Information Sciences 384 (2017) 145–156

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Graph theoretical analysis of Alzheimer's disease: Discrimination of AD patients from healthy subjects

Mahdi Jalili

Department of Electrical and Computer Engineering, School of Engineering RMIT University, Melbourne, Australia

ARTICLE INFO

Article history:
Received 1 November 2015
Revised 1 August 2016
Accepted 14 August 2016
Available online 16 August 2016

Keywords:
EEG
Alzheimer's disease
Classification
Feature selection
Network science
Graph theory

ABSTRACT

Tools available in graph theory have been recently applied to signals recorded from the human brain, where its cognitive functions are linked to topological properties of connectivity networks. In this work, we consider resting-state electroencephalography (EEG) signals recorded from healthy subjects and patients suffering from Alzheimer's disease (AD) in two conditions: eyes-open and eyes-closed. The EEGs are used to construct functional brain networks in which the nodes are EEG sensor locations and edges represent functional connectivity between them. The networks are then tested for a number of neurobiologically relevant graph theory metrics. The analyses show that the network properties are stable across all conventional frequency bands. AD brains in eyes-closed condition show significantly reduced local efficiency and modularity measures ($P < 0.05$; Wilcoxon's ranksum test). We then use the network metrics as features for discriminating AD from healthy controls. Three feature selection methods (Genetic Algorithms (GA), Binary Particle Swarm Optimization (BPSO) and Social Impact Theory based Optimization (SITO)) are used to select the best feature set. GA with support vector machines (as classifier) results in an accuracy of 83% in eyes-close beta band. The set of optimal features include edge betweenness centrality, global efficiency, modularity and synchronizability.

1. Introduction

Human brain is a complex network comprised of a number of nodes (brain regions) connected through anatomical/functional links. Modelling the brain as a networked structure and studying its static and dynamic properties has received significant attention in recent years, which is mainly due to recent advances in network science tools and data processing techniques [8]. Data recorded from Electroencephalography (EEG), Magnetocephalography (MEG) and functional Magnetic Resonance Imaging (fMIR) can be used to extract functional brain networks [2,16,34], while anatomical brain networks can be extracted through Diffusion Tensor Imaging (DTI) technique [12]. In order to extract functional connectome of the brain, the first step is to obtain weighted connectivity matrices from EEG, MEG or fMIR data. To this end, linear techniques such as Pearson correlation and coherence analysis or nonlinear methods such as synchronization likelihood can be used. Often, the weighted connectivity matrices are binarized to obtain reliable network measures. To binarize a weighted network, one can use methods such as sparsity thresholding [1] or minimum spanning three [35].

Studying network properties in various brain disorders have revealed disease-specific abnormalities. Disorders such as schizophrenia [18], Alzheimer's disease [34] and epilepsy [6] show altered functional networks. AD symptoms are the most cause of dementia in the brain, often leading to the death. Previous studies showed that functional and anatomical networks

E-mail address: mahdi.jalili@rmit.edu.au

http://dx.doi.org/10.1016/j.ins.2016.08.047
0020-0255/

Table 1
Demographic data of subjects used in this study.

	AD patients	Healthy controls	*P*-value (Wilcoxon's ranksum test)
Size	23 (9 female)	25 (11 female)	
Age	72.09 ± 10.05	67.96 ± 11.19	0.27
MMSE	22.43 ± 4.34	29 ± 1	< 0.0001

in AD patients demonstrate altered network properties such as average path length, clustering coefficient, small-worldness and synchronizability [10,33–35,40]. EEG, as a non-invasive and cheap neuroimaging modality, has been used to study the AD mechanisms in the brain [3,5,24]. EEG-based networks have been studied in AD. For example, Stam et al. studied AD networks extracted from EEG and found that AD networks show longer shortest path length than healthy controls in beta band [34]. The clustering coefficient also altered in AD networks [40]. Tahaei et al. showed decreased synchronizability (as measured by the eigenratio of the Laplacian matrix of the connection graph) in AD subjects [35]. Afshari and Jalili studied directed networks in AD brains and showed abnormalities in global and local efficiency measures [4].

Previous studies on EEG-based brain functional networks were mainly in resting-state eyes-closed condition or when the subjects were performing a cognitive task. In this work, we study the resting-state with both eyes-open and eyes-closed conditions. We consider EEGs recorded from 25 healthy control subjects and 23 patients suffering from AD and analyse the network properties in different frequency bands. EEG signs have been previously used to classify AD patients from healthy control subjects [26,32,36]. We use EEG-based graph theory metrics to discriminate AD patients from healthy controls, which to the best of our knowledge, has not been studied in previous research works. Previous research works have used graph theory features of AD networks obtained from fMRI or MEG to classify AD from healthy controls. Zanin et al. studied MEG-based networks in Mild Cognitive Impairment (MCI), which is known to be early stages of AD development [39]. They showed that one can correctly classify MCI patients from healthy subjects with an accuracy of about 80% for networks with small to medium density values. By studying fMIR-based networks of MCI patients, Wang et al. showed significant decrease of global and functional connectivity in MCI [38]. Jie et al. also studied fMRI-based functional networks of MCI patients and obtained an accuracy of 89% using multi-kernel Support Vector Machines (SVM) classifier [21].

In this work, a number of graph theory features of functional networks are used to classify AD patients from healthy subjects. Also, a number of feature selection methods are used to choose a set of optimal features for the classification task. The feature selection algorithms are population-based optimization methods, which have been frequently applied to this problem [27,30]. We find that the AD patients have statistically different local connectivity than healthy controls in the eyes-closed condition. Furthermore, SVM with optimal features can correctly predict up to 83% of the subjects.

2. Methods

2.1. Subjects and EEG recording

The EEGs of 23 newly diagnosed patients suffering from AD symptoms and 25 healthy controls are considered in this study. The subjects were recruited from the Memory Clinic of the Neurology Department (CHUV, Lausanne). The AD and control groups are not different in their age and educational level. AD patients show significantly less Mini Mental Test Examination (MMSE) scores than control subjects. Table 1 shows demographic information of the subjects used in this study. All the patients, caregivers, and control subjects gave written informed consent. All the applied procedures conform to the Declaration of Helsinki (1964) by the World Medical Association concerning human experimentation and were approved by the local Ethics Committee of Lausanne University.

The EEGs were recorded in two resting-state conditions: eyes-closed and eyes-open. The data were collected while subjects were sitting relaxed in a semi-dark room. To record the EEG data for duration of 3–4 min for each subject, the 128-channel Geodesic Sensor Net (EGI, USA) machine was used. The recordings were made with vertex reference at a sampling frequency of 500 Hz, and were further filtered (FIR, band-pass of 1–50 Hz; 50 Hz notch filter) and re-referenced against the common average reference. Then, the data were segmented into non-overlapping epochs each with 1 s length (500 samples in each epoch). All computations were first performed on individual epochs and then averaged over all artifact-free epochs in order to obtain the measures for each individual subject. The data recorded from outer electrodes (18 in total) were removed due to lower signal-to-noise ratio in these electrodes. The EEGs were filtered in conventional frequency bands, i.e., delta (1–3 Hz), theta (3–7 Hz), alpha (7–13 Hz), beta (13–30 Hz), and gamma (30–50 Hz). These subjects have been used in our previous studies for studying topography of synchronization maps and synchronizability of AD brains [9,17,24,35].

2.2. Constructing brain functional networks

In order to study network properties of functional brain networks, the first step is to obtain the connectivity matrices showing inter-relations between brain regions. There are a number of methods in order to obtain the connectivity matrices (see a recent review of connectivity estimation methods in [17]); here we use Pearson correlation coefficient between the

time series, which has been frequently used for this purpose. The correlation coefficient between sensors i and j can be obtained as

$$r_{ij} = \frac{\mathrm{cov}(i, j)}{\sqrt{\mathrm{var}(i)\,\mathrm{var}(j)}}, \tag{1}$$

where cov(i,j) is the covariance between nodes i and j, and var(i) is the variance of node i.

Often, binary networks are constructed from the weighted connectivity matrices to reduce the noise level. This can be achieved by thresholding the correlation matrices; there is a link between two nodes if their correlation weight is larger than a certain threshold. A good strategy to threshold the weighted correlation matrices is to threshold them for a certain density value [1]. In this method, known as sparsity thresholding, the extracted networks will all have the same density (i.e., number of edges), and thus the comparison will be unbiased to the density (which has significant influence on many network properties). Here we fix the density value at 0.05 to be consistent with real cortical networks that have been shown to have connection densities in this range.

2.3. Graph theoretical metrics

Here we study a number of neurobiologically relevant static and dynamical network measures. These measures correspond to integration/segregation phenomena, communicability of brain regions, hierarchal structure and synchronization. Having a binary undirected graph, the simplest measure is node degree, which is the number of links the node receives. Here we use the maximum degree and the standard deviation of nodes' degree as indicators of heterogeneity in the degree distribution. Segregation of information is one of the distinct abilities of the brain. This means that the brain processes information in a specialized manner and each region is responsible for processing special kind of information. In this work, we consider two graph theory metrics measuring segregation properties of a network: local efficiency and transitivity. These measures mainly take into account local connections in the network. Local efficiency of node i is computed as

$$LE_i = \frac{1}{d_i(d_i-1)} \sum_{j=G_i} \frac{1}{l_{i,j}}, \tag{2}$$

where d_i is degree of node i (the number of nodes connected to node i), l_{ij} is the length of the shortest path between nodes i and j and G_i is the graph of neighbours of nodes i excluding node i. The local efficiency of the network is obtained by making average over all the nodes, as

$$LE = \frac{1}{N} LE_i, \tag{3}$$

where N is the total number of nodes. Transitivity (or clustering coefficient) of a network measures to how much extent neighbours of two connected nodes are interconnected. It can be computed by counting all triangles in the network and comparing it with all triple counts, as

$$T = \frac{1}{N} \sum_k \frac{\sum_{i,j} a_{ij} a_{ik} a_{jk}}{d_k(d_k-1)}, \tag{4}$$

where a_{ij} is the (i,j) entry of the binary connection matrix; $a_{ij} = a_{ji} = 1$ if there is a link between nodes i and j, and $a_{ij} = a_{ji} = 0$, otherwise. There are no self-loops in the network, and thus $a_{ii} = 0$.

Other prominent feature of the brain is its ability in integration of information separately processed in different brain regions. To do this, the brain network should provide efficient communication between its regions. Here we consider three metrics that quantify communication efficiency in networks: global efficiency, edge betweenness centrality and node betweenness centrality. Global efficiency of a network is inversely proportional to its average shortest path length and is defined as

$$GE = \frac{1}{N(N-1)} \sum_{i,j} \frac{1}{l_{i,j}}. \tag{5}$$

In order to take into account the significance (centrality) of nodes and edges in the brain networks, their betweenness centrality measures can be considered. Let's denote the edge between nodes i and j by e_{ij}. Edge-betweenness centrality EBC_{ij} of the network is defined by

$$EBC_{ij} = \sum_{p \neq u} \frac{\Gamma_{pu}(e_{ij})}{\Gamma_{pu}}, \tag{6}$$

where Γ_{pu} is the number of shortest paths between nodes p and u in the graph and $\Gamma_{pu}(e_{ij})$ is the number of these shortest paths making use of the edge e_{ij}. Node-betweenness centrality NBC_i is a centrality measure of node i in a graph, showing

the number of shortest paths making use of node i (except those between the ith node with the other nodes). One can compute it as

$$NBC_i = \sum_{j \neq i \neq k} \frac{\Gamma_{jk}(i)}{\Gamma_{jk}}, \tag{7}$$

where Γ_{jk} is the number of shortest paths between nodes j and k and $\Gamma_{jk}(i)$ is the number of these shortest paths making use of the node i.

We also obtain the modularity index that reveals a hierarchal structure of a network. In order to capture the degree of modularity in a network composed of M modules, the following index has been proposed

$$Q = \sum_{i \in M} \left[q_{ii} - \left(\sum_{j \in M} q_{ij} \right)^2 \right], \tag{8}$$

where the network is fully partitioned into M non-overlapping modules (clusters), and q_{ij} represents the proportion of all links connecting nodes in module i with those in module j. The modularity index is computed by estimating the optimal modular structure for a given network.

Networks may undergo random and/or intentional failures in their components, and resiliency against such failures is of high importance for their proper functioning. Degree-degree correlation has a significant role in determining resiliency of complex networks, which can be quantified by calculating the assortativity measure, as defined by

$$r = \frac{\frac{1}{E}\sum_{j>i} d_i d_j a_{ij} - \left[\frac{1}{E} \sum_{j>i} \frac{1}{2}(d_i + d_j) a_{ij} \right]^2}{\frac{1}{E}\sum_{j>i} \frac{1}{2}(d_i^2 + d_j^2) a_{ij} - \left[\frac{1}{E} \sum_{j>i} \frac{1}{2}(d_i + d_j) a_{ij} \right]^2}, \tag{9}$$

where E is the total number of the edges in the network. If $r > 0$, the network is assortative, whereas $r < 0$ indicates a disassortative network. For $r = 0$ there is no correlation between the node-degrees. In assortative networks, high-degree nodes often tend to interconnect, whereas in disassortative networks, nodes with high degree tend to connect to low-degree nodes. Assortative networks are likely to consist of mutually coupled hub nodes with high degrees and to be resilient against random failures. In contrast, disassortative networks are likely to have vulnerable high-degree nodes.

The above metrics measure static properties of networks. However, the brain is composed of dynamical units and phenomena like synchronization is often linked to its function. Brain disorders such as schizophrenia, AD, epilepsy and Parkinson's disease have been linked to abnormal synchronization levels [19,24,25]. The eigenratio of the Laplacian matrix is often used as a synchronizability index [35], which is the ease by which synchronization can be achieved. Having the binary adjacency matrix $A = (a_{ij})$, the Laplacian matrix is defined as $S = (s_{ij})$, where $s_{ij} = -a_{ij}$ for $i \neq j$ and $s_{ii} = d_i$. S is a symmetric zero-row sum matrix, and thus all its eigenvalues are non-negative real values. Let's denote the eigenvalues of the Laplacian matrix S as $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_N$. The eigenratio, as a measure of network synchronizability is obtained as

$$R = \frac{\lambda_N}{\lambda_2}, \tag{10}$$

and networks with smaller R are better synchronizable.

2.4. Statistical assessments

In order to assess whether AD and control groups have statistically significant network properties, non-parametric Wilcoxon's ranksum test is used. The results with $P < 0.05$ are considered to be statistically significant. All computations were performed in MatLab.

2.5. Discrimination of AD patients from healthy controls based on graph measures

We take the above graph theory measures (assortativity, edge betweenness centrality, node betweenness centrality, global efficiency, local efficiency, modularity index, transitivity, maximum degree, standard deviation of node degrees, and synchronizability index) as features for classifying AD and healthy subjects. The classification task is a 2-class problem (AD group and healthy controls) with 10 features. We use a number of techniques in order to select a subset of features that result in the best performance, i.e., the least misclassification rate [13,20]. To this end, techniques based on genetic Algorithms (GA) [14], Binary Particle Swarm Optimization (BPSO) [23] and Social Impact Theory based Optimization (SITO) [29] are used. In the following we provide brief explanations of these techniques.

GA [14] is a stochastic optimization method that has been frequently used for optimizing non-convex problems and has good performance on large spaces. In this method, the candidate solutions are encoded in a binary solution that is called chromosome. The algorithm uses a number of operations on the chromosomes (selection, crossover, mutation and

replacement) to find a good solution. PSO performs the search using entities called particles [22], and each solution is represented as a particle's position vector ($\boldsymbol{x}_i$). In each iteration t of the optimization process, a velocity vector $\boldsymbol{v}_i$ is computed for each particle i, as

$$\boldsymbol{v}_i(t+1) = \omega(t)\boldsymbol{v}_i(t) + c_1 u(0,1)(\boldsymbol{p}_i(t) - \boldsymbol{x}_i(t)) + c_2 U(0,1)(\boldsymbol{g}(t) - \boldsymbol{x}_i(t)), \tag{11}$$

where $\boldsymbol{v}_i(t)$ is the current velocity vector, $\omega(t)$ is the inertia weight, $\boldsymbol{p}_i(t)$ is the best solution found by particle i up to iteration t, $\boldsymbol{g}(t)$ is the best solution found by all particles up to time t, and c_1 and c_2 are the weight factors for the personal best and the global best solutions, respectively. $U(a,b)$ indicate a value picked from a uniform distribution between a an b. The inertia weight constantly decreases from $\omega_{\max}$ to $\omega_{\min}$ using the following equation

$$\omega(t) = \frac{\omega_{\max} - \omega_{\min}}{t_{\max}} t, \tag{12}$$

where $t_{\max}$ is the maximum number of iterations. For the feature selection problem, each candidate solution is encoded as a position of a particle in a multidimensional binary space, where 1 s (0 s) corresponds to selection (elimination) of the corresponding features in the dataset.

SITO, as a recently developed population-based optimization technique, is inspired by opinion formation process in social networks [29] and has been used for the feature selection task [28,29]. In SITO, each agent is initialized with a random binary opinion vector representing to a candidate solution. There is a graph connecting the agents (often chosen as a random graph), and the neighbours of agent i are divided into two categories: persuasive (PER_i) and supportive (SUP_i). Persuasive agents have the same opinion as agent i, whereas that of supportive agents is opposite. In the opinion formation phase, the impact value I_i is computed as

$$I_i(t) = \frac{1}{|PER_i|} \sum_{j \in PER_i} f_{ij} - \frac{1}{|SUP_i|} \sum_{j \in SUP_i^d} f_{ij}, \tag{13}$$

where f_{ij} is the impact strength of agent j on agent i that is obtained as

$$f_{ij} = \max(h(\boldsymbol{x}_i) - h(\boldsymbol{x}_j), 0). \tag{14}$$

and $f(.)$ is the cost function to be minimized (the misclassification rate here) and $\boldsymbol{x}_i$ is the opinion value of agent i. $\boldsymbol{x}_i$ is then updated as

$$\boldsymbol{x}_i(t+1) = \begin{cases} 1 - \boldsymbol{x}_i(t), & if\ I_i > 0 \\ \boldsymbol{x}_i(t), & otherwise \end{cases}. \tag{15}$$

In order to perform the classification task, three well-known classifiers are used: Support Vector Machines (SVM), Naïve Bayes and k Nearest Neighbour (NN) [7].

3. Results and discussion

Figs. 1–8 compare AD and healthy control subjects in terms of different graph theory metrics. The networks show stable properties across the frequency bands. AD brains show significantly (almost 10%) decreased local efficiency in the eyes-closed condition, but the decrease in the eyes-open condition is not significant (Fig. 1). The changes in the transitivity profile are not statistically significant in either of the conditions. Although the global connectivity measures (global efficiency, edge betweenness centrality and node betweenness centrality) are increased in AD brains as compared to healthy brains, the changes are not statistically significant. EEG-based functional networks of AD show decreased modularity than healthy control subjects in the eyes-closed condition (Fig. 6). The modularity index takes into account community structure in networks, and thus is more influenced by local connections than long-range ones. A community is defined as a subset of nodes with dense intra-community connections and sparse inter-community links. Our results indicate that AD brains have stronger community structure in the eyes-closed condition. The groups are not significantly different in terms of the assortativity index (Fig. 7) or the synchronizability (Fig. 8).

We use three feature selection algorithms (GA, BPSO and SITO) along with three well-known classifiers (3NN, Naïve Bayes and SVM) to minimize (maximize) the misclassification rate (classification accuracy). The population size for the feature selection algorithms are fixed at 10 and binary random vectors are considered as initial solution vectors. The mutation operator, which is part of all three feature selection methods, is chosen as

$$\rho(t) = \frac{\rho_{\max} - \rho_{\min}}{t_{\max}} t, \tag{16}$$

where $\rho(t)$ is the mutation rate at iteration t, $t_{\max}$ is the maximum number of iterations, and $\rho_{\max} = 3/10$ (there are 10 features for the classification task) and $\rho_{\min} = 0$.

A 2-point crossover is used for GA and Roulette-Wheel method is used to implement the selection operator. These methods have been previously shown to perform well in the feature selections tasks [11,15,31]. BPSO is initialized with 10 particles, random initial positions and velocities, and with $V_{min} = -6$, $V_{max} = +6$, $c_1 = 2.0$, $c_2 = 1.5$, $\omega_{min} = 0.5$ and $\omega_{max} = 0.995$, as suggested in [37] for the feature selection task. We use random topology as a connection graph for SITO and the parameters

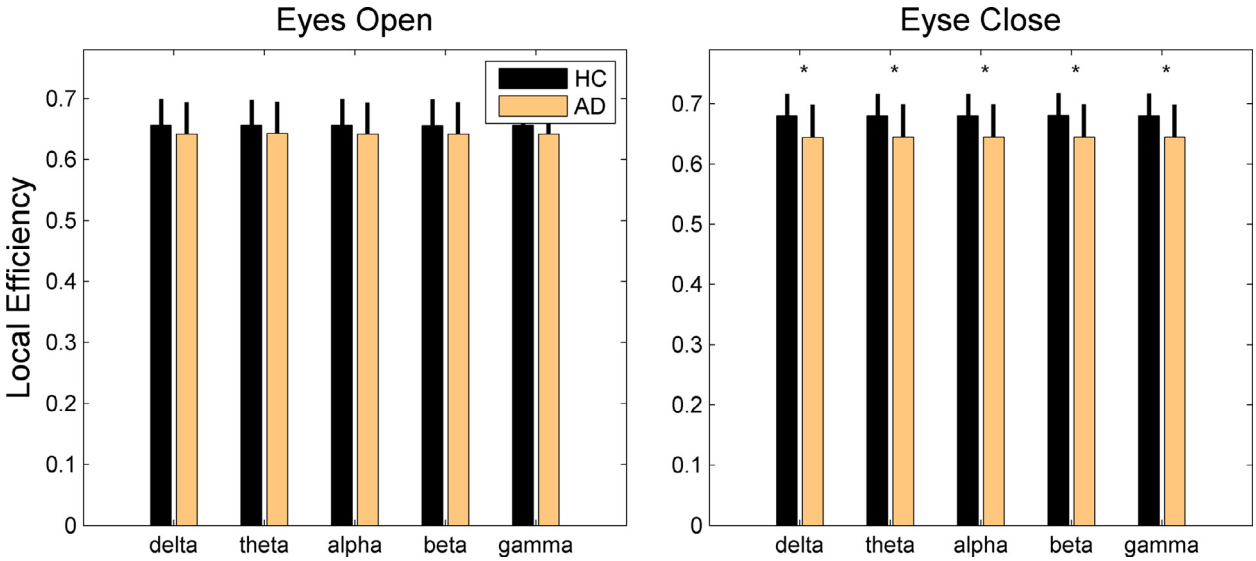

Fig. 1. Local efficiency of EEG-based functional networks in healthy controls (HC) and patients suffering from Alzheimer's disease (AD) in different frequency bands including delta (1–3 Hz), theta (3–7 Hz), alpha (7–13 Hz), beta (13–30 Hz) and gamma (30–50 Hz) bands. The EEGs have been recorded in two conditions: resting state with eyes open and eyes close. Graphs show mean values along with the standard deviations. Asterisk above the bars shows the cases for which the two groups (HC and AD) are significantly different ($P < 0.05$; Wilcoxon's ranksum test).

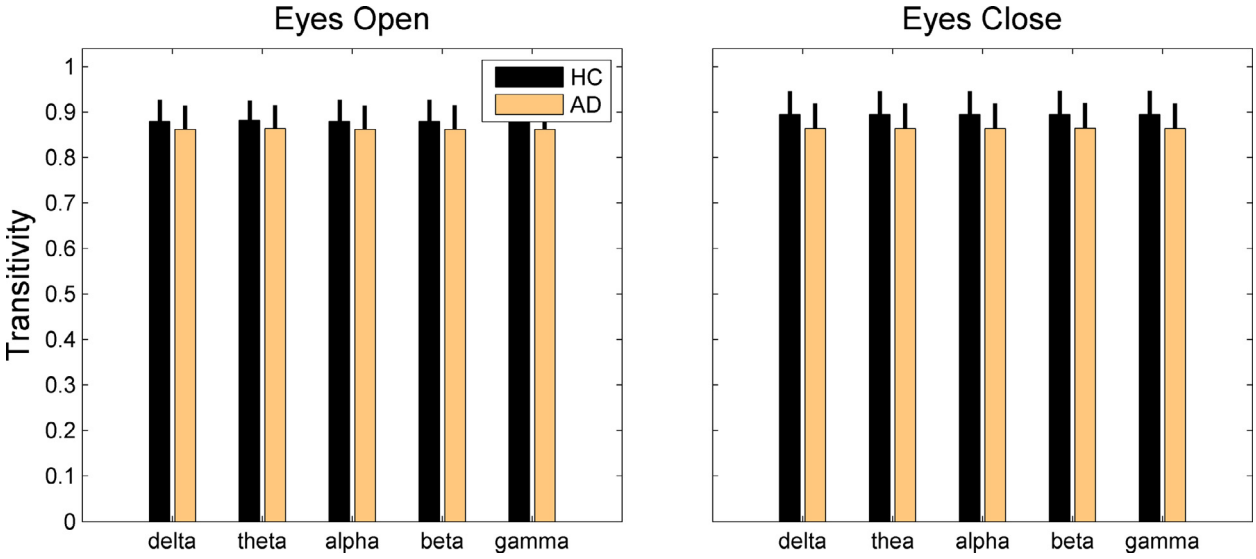

Fig. 2. Transitivity of networks in HC and AD groups. Other designations are as Fig. 1.

are set as suggested in [28,29]. Since the data has rather small sample size (23 samples in AD group and 25 samples in control group), we use leave-one-out cross-validation algorithm. In this method, one of them is taken as a test data (with unknown class), and the classifier is trained with the remaining data (47 samples). Then, the trained model is verified on the test sample, and if its class is incorrectly predicted, a misclassification is counted. This procedure is repeated until all samples are taken as sample data.

We assess the performance of the feature selection methods using 3NN, Naïve Bayes and SVM, and the results are shown in Table 2. Note that a random classification results in 50% misclassification rate for a 2-class classification problem. These results show that GA performs better than BPSO and SITO in all classifiers and recording conditions. Furthermore, SVM is the top-performer classifier. The network features in the eyes-open condition result in worse classification rates than the eyes-closed cases (data not shown here), which is expected since the groups do not show significant differences in network

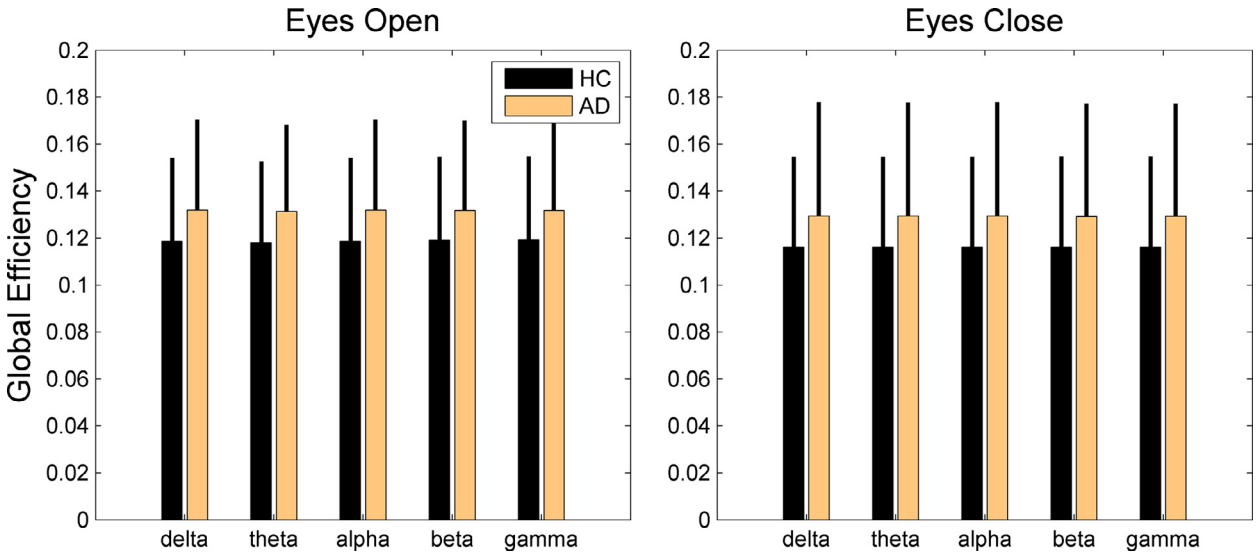

Fig. 3. Global efficiency of networks in HC and AD groups. Other designations are as Fig. 1.

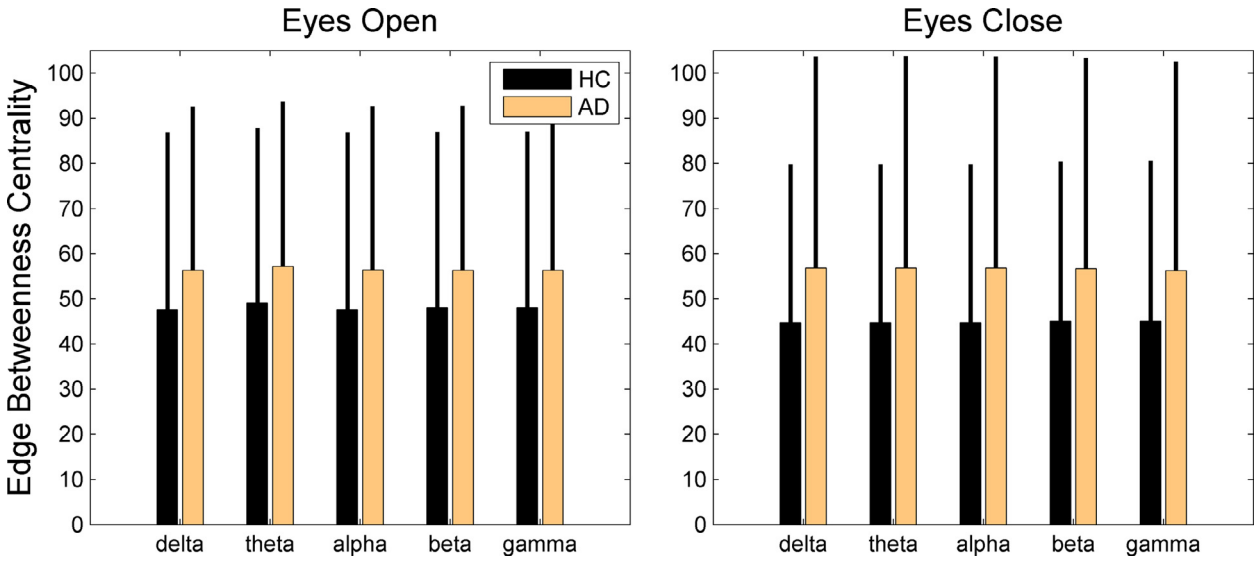

Fig. 4. Edge betweenness centrality of networks in HC and AD groups. Other designations are as Fig. 1.

features for the eyes-open condition. We also study the optimal set of graph theory features found by the top-performer method (GA as feature selector and SVM as classifier). Table 3 shows these features for different frequency bands. It is seen that some of the features (edge betweenness centrality, global efficiency and synchronizability) are among the optimal set in all frequency bands, indicating that these features can be considered as graph theory marker of AD.

We next apply a brute-force search strategy to find the best feature set. Table 4 shows the minimum misclassification rate using brute-force feature selection strategy in the eyes-closed condition. This analysis shows that SVM applied on the eyes-closed data in beta band results in the best classification rate, where 83% of the subjects are correctly classified. In this case, the optimal feature set includes edge betweenness centrality, global efficiency, modularity and synchronizability.

In sum, our analysis show that resting state EEGs of AD patients have abnormal graph theoretical properties, which is more pronounced in the eyes-closed condition. The features extracted from network analysis can be used to discriminate AD from healthy controls with medium classification performance. In order to confirm these results, the analysis should be repeated with larger sample sizes and other neuroimaging modalities such as EMG and fMRI.

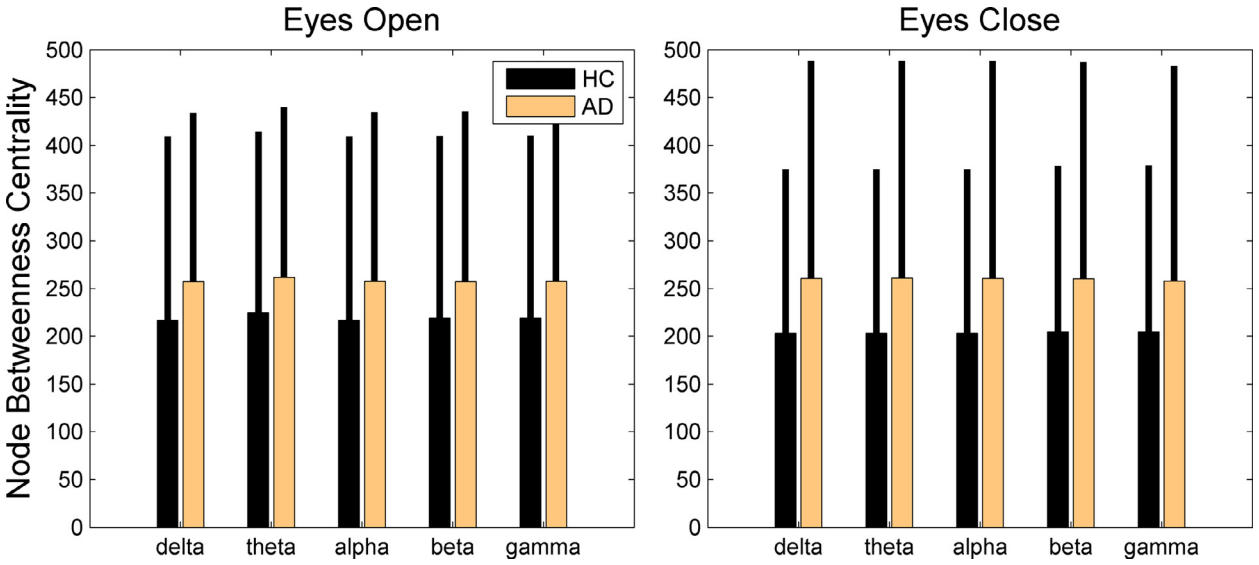

Fig. 5. Node betweenness centrality of networks in HC and AD groups. Other designations are as Fig. 1.

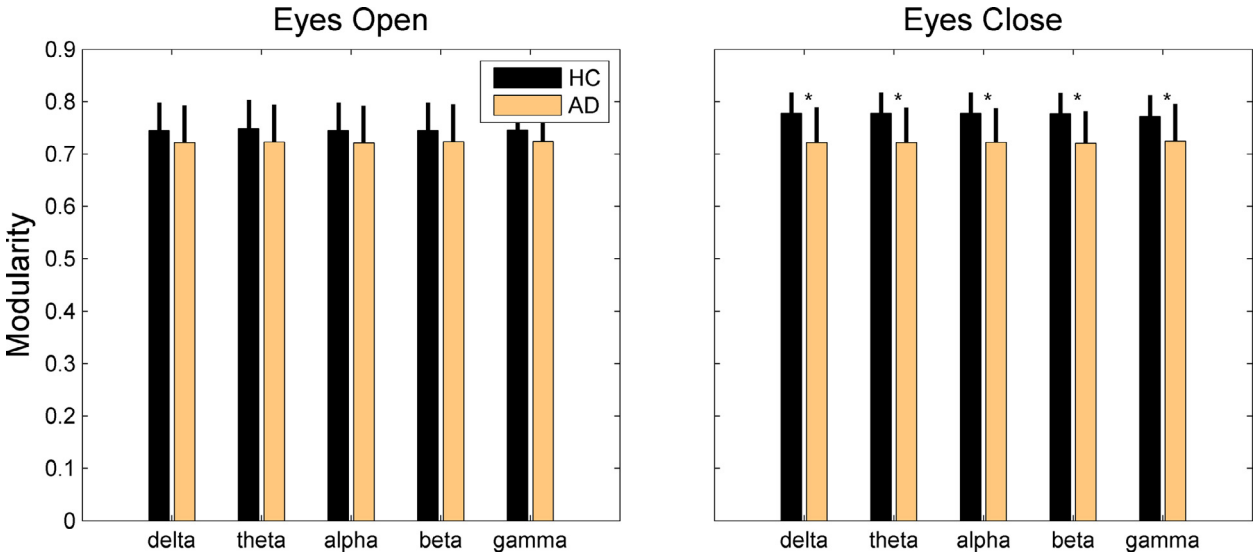

Fig. 6. Network modularity in HC and AD groups. Other designations are as Fig. 1.

Table 2
Misclassification rate (mean value and standard deviation over 10 independent runs each with 50 iterations) for eyes-closed condition. Three feature selection methods (binary particle swarm optimization (BPSO), genetic algorithms (GA) and social impact theory based optimization (SITO)) and two classifiers (3NN and Naïve Bayes). The bold cases indicate the classifiers with the best performance.

Band	BPSO			GA			SITO			All Features		
	3NN	Naïve Bayes	SVM	3NN	Naïve Bayes	SVM	3NN	Naïve Bayes	SVM	3NN	Naïve Bayes	SVM
delta	0.39 ± 0.04	0.27 ± 0.03	0.22 ± 0.01	0.31 ± 0.01	0.24 ± 0.00	$\mathbf{0.21 \pm 0.00}$	0.35 ± 0.02	0.26 ± 0.02	0.25 ± 0.01	0.42 ± 0.02	0.34 ± 0.02	0.31 ± 0.03
theta	0.40 ± 0.04	0.27 ± 0.03	0.23 ± 0.01	0.34 ± 0.00	0.24 ± 0.01	$\mathbf{0.24 \pm 0.01}$	0.36 ± 0.02	0.27 ± 0.02	0.25 ± 0.02	0.43 ± 0.01	0.33 ± 0.03	0.30 ± 0.03
alpha	0.38 ± 0.04	0.27 ± 0.03	0.24 ± 0.01	0.34 ± 0.00	0.24 ± 0.00	$\mathbf{0.21 \pm 0.00}$	0.38 ± 0.03	0.27 ± 0.03	0.25 ± 0.01	0.41 ± 0.03	0.32 ± 0.02	0.31 ± 0.02
beta	0.40 ± 0.03	0.28 ± 0.04	0.26 ± 0.01	0.34 ± 0.00	0.24 ± 0.00	$\mathbf{0.17 \pm 0.00}$	0.35 ± 0.01	0.25 ± 0.02	0.25 ± 0.01	0.43 ± 0.02	0.33 ± 0.02	0.31 ± 0.02
gamma	0.39 ± 0.04	0.28 ± 0.04	0.26 ± 0.02	0.33 ± 0.02	0.24 ± 0.00	$\mathbf{0.21 \pm 0.00}$	0.39 ± 0.05	0.27 ± 0.03	0.26 ± 0.01	0.40 ± 0.02	0.34 ± 0.03	0.32 ± 0.02

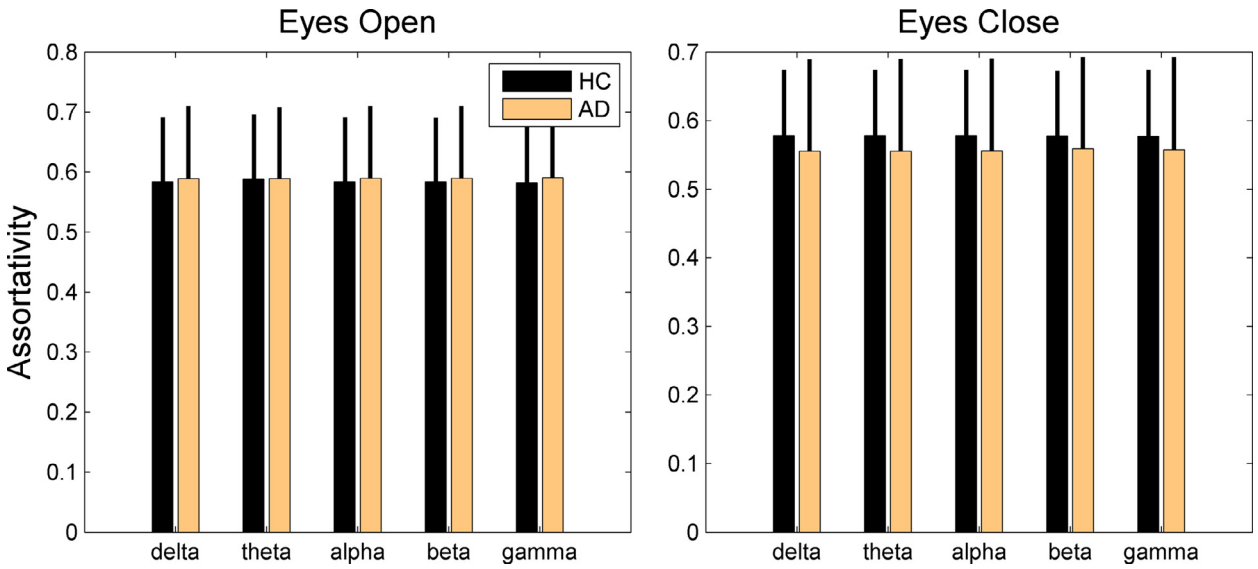

Fig. 7. Assortativity of EEG-based functional networks in HC and AD groups. Other designations are as Fig. 1.

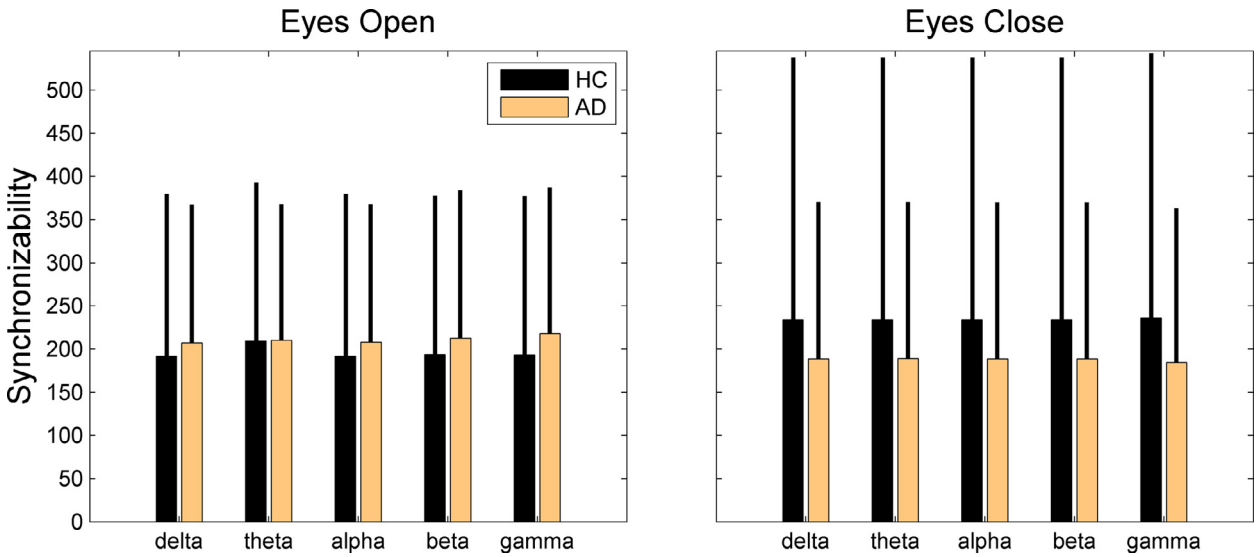

Fig. 8. Network synchronizability in HC and AD groups. Other designations are as Fig. 1.

Table 3
The optimal set of features that are selected by GA as feature selection method and SVM as classifier. EBC: edge betweenness centrality, GE: global efficiency, LE: local efficiency, Q: modularity, d_{max}: maximum degree, R: synchronizability.

Band	Selected features
delta	EBC, GE, LE, Q, d_{max}, R
theta	EBC, GE, R
alpha	EBC, GE, Q, d_{max}, R
beta	EBC, GE, Q, R
gamma	EBC, GE, d_{max}, R

Table 4
Misclassification rate for eyes-closed condition. Brute-force search was used to determine the set of best-performing features by examining all $210 = 1024$ cases. Bold values show the best discrimination performance in each frequency band.

Band	3NN	Naïve Bayes	SVM
delta	0.31 ± 0.01	0.24 ± 0.00	$\mathbf{0.21 \pm 0.00}$
theta	0.34 ± 0.03	0.24 ± 0.01	$\mathbf{0.24 \pm 0.01}$
alpha	0.34 ± 0.02	0.24 ± 0.00	$\mathbf{0.21 \pm 0.00}$
beta	0.34 ± 0.03	0.24 ± 0.00	$\mathbf{0.17 \pm 0.00}$
gamma	0.31 ± 0.02	0.24 ± 0.00	$\mathbf{0.21 \pm 0.00}$

4. Conclusion

The human brain can be modelled as a complex network where the nodes represent distinct brain regions and the links represent functional/anatomical connectivity between them. To construct brain networks, one can use brain neuroimaging modalities such as EEG, EMG, fMRI and DTI. In this work we considered EEGs recorded from 23 AD patients and 25 healthy control subjects in two conditions: resting-state eyes-open and eyes-closed. The network structures were extracted by applying Pearson correlation coefficient to the filtered EEG time series and sparsity thresholding method to binarize the connectivity matrices. A number of network metrics were obtained including maximum degree, standard deviation of degree distribution, local efficiency, transitivity, global efficiency, edge betweenness centrality, nodes betweenness centrality, modularity, assortativity and synchronizability. Our analysis showed that AD networks have significantly reduced local efficiency and modularity in eyes-closed condition. The network properties also showed stable profile in the network measures across all frequency bands. We used the network metrics as features to discriminate AD patients from healthy subjects. To this end, three feature selection methods (GA, BPSO and SITO) were used along with three well-known classifiers. We also applied a brute-force search in the feature space and found the best set of features for this 2-class classification problem. Using the optimal set of features, SVM could correctly classify 83% of the subjects in beta frequency band of eyes-closed condition. The optimal set of features included edge betweenness centrality, global efficiency, modularity and synchronizability.

Acknowledgments

The author would like to thank Dr. Maria G. Knyazeva for EEG recording and preprocessing, and Homayoun Hamed Moghaddam Rafati for his help in implementing the feature selection and classification algorithms. This research was partially by Australian Research Council through grant number DE140100620.

References

[1] S. Achard, E. Bullmore, Efficiency and cost of economical brain functional networks, PLoS Comput. Biol. 3 (2007) 174–183.
[2] S. Achard, R. Salvador, B. Whitcher, J. Suckling, E. Bullmore, A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs, J. Neurosci. 26 (2006) 63–72.
[3] G. Adler, S. Brassen, A. Jajcevic, EEG coherence in Alzheimer's dementia, J. Neural Transm. 110 (2003) 1051–1058.
[4] S. Afshari, M. Jalili, Directed functional networks in Alzheimer's disease: disruption of global and local connectivity measures, IEEE J. Biomed. Health Inform. (2016).
[5] E. Barzegaran, B. van Damme, R. Meuli, M.G. Knyazeva, Perception-related EEG is more sensitive to Alzheimer's disease effects than resting EEG, Neurobiol. Aging 43 (2016) 129–139.
[6] B.C. Bernhardt, L. Bonilha, D.W. Gross, Network analysis for a network disorder: the emerging role of graph theory in the study of epilepsy, Epilepsy Behav. 50 (2015) 162–170.
[7] C.M. Bishop, Pattern Recognition and Machine Learning, first ed., Springer, Singapore, 2006.
[8] E. Bullmore, O. Sporns, The economy of brain network organization, Nat. Rev. Neurosci. (2012).
[9] C. Carmeli, E. Fornari, M. Jalili, R. Meuli, M.G. Knyazeva, Structural covariance of superficial white matter in mild Alzheimer's disease compared to normal aging, Brain Behav. 4 (2014) 721–737.
[10] W. de Haan, Y.A.L. Pijnenburg, R.L.M. Strijers, Y. van der Made, W.M. van der Flier, P. Scheltens, C.J. Stam, Functional neural network analysis in frontotemporal dementia and Alzheimer's disease using EEG and graph theory, BMC Neurosci. 10 (2009) 101-101.
[11] H.A. Firpi, E. Goodman, Swarmed Feature Selection, in: Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop, 2004, pp. 112–118.
[12] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C.J. Honey, V.J. Wedeen, O. Sporns, Mapping the structural core of human cerebral cortex, PLoS Biol. 6 (2008) e169.
[13] L. Hedjazi, J. Aguilar-Martin, M.-V. Le Lannb, T. Kempowsky-Hamonb, Membership-margin based feature selection for mixed type and high-dimensional data: theory and applications, Inf. Sci. 322 (2015) 174–196.
[14] J.H. Holland, Genetic algorithms, Sci. Am. 267 (1992) 66–72.
[15] F. Hussein, N. Kharma, R. Ward, Genetic algorithms for feature selection and weighting, a review and study, in: Proceedings of Sixth International Conference on Document Analysis and Recognition, 2001, pp. 1240–1244.
[16] M. Jalili, Resiliency of EEG-based brain functional networks, PLoS One 10 (2015) e0135333.
[17] M. Jalili, Functional brain networks: does the choice of dependency estimator and binarization method matter? Sci. Rep. 6 (2016) 29780.
[18] M. Jalili, M.G. Knyazeva, EEG brain functional networks in schizophrenia, Comput. Biol. Med. 41 (2011) 1178–1186.
[19] M. Jalili, S. Lavoie, P. Deppen, R. Meuli, K.Q. Do, M. Cuenod, M. Hasler, O. De Feo, M.G. Knyazeva, Dysconnection topography in schizophrenia with state-space analysis of EEG, PLoS One 2 (2007) e1059.
[20] R. Jensen, N.M. Parthaláin, Towards scalable fuzzy-rough feature selection, Inf. Sci. 323 (2015) 1–15.

[21] B. Jie, D. Zhang, W. Gao, Q. Wang, C.-Y. Wee, D. Shen, Integration of network topological and connectivity properties for neuroimaging classification, IEEE. Trans. Biomed. Eng. 61 (2014) 576–589.
[22] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proceedings of IEEE International Conference on Neural Networks, 1995, pp. 1942–1948.
[23] J. Kennedy, R.C. Eberhart, A discrete binary version of the particle swarm algorithm, in: Proceedings of IEEE International Conference on Computational Cybernetics and Simulation, 1997, pp. 4104–4108.
[24] M.G. Knyazeva, M. Jalili, A. Brioschi, I. Bourquin, E. Fornari, M. Hasler, R. Meuli, P. Maeder, J. Ghika, Topography of EEG multivariate phase synchronization in early Alzheimer's disease, Neurobiol. Aging 31 (2010) 1132–1144.
[25] M.G. Knyazeva, M. Jalili, R.S. Frackowiak, A.O. Rossetti, Psychogenic seizures and frontal disconnection: EEG synchronisation study, J. Neurol. Neurosurg. Psychiatry 82 (2011) 505–511.
[26] S. Lemm, B. Blankertz, G. Curio, K. Muller, Spatio-spectral filters for improving the classification of single trial EEG, IEEE. Trans. Biomed. Eng. 52 (2005) 1541–1548.
[27] F. Li, Z. Zhang, C. Jin, Feature selection with partition differentiation entropy for large-scale data sets, Inf. Sci. 329 (2016) 690–700.
[28] M. Macaš, Opinion Formation Inspired Search Strategies for Feature Selection, Czech Technical University, 2012.
[29] M. Macaš, Binary social impact theory based optimization and its applications in pattern recognition, Neurocomputing 132 (2014) 85–96.
[30] P. Moradi, M. Gholampour, A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy, Appl. Soft Comput. 43 (2016) 117–130.
[31] I.-S. Oh, J.-S. Lee, B.-R. Moon, Hybrid genetic algorithms for feature selection, IEEE Trans. Pattern Anal. Mach. Intell. 26 (2004) 1424–1437.
[32] C. Pedreiraa, A.E. Vaudano, R.C. Thornton, U.J. Chaudhary, S. Vulliemoz, H. Laufs, R. Rodionov, D.W. Carmichael, S.D. Lhatoog, M. Guye, R.Q. Quiroga, L. Lemieux, Classification of EEG abnormalities in partial epilepsy with simultaneous EEG-fMRI recordings, NeuroImage 99 (2014) 461–476.
[33] E.J. Sanz-Arigita, M.M. Schoonheim, J.S. Damoiseaux, S.A. Rombouts, E. Maris, F. Barkhof, P. Scheltens, C.J. Stam, Loss of 'small-world' networks in Alzheimer's disease: graph analysis of FMRI resting-state functional connectivity, PLoS One 5 (2010) e13788.
[34] C.J. Stam, B.F. Jones, G. Nolte, M. Breakspear, P. Scheltens, Small-world networks and functional connectivity in Alzheimer's disease, Cereb. Cortex. 17 (2007) 92–99.
[35] M.S. Tahaei, M. Jalili, M.G. Knyazeva, Synchronizability of EEG-based functional networks in early Alzheimer's disease, IEEE Trans. Neural Syst. Rehabil. Eng. 20 (2012) 636–641.
[36] L.J. Trejo, K. Kubitz, R. Rosipal, R.L. Kochavi, L.D. Montgomery, EEG-based estimation and classification of mental fatigue, Psychology 6 (2015) 572–589.
[37] A. Unler, A. Murat, A discrete particle swarm optimization method for feature selection in binary classification problems, Eur. J. Oper. Res. 206 (2010) 528–539.
[38] J. Wang, X. Zuo, Z. Dai, M. Xia, Z. Zhao, X. Zhao, J. Jia, Y. Han, Y. He, Disrupted functional brain connectome in individuals at risk for Alzheimer's disease, Biol. Psychiatry 73 (2013) 472–481.
[39] M. Zanin, P. Sousa, D. Papo, R. Bajo, J. Garcıa-Prieto, F. del Pozo, E. Menasalvas, S. Boccaletti, Optimizing functional network representation of multivariate time series, Sci. Rep. 2 (2012) 630.
[40] X. Zhao, Y. Liu, X. Wang, B. Liu, Q. Xi, Q. Guo, H. Jiang, T. Jiang, P. Wang, Disrupted small-world brain networks in moderate Alzheimer's disease: a resting-state fMRI study, PLoS One 7 (2012) e33540.

Mahdi Jalili received his B.S. degree in electrical engineering from Tehran Polytechnique in 2001, his M.S. degree in electrical engineering from the University of Tehran in 2004, and his PhD from Swiss Federal Institute of Technology Lausanne (EPFL) in 2008. He then joined Sharif University of Technology as assistant professor. He is now with the School of Electrical and Computer Engineering, RMIT University, Melbourne, Australia and holds Australian Research Council DECRA (Discovery Early Carrier Research Award) Fellowship and RMIT's Vice-Chancellor Research Fellowship. His research interests are in network science, dynamical systems, social networks analysis and mining, and human brain functional connectivity analysis. Dr. Jalili is a senior member of IEEE and associate editor of IEEE Canadian Journal of Electrical and Computer Engineering and an editorial board member of Complex Adaptive Systems Modelling.