the average number of voxels across subjects divided by the indicative size and rounded to the closest integer. Subsequently, we extracted the boundary voxels of each ROI with white matter and we applied eigen-decomposition to the covariance matrix to define the best fit plane for the boundary voxels. The coordinates of each voxel in the ROI was projected on the first eigen-vector so that we could sub-divide them based on their projected position. Finally, they were subdivided into the predefined number of sub-regions so that each sub-region has equal number of voxels (± 1). This way guarantees that all sub-regions within the original ROI are surrounded with relatively equal number of white matter voxels. Therefore, tracts between each sub-region and the rest ROIs can be identified. An example of the subparcellation for a subject is shown in Fig. 1.

2.3 From Brain Graphs to Hierarchical Random Graphs

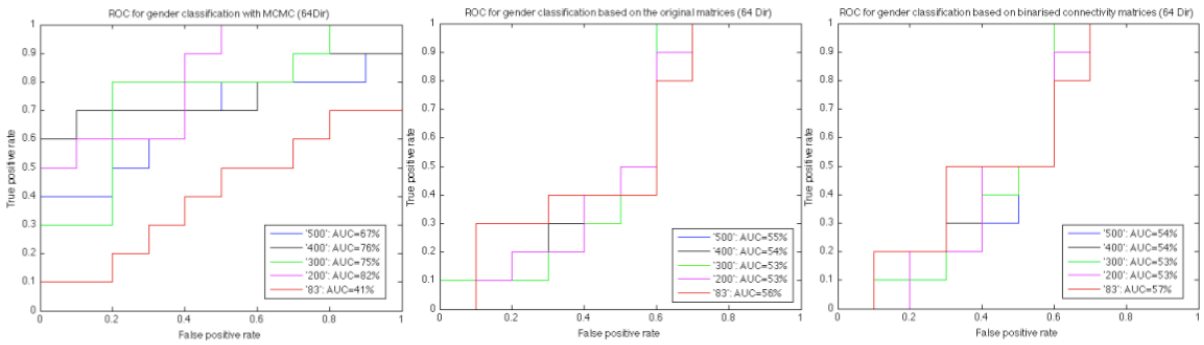
In networks with hierarchical organization, nodes are subdivided into groups that are further subdivided into more groups and so forth over multiple scales. In brain networks, this implies that connections are dense within groups of areas and sparse between them. Clauset et al. showed that this property alone is able to explain both qualitatively and quantitatively a number of topological and statistical properties of the original graph, such as their degree distribution, clustering coefficient and so on [2]. Their approach offers two major strengths: Firstly, it can capture both clusters of nodes, 'assortative', and disassociated nodes, 'disassortative' structures, as well as arbitrary mixtures of the two. Secondly, it does not depend on one hierarchical model but it generates a series of hierarchical models, dendrograms, and samples along them to create a consensus dendrogram that expresses the network's topology.

Let us represent a structural brain network as a graph G with n nodes. The observed network data can be fitted to a random binary dendrogram based on a Monte Carlo sampling algorithm over a maximum-likelihood approach. A binary dendrogram has n leaves corresponding to the n nodes of the graph G, and each of the $n-1$ internal nodes have exactly two descendants. Each internal node r is associated with a probability p_r. Under a maximum-likelihood approach the probability p_r is estimated as the fraction of edges between the two sub-trees:

$$\bar{p}_r = \frac{E_r}{L_r R_r} \tag{1}$$

where E_r are the edges between left and right sub-trees and L_r and R_r are the number of nodes in left and right sub-tree, respectively. The likelihood of the dendrogram at this maximum is given below:

$$L(D) = \prod_{r \in D} \left[\bar{p}_r^{\bar{p}_r} (1-\bar{p}_r)^{1-\bar{p}_r} \right]^{L_r R_r}. \tag{2}$$



(a) MCMC (b) Original Matrices (c) Binarised Matrices

Figure 3: ROC curves for gender classification.

The overall likelihood of the dendrogram reflects how well the dendrogram fits the graph data under the assumption that the original graph has hierarchical organization. Here we used the logarithm of the likelihood to avoid numerical instabilities due to very small numbers:

$$\log L(D) = -\sum_{r \in D} L_r R_r h(\bar{p}_r) \tag{3}$$

where h is the Gibbs-Shannon entropy function: $h(p) = -p\log p - (1-p)\log(1-p)$. Note that dendrograms with high probability are those that partition the vertices into groups that are either very well interconnected or disconnected. Subsequently, the MCMC method is used to sample dendrograms and accepts them according to the Metropolis-Hasting rules [2]. The Markov chain consists of re-arrangements of subtrees of the dendrogram by choosing a random node and exchanging any of its children with its parent's child.

Once MCMC reaches equilibrium, dendrograms can be sampled at regular interval from the Markov chain. For each sampled dendrogram a probability connectivity matrix can be created with values that reflect the probability of each pair of nodes/areas to be connected. The probability between each pair of nodes i and j is equal to the probability p_r of the lowest common ancestor of the nodes i and j in the sampled dendrogram. A consensus probability matrix is estimated by averaging these matrices across dendrogram samples.

3 Results

We used diffusion weighted images (DWI) that have been acquired from 20 normal volunteers (10 males, 10 females) with the following imaging parameters: 64 non-collinear directions, in 72 slices, slice thickness $2mm$, FOV $224mm$, matrix 128x128, voxel size 1.75x1.75x$2mm^3$, b value 1000 s/mm^2 (Philips 3Tesla). Based on the original 83-ROIs segmentation, we created four sub-parcellations with an indicative size of region of 500, 400, 300 and 200 voxels per area. This resulted in four segmentations with 243, 295, 381 and 564 regions, respectively, Fig. 1. We run the adapted probabilistic tractography [4], which estimated the connectional strength between each pair of regions and provided with the corresponding connectivity matrices for each sub-parcellation, Fig. 2. For each subject and each parcellation, we run the MCMC method until the algorithm reached equilibrium. Once the MCMC has reach equilibrium, the probability that there is a connection between each pair of nodes is estimated and it is averaged across a predefined number of dendrograms (5000). These new connectivity matrices define the probability that two nodes are connected.

We performed gender classification by separating subjects into two groups according to their gender and averaging the mean-subject probability connectivity matrices. We used leave-one-out cross validation to classify the subjects based on the normalised Euclidean distance between the average probability maps of each group and the leave-one-out subject. If the subject's distance from the male group is higher than its distance from the female group, we classify the subject as female and vice-versa. To demonstrate the efficiency of our methodology we also performed classification with the original connectivity matrices and the same classifier. In Fig. 3, the Receiver Operating Characteristic (ROC) curve is shown for each of the sub-parcellations with and without MCMC. The Area Under the Curve (AUC) is a measure of the optimum performance of the classifier. Our results suggest that age classification is not successful based on the original parcellation alone neither with the application of the MCMC algorithm or without, Fig. 3 . However, when we analyzed the sub-parcelated connectivity matrices with the MCMC approach, we could classify the subjects with up to 82% classification rate. Gender classification based on the original connectivity matrices was unsuccessful for all the different sub-parcellation scenarios. Since MCMC has as input binary matrices (one when a connection exist, zero for the absence of connection), we also classified the binarised version of the original connectivity matrices. There was no significant difference in performance between the original weighted graphs and the binarised versions, Fig. 3.

4 Discussion and Conclusions

We applied our approach successfully in a gender classification paradigm of 20 subjects with diffusion data in 64 directions, Fig. 3. Our results showed that although classification performance did not improve in the case of the original 83-ROIs parcellation, it was significantly enhanced in all other sub-parcellations. This suggests that hierarchical organization of anatomical networks derived from DWI can be confounded by the uneven sub-parcellation in 83-ROIs with size that varies from 20 to over 8000 voxels per region. On the other hand, MCMC applied on connectivity matrices derived from the sub-parcellation has a significant improvement in classification over the original connectivity matrices, Fig. 3. These results demonstrate the potential of the MCMC to identify missing links and false connections in whole-brain structural connectivity data.

References

[1] TEJ Behrens, MW Woolrich, M Jenkinson, H Johansen-Berg, RG Nunes, S Clare, PM Matthews, JM Brady, and SM Smith. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnet Reson Med*, 50(5):1077–1088, 2003.

[2] A Clauset, C Moore, and M Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.

[3] H Johansen-Berg and T Behrens. Diffusion MRI: from quantitative measurement to in-vivo neuroanatomy. *Academic Press*, page 490, 2009.

[4] E Robinson, A Hammers, A Ericsson, AD Edwards, and D Rueckert. Identifying population differences in whole-brain structural networks: a machine learning approach. *NeuroImage*, 50(3):910–919, 2010.