# Ordinal Patterns for Connectivity Networks in Brain Disease Diagnosis

Mingxia Liu, Junqiang Du, Biao Jie, and Daoqiang Zhang(✉)

School of Computer Science and Technology,
Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
dqzhang@nuaa.edu.cn

**Abstract.** Brain connectivity networks have been widely used for diagnosis of brain-related diseases, *e.g.*, Alzheimer's disease (AD), mild cognitive impairment (MCI), and attention deficit hyperactivity disorder (ADHD). Although several network descriptors have been designed for representing brain connectivity networks, most of them not only ignore the important weight information of edges, but also cannot capture the modular local structures of brain connectivity networks by only focusing on individual brain regions. In this paper, we propose a new network descriptor (called ordinal pattern) for brain connectivity networks, and apply it for brain disease diagnosis. Specifically, we first define ordinal patterns that contain sequences of weighted edges based on a functional connectivity network. A frequent ordinal pattern mining algorithm is then developed to identify those frequent ordinal patterns in a brain connectivity network set. We further perform discriminative ordinal pattern selection, followed by a SVM classification process. Experimental results on both the ADNI and the ADHD-200 data sets demonstrate that the proposed method achieves significant improvement compared with state-of-the-art brain connectivity network based methods.

## 1 Introduction

As a modern brain mapping technique, functional magnetic resonance imaging (fMRI) is an efficient as well as non-invasive way to map the patterns of functional connectivity of the human brain [1,2]. In particular, the task-free (resting-state) functional magnetic resonance imaging (rs-fMRI) have a small-world architecture, which can reflect a robust functional organization of the brain. Recent studies [3–6] show great promises of brain connectivity networks in understanding brain diseases (*e.g.*, AD, MCI, and ADHD) pathology by exploring anatomical connections or functional interactions among different brain regions, where brain regions are treated as nodes and anatomical connections or functional associations are regarded as edges.

Several network descriptors have been developed for representing brain connectivity networks, such as node degrees [3], clustering coefficients [4], and sub-networks [7]. Most of existing descriptors are designed on un-weighted brain connectivity networks, where the valuable weight information of edges are ignored.

M. Liu and J. Du—These authors contribute equally for this paper.


S. Ourselin et al. (Eds.): MICCAI 2016, Part I, LNCS 9900, pp. 1–9, 2016.
DOI: 10.1007/978-3-319-46720-7_1

Actually, different edges are usually assigned different weights to measure the connectivity strength between pairs of nodes (*w.r.t.* brain regions). However, previous studies usually simply apply thresholds to transform the original weighted networks into un-weighted ones [2,5], which may lead to sub-optimal learning performance. In addition, existing descriptors mainly focus on individual brain regions other than local structures of brain networks, while many evidences have declared that some brain diseases (*e.g.*, AD and MCI) are highly related to modular local structures [8]. Unfortunately, it is hard to capture such local structures using existing network descriptors.

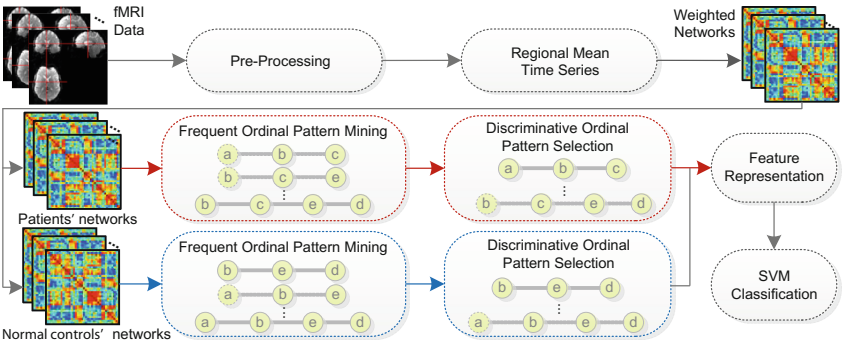


**Fig. 1.** An overview of ordinal pattern based learning for brain disease diagnosis.

In this paper, we propose a new network descriptor, *i.e.*, ordinal pattern, for brain connectivity networks. The basic idea of the ordinal pattern is to construct a sequence of weighted edges on a weighted network by considering both the edge weights and the ordinal relations between edges. Compared with conventional network descriptors, ordinal patterns are directly constructed on weighted networks, which can naturally preserve the weight information and local structures of original networks. Then, an ordinal pattern based learning method is developed for brain disease diagnosis. Figure 1 presents the schematic diagram of the proposed framework with each network representing a specific subject. We first construct ordinal patterns on patients' and normal controls' (NCs) brain connectivity networks separately. A frequent ordinal pattern mining algorithm is then developed to identify ordinal patterns that frequently occur in patients' and NCs' brain networks. We then select the most discriminative ordinal patterns from those frequent ordinal patterns, and regard them as feature representation for subjects. Finally, we learn a support vector machine (SVM) classifier for brain disease diagnosis, by using ordinal pattern based feature representation.

## 2 Method

### 2.1 Data and Preprocessing

The first data set contains rs-fMRI data from the ADNI[1] database with 34 AD patients, 99 MCI patients, and 50 NCs. The rs-fMRI data were pre-processed by brain skull removal, motion correction, temporal pre-whitening, spatial smoothing, global drift removal, slice time correction, and band pass filtering. By warping the automated anatomical labelling (AAL) [9] template, for each subject, we concatenate the brain space of rs-fMRI scans into 90 regions of interest (ROIs). For each ROI, the rs-fMRI time series of all voxels were averaged to be the mean time series of the ROI. With ROIs as nodes and Pearson correlations between pair of ROIs as connectivity weights, a functional full connected weighted network is constructed for each subject. The second data set is ADHD-200 with the Athena preprocessed rs-fMRI data, including 118 ADHD patients and 98 NCs (detailed description of data acquisition and post-processing are given online[2].

### 2.2 Ordinal Pattern and Frequent Ordinal Pattern

**Definition 1: Ordinal Pattern.** Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathbf{w}\}$ denote a weighted network, where $\mathcal{V}$ is a set of nodes, $\mathcal{E}$ is a set of edges, and $\mathbf{w}$ is the weight vector for those edges with the $i$-th element $w(e_i)$ representing the weight value for the edge $e_i$. If $w(e_i) > w(e_j)$ for all $0 < i < j \leq M$, an *ordinal pattern* ($op$) of $\mathcal{G}$ is defined as $op = \{e_1, e_2, \cdots, e_M\} \subseteq \mathcal{E}$, where $M$ is the number of edges in $op$.

An illustration of the proposed ordinal patterns is given in Fig. 2(a), where a weighted network contains 5 nodes and 7 edges. We can get ordinal patterns that contain two edges, *e.g.*, $op^1 = \{e_{a-b}, e_{b-c}\}$ and $op^2 = \{e_{b-c}, e_{c-e}\}$. The ordinal pattern $op^1$ actually denotes $w(e_{a-b}) > w(e_{b-c})$. We can further obtain

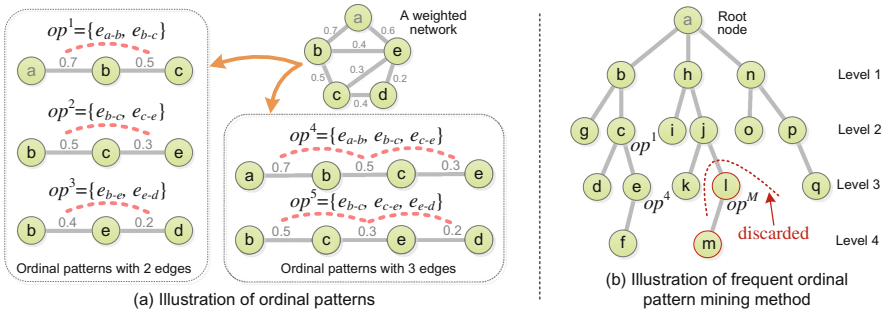


**Fig. 2.** Illustration of (a) ordinal patterns, and (b) frequent ordinal pattern mining method.

[1] http://adni.loni.usc.edu/.
[2] http://www.nitrc.org/plugins/mwiki/index.php/neurobureau:AthenaPipeline.

ordinal patterns containing three edges, *e.g.*, $op^4 = \{e_{a-b}, e_{b-c}, e_{c-e}\}$. Hence, the proposed ordinal pattern can be regarded as the combination of some ordinal relations between pairs of edges. We only consider connected ordinal patterns in this study. That is, an ordinal pattern is connected if and only if the edges it contains can construct a connected sub-network. Different from conventional methods, the ordinal pattern is defined on a weighted network directly to explicitly utilize the weight information of edges. Also, as a special sub-network, an ordinal pattern can model the ordinal relations conveyed in a weighted network, and thus, can naturally preserve the local structures of the network.

**Definition 2: Frequent Ordinal Pattern.** Let $\mathcal{D} = \{\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N\}$ represent a set of $N$ weighted networks. Given an ordinal pattern $op$, the frequency ratio of $op$ is defined as follows

$$f(op|\mathcal{D}) = \frac{|\mathcal{G}_n|op \text{ is an ordinal pattern of } \mathcal{G}_n, \mathcal{G}_n \in \mathcal{D}|}{|\mathcal{D}|} \tag{1}$$

If $f(op|\mathcal{D}) > \theta$ where $\theta$ is a pre-defined threshold value, the ordinal pattern $op$ is called as a *frequent ordinal pattern* of $\mathcal{D}$.

We can see that frequent ordinal patterns are ordinal patterns that frequently appear in a weighted network set. For instance, a frequent ordinal pattern in a brain network set may represent common functional or structural information among subjects. Besides, frequent ordinal patterns have an appealing property that plays an important role in data mining process. Specifically, for two ordinal patterns $op^i = \{e_1^i, e_2^i, \cdots, e_M^i\}$ and $op^j = \{e_1^j, e_2^j, \cdots, e_M^j, e_{M+1}^j\}$, if $e_m^i = e_m^j$ $(\forall m \in \{1, 2, \cdots, M\})$, $op^i$ is called the parent of $op^j$, and $op^j$ is called a child of $op^i$. As shown in Fig. 2(a), $op^1 = \{e_{a-b}, e_{b-c}\}$ is the parent of $op^4 = \{e_{a-b}, e_{b-c}, e_{c-e}\}$. It is easy to prove that the frequency ratio of an ordinal pattern is no larger than the frequency ratios of its parents. That is, if an ordinal pattern is not a frequent ordinal pattern, its children and descendants are not frequent ordinal patterns, either.

### 2.3 Ordinal Pattern Based Learning

**Ordinal Pattern Construction:** Using the above-mentioned preprocessing method, we can construct one brain connectivity network for each subject, with each node denoting a ROI and each edge representing Pearson correlation between a pair of ROIs. We then construct ordinal patterns on patients' and normal controls' (NCs) brain connectivity networks separately. Given all training subjects, we can obtain a brain network set with patients' and NCs' networks.

**Frequent Ordinal Pattern Mining:** We then propose a frequent ordinal pattern mining algorithm to identify ordinal patterns that are frequently occur in a brain network set, by constrcuing a deep first search (DFS) tree. We first

randomly choose an edge whose frequency ratio is larger than a threshold $\theta$ as the root node. As illustrated in Fig. 2(b), a path from the root node to the current node forms a specific ordinal pattern, *e.g.*, $op^1 = \{e_{a-b}, e_{b-c}\}$. We then record the number of occurrences and compute the frequency ratio of this ordinal pattern in a network set (with each network corresponding to a subject). If its frequency ratio defined in Eq. (1) is larger than $\theta$, the ordinal pattern (*e.g.*, $op^1$) is a frequent ordinal pattern and its children (*e.g.*, $op^4$) will be further searched. Otherwise, the ordinal pattern (*e.g.*, $op^M$) is not a frequent ordinal pattern, and its descendants will be discarded directly. The max depth of a DFS tree is limited by the level number. For example, if the level is 3, the frequent ordinal patterns contain at most 3 edges. Obviously, more levels bring more frequent ordinal patterns as well as more run-time.

**Discriminative Ordinal Pattern Selection:** There are a number of frequent ordinal patterns, and some of them could have less discriminative power. Accordingly, we perform a discriminative ordinal pattern selection process on those frequent ordinal patterns. Specifically, we first mine frequent ordinal patterns from the patients' brain network set and the NCs' brain network set separately. According to the discriminative power, we select the most discriminative ordinal patterns from all frequent ordinal patterns in both patients' and NCs' sets. The ratio score [10] is used to evaluate the discriminative power of frequent ordinal patterns. Given a frequent ordinal pattern $op^i$ mined from the patients' brain network set (denoted as $\mathcal{D}^+$), the ratio score of $op^i$ is defined as

$$RS(op^i) = \log \frac{|\mathcal{G}_n | op^i \text{ is an ordinal pattern of } \mathcal{G}_n, \mathcal{G}_n \in \mathcal{D}^+|}{|\mathcal{G}_n | op^i \text{ is an ordinal pattern of } \mathcal{G}_n, \mathcal{G}_n \in \mathcal{D}^-| + \epsilon} \times \frac{|\mathcal{D}^-|}{|\mathcal{D}^+|} \tag{2}$$

where $\mathcal{D}^-$ means the NCs' brain network set, and $\epsilon$ is a small value to prevent the denominator to be 0. Similarly, the frequent ordinal pattern $op^j$ mined from the NCs' brain network set (*i.e.*, $\mathcal{D}^-$), its ratio score is computed as

$$RS(op^j) = \log \frac{|\mathcal{G}_n | op^j \text{ is an ordinal pattern of } \mathcal{G}_n, \mathcal{G}_n \in \mathcal{D}^-|}{|\mathcal{G}_n | op^j \text{ is an ordinal pattern of } \mathcal{G}_n, \mathcal{G}_n \in \mathcal{D}^+| + \epsilon} \times \frac{|\mathcal{D}^+|}{|\mathcal{D}^-|} \tag{3}$$

**Classification:** A total of $k$ discriminative ordinal patterns are first selected, with half from patients' and the other half from NCs' brain connectivity network sets. We then combine those discriminative ordinal patterns to construct a feature matrix for representing subjects. Specifically, given $|\mathcal{D}|$ brain connectivity networks (with each network corresponding to a specific subject) and $k$ selected discriminative ordinal patterns, we denote the feature matrix as $\mathbf{F} \in \mathbb{R}^{|\mathcal{D}| \times k}$, where the element $F_{ij}$ represents the $j$-th feature of the $i$-th subject. Specifically, if the $j$-th discriminative ordinal pattern appears in the brain connectivity network of the $i$-th subject, $F_{i,j}$ is equal to 1, and otherwise 0. Finally, we adopt an SVM classifier to identify AD/MCI/ADHD patients from NCs.

## 3 Experiments

**Experimental Settings:** We perform three classification tasks, *i.e.*, AD vs. NC, MCI vs. NC and ADHD vs. NC classification, by using a 10-fold cross-validation strategy. Note that those discriminative ordinal patterns are selected only from training data. Classification performance is evaluated by accuracy (ACC), sensitivity (SEN), specificity (SPE) and area under the ROC curve (AUC). The parameter $\epsilon$ in ratio score in Eqs. (2) and (3) is set as 0.1 empirically. With a inner cross-validation strategy, the level number in our frequent ordinal pattern mining algorithm is chosen from $[2, 6]$ with step 1, and the number of discriminative ordinal patterns are chosen from $[10, 100]$ with step 10.

We compare our method with two widely used network descriptors in brain connectivity network based studies, including cluster coefficients [4] and discriminative sub-networks [7]. Since these two descriptors require a thresholding process, we adopt both single-threshold and multi-thresholds [5,11] strategies to transform weighted networks to un-weighted ones. In summary, there are four competing methods, including (1) clustering coefficients (CC) with single-threshold, (2) clustering coefficient using multi-thresholds (CCMT), (3) discriminative sub-networks (DS) with single-threshold, and (4) discriminative sub-networks using multi-thresholds (DSMT). The linear SVM with the default parameter (*i.e.*, $C = 1$) is used as the classifier in different methods.

**Results:** Experimental results are listed in Table 1, from which we can see that our method consistently achieves the best performance in three tasks. For instance, the accuracy achieved by our method is 94.05 % in AD vs. NC classification, which is significantly better than the second best result obtained by DSMT. This demonstrates that the ordinal patterns are discriminative in distinguishing AD/MCI/ADHD patients from NCs, compared with conventional network descriptors.

We further plot those top 2 discriminative ordinal patterns identified by our method in three tasks in Fig. 3. For instance, the most discriminative ordinal pattern for AD, shown in top left of Fig. 3(a), can be recorded as $op = \{e_{DCG.L-ACG.L}, e_{ACG.L-ROL.L}, e_{ROL.L-PAL.R}, e_{PAL.R-LING.L}, e_{PAL.R-MOG.R}\}$.

**Table 1.** Comparison of different methods in three classification tasks

| Method | AD vs. NC | | | | MCI vs. NC | | | | ADHD vs. NC | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ACC (%) | SEN (%) | SPE (%) | AUC (%) | ACC (%) | SEN (%) | SPE (%) | AUC (%) | ACC (%) | SEN (%) | SPE (%) | AUC (%) |
| CC | 72.62 | 73.53 | 67.94 | 70.94 | 71.14 | 72.73 | 68.00 | 68.69 | 71.29 | 72.03 | 70.41 | 70.51 |
| CCMT | 80.95 | 82.35 | 80.00 | 76.35 | 74.50 | 75.76 | 72.00 | 74.79 | 74.53 | 75.43 | 73.47 | 77.64 |
| DS | 76.19 | 76.47 | 76.00 | 75.59 | 77.18 | 78.79 | 74.00 | 74.89 | 81.01 | 81.36 | 80.61 | 80.82 |
| DSMT | 85.71 | 85.29 | 86.00 | 87.59 | 79.19 | 80.81 | 76.00 | 76.99 | 83.79 | 84.74 | 82.65 | 84.63 |
| Proposed | **94.05** | **96.77** | **92.45** | **96.35** | **88.59** | **87.27** | **92.31** | **84.57** | **87.50** | **88.89** | **85.85** | **87.37** |

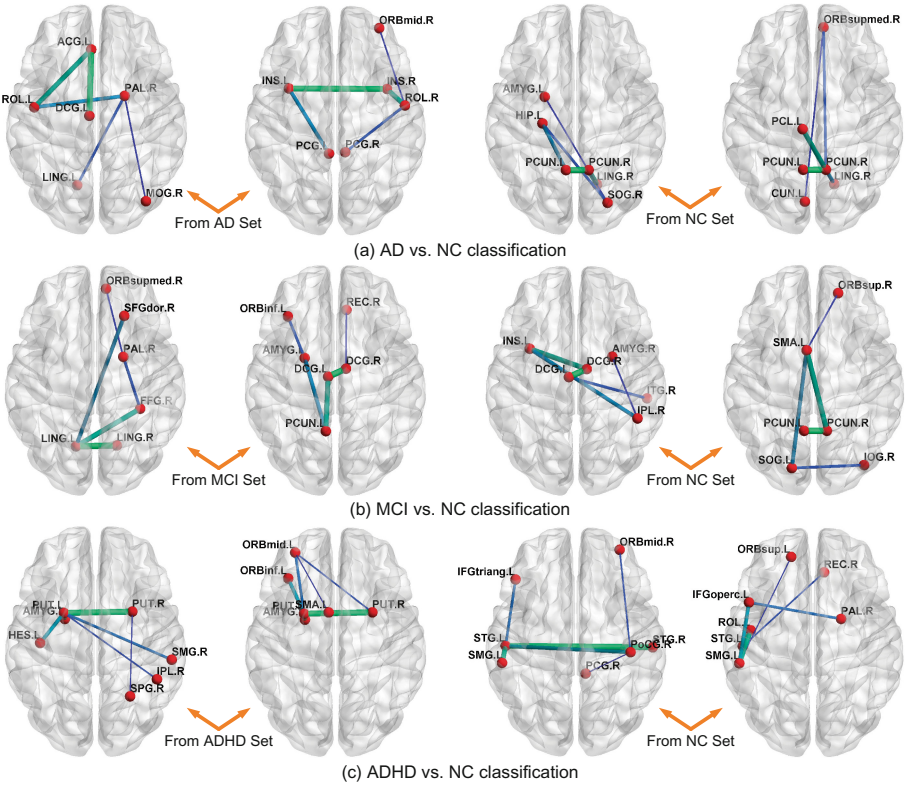


**Fig. 3.** The most discriminative ordinal patterns identified by the proposed method in three tasks. In each row, the first two columns show those top 2 discriminative ordinal patterns selected from positive classes (*i.e.*, AD, MCI, and ADHD), while the last two columns illustrate those selected from the negative class (*i.e.*, NC).

These results imply that the proposed ordinal patterns do reflect some local structures of original brain networks.

We investigate the influence of frequent ordinal pattern mining level and the number of selected discriminative ordinal patterns, with results shown in Fig. 4. From this figure, we can see that our method achieves relatively stable results when the number of selected ordinal patterns is larger than 40. Also, our method achieves overall good performance when the level number in the frequent ordinal pattern mining algorithm are 4 in AD/MCI vs. NC classification and 5 in ADHD vs. NC classification, respectively.

We perform an additional experiment by using weights of each edge in ordinal patterns as raw features, and achieve the accuracies of 71.43 %, 67.11 %, and 69.91 % in AD vs. NC, MCI vs. NC and ADHD vs. NC classification, respectively. We further utilize a real valued network descriptor based on ordinal patterns (by taking the product of weights in each ordinal pattern), and obtained the accuracies of 78.52 %, 72.37 %, and 72.69 % in three tasks, respectively.

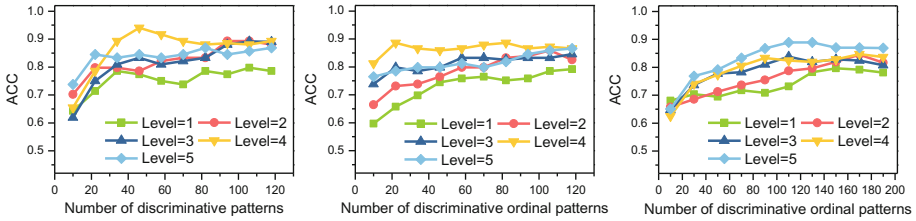


**Fig. 4.** Influence of the level number in frequent ordinal pattern mining method and the number of discriminative ordinal patterns in AD vs. NC (left), MCI vs. NC (middle), and ADHD vs. NC (right) classification.

## 4 Conclusion

In this paper, we propose a new network descriptor (*i.e.*, ordinal pattern) for brain connectivity networks. The proposed ordinal patterns are defined on weighted networks, which can preserve the weights information of edges and the local structure of original brain networks. Then, we develop an ordinal pattern based brain network classification method for the diagnosis of AD/MCI and ADHD. Experimental results on both ADNI and ADHD-200 data sets demonstrate the efficacy of our method.

**Acknowledgment.** This study was supported by National Natural Science Foundation of China (Nos. 61422204, 61473149, 61473190, 61573023), the Jiangsu Natural Science Foundation for Distinguished Young Scholar (No. BK20130034), and the NUAA Fundamental Research Funds (No. NE2013105).

## References

1. Robinson, E.C., Hammers, A., Ericsson, A., Edwards, A.D., Rueckert, D.: Identifying population differences in whole-brain structural networks: a machine learning approach. NeuroImage **50**(3), 910–919 (2010)
2. Sporns, O.: From simple graphs to the connectome: networks in neuroimaging. NeuroImage **62**(2), 881–886 (2012)
3. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. NeuroImage **52**(3), 1059–1069 (2010)
4. Wee, C.Y., Yap, P.T., Li, W., Denny, K., Browndyke, J.N., Potter, G.G., Welsh-Bohmer, K.A., Wang, L., Shen, D.: Enriched white matter connectivity networks for accurate identification of MCI patients. NeuroImage **54**(3), 1812–1822 (2011)
5. Jie, B., Zhang, D., Wee, C.Y., Shen, D.: Topological graph kernel on multiple thresholded functional connectivity networks for mild cognitive impairment classification. Hum. Brain Mapp. **35**(7), 2876–2897 (2014)
6. Liu, M., Zhang, D., Shen, D.: Relationship induced multi-template learning for diagnosis of Alzheimer disease and mild cognitive impairment. IEEE Trans. Med. Imaging **35**(6), 1463–1474 (2016)
7. Fei, F., Jie, B., Zhang, D.: Frequent and discriminative subnetwork mining for mild cognitive impairment classification. Brain Connect. **4**(5), 347–360 (2014)

8. Brier, M.R., Thomas, J.B., Fagan, A.M., Hassenstab, J., Holtzman, D.M., Benzinger, T.L., Morris, J.C., Ances, B.M.: Functional connectivity and graph theory in preclinical Alzheimer's disease. Neurobiol. Aging **35**(4), 757–768 (2014)
9. Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M.: Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage **15**(1), 273–289 (2002)
10. Yan, X., Cheng, H., Han, J., Yu, P.S.: Mining significant graph patterns by leap search. In: Proceedings of ACM SIGMOD International Conference on Management of Data, pp. 433–444. ACM (2008)
11. Sanz-Arigita, E.J., Schoonheim, M.M., Damoiseaux, J.S., Rombouts, S., Maris, E., Barkhof, F., Scheltens, P., Stam, C.J., et al.: Loss of 'small-world' netowrks in Alzheimer's disease: graph analysis of FMRI resting-state functional connectivity. PLoS ONE **5**(11), e13788 (2010)