Neurocomputing 170 (2015) 32–46

Contents lists available at ScienceDirect

Neurocomputing

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

A heuristic branch-and-bound based thresholding algorithm for unveiling cognitive activity from EEG data

M. Thilaga [a,*], R. Vijayalakshmi [a], R. Nadarajan [a], D. Nandagopal [b], B. Cocks [b], C. Archana [a], N. Dahal [b]

[a] Department of Applied Mathematics and Computational Sciences, PSG College of Technology, Tamil Nadu, India
[b] Division of Information Technology, Engineering and the Environment, University of South Australia, Australia

ARTICLE INFO

Article history:
Received 10 October 2014
Received in revised form
7 February 2015
Accepted 3 March 2015
Available online 10 July 2015

Keywords:
Cognition
Branch-and-bound
Electroencephalography
Functional brain networks
Graph theory
Thresholding

ABSTRACT

One of the biggest challenges in the field of computational neuroscience from the perspective of complex network analysis is the measurement of dynamic local and global interactions of the brain regions during cognitive function. Graph theoretic analysis has been extensively applied to study the dynamics of functional brain networks in the recent years. The selection of appropriate thresholding methods to construct weighted/unweighted subnetworks to detect cognitive load induced changes in brain's electrical activity remains an open challenge in the functional brain network research. This paper reviews the application of statistical and information theoretic metrics to construct the functional brain networks, proposes a novel Branch-and-Bound based thresholding algorithm that extracts the influential subnetwork, and applies efficient computational techniques and complex network metrics to detect and quantify the cognitive activities. The empirical analyses showcase the efficiency of the proposed thresholding algorithm by highlighting the changing neuronal patterns during cognitive activity when compared to that of baseline activity. Statistical evaluation of the results further validates the efficiency of the proposed method as well. The results demonstrate the ability of the proposed algorithm in detecting subtle cognitive load induced changes in functional brain networks.

1. Introduction

Graph theory, a branch of discrete mathematics, offers a broad range of theoretical tools that are used to model, describe and quantify the topologically and functionally significant patterns in a network. Network research in the recent past has realized that the complex network analysis is a major theme due to its widespread application in areas such as traffic networks, World Wide Web, social networks, communication networks, and biological networks [1–5]. As a result, there is an increasing interest in developing novel methods and measures that describe various characteristics of complex networks such as the structure of the network, patterns of structural/functional interactions, cohesive subgroups, and quantification of individual node's significance. Since human brain is one of the most complex and adaptive systems with spatially distributed but functionally connected regions, all these characteristics are explored by many researchers.

The functional connectivity of the brain network is indirectly determined by billions of neurons constituting the anatomical brain structure. From the network perspective, the Functional Brain Network (FBN) is characterized as a graph/network where neuronal populations, brain regions or cortical areas are modeled as nodes (vertices) and axons, synapses, fiber tracts or statistical or causal relationships that describe functional association as edges (links) [1,6]. Extraction of prominent features from the functional brain connectivity patterns to identify relationships among the various brain regions remains a challenging issue. Over the past decades, discovering significant features from large spatiotemporal data produced by different neuroimaging techniques such as the functional magnetic resonance imaging (fMRI), structural Magnetic Resonance Imaging (MRI), Diffusion Tensor Imaging (DTI), and neurophysiological recordings such as ElectroEncephaloGraphy (EEG) and MagnetoEncephaloGraphy (MEG) have inspired many researchers [7–10].

Many of the existing studies on human FBNs focus on analyzing the disorders of the human brain using various techniques [11]. To

* Corresponding author.
E-mail addresses: thilagaselvan@gmail.com (M. Thilaga), rvpsgtech@gmail.com (R. Vijayalakshmi), rn@amc.psgtech.ac.in (R. Nadarajan), nanda.nandagopal@unisa.edu.au (D. Nandagopal).
URLS: http://www.psgtech.edu/department/appmaths/vijayalaksmi_r.html, http://scholar.google.co.in/citations?hl=en&user=zBbads4AAAAJ, https://ep.unisa.edu.au/view/view.php?t=bDZMfxFuGcJwle8kK16R, https://groups.google.com/forum/#!topic/unisa-cis-seminars/lqb6w31aYAU, https://ep.unisa.edu.au/view/view.php?id=32962 (R. Vijayalakshmi), http://www.unisanet.unisa.edu.au/staff/homepage.asp?Name=Nanda.Nandagopal (D. Nandagopal).

http://dx.doi.org/10.1016/j.neucom.2015.03.095
0925-2312/

understand and identify the underpinning of the neuronal activities of the brain network interactions, an extensive study of normal brain functioning is required. Such studies would aid better understanding of the dynamic behavior of the normal brain during cognition when compared to the baseline activity and thus assist in the development of methodologies to augment cognition as well [12,13].

One of the most significant research challenges in cognitive neuroscience is to characterize the functional interactions between various brain regions during cognitive activity. In order to construct an influential subnetwork from the fully connected FBN, a popular technique called thresholding is used in the literature. It retains/extracts only the influential/significant connections of the fully connected weighted FBNs for further analysis. The thresholding methods play a central role in construction of subnetworks. Selection of efficient thresholding method is considered significant in network analysis since it may affect both density and topology of the network. Thus, developing novel strategies for thresholding the FBNs remains an open research challenge [14].

Many studies on FBN analysis using network theory have explored unweighted networks [1,3,14,15]. These studies have used various thresholding methods to convert the weighted networks into binary networks where the presence or absence of an edge is represented by 1 or 0 respectively in the adjacency matrix of the network. These thresholding methods ignore the individual strengths of the links from analysis, and hence, accurate results are not guaranteed. Thus, many researchers have changed their focus to consider the strengths of the interactions/associations of the edges as well during the weighted FBN analysis [16,17].

This paper proposes an Influential functional brain network Extraction and Analysis (IEA) framework. It includes a heuristic Branch-and-Bound based thresholding algorithm called the Weighted Subgraph Extraction (WeSE) algorithm to construct the Influential Functional Brain Networks (IFBN) for the fully connected weighted networks. It is later used for investigating the behavior of FBNs during the baseline and cognitive load states. Branch-and-Bound is an efficient algorithmic design technique that performs systematic enumeration in the solution space by using the upper and lower bounds thereby retaining the small subsets of efficient candidate solutions [18]. This paper also explores complex network measures and graph mining algorithms to identify functional interactions and discover the hidden patterns and correlations from the weighted influential FBNs. These results are further validated using inferential statistical methods.

The rest of the paper is organized as follows. A survey on the current approaches to thresholding techniques applied to FBN analysis, their limitations, and an overview of complex network metrics is presented in Section 2. The proposed framework that includes the novel thresholding algorithm WeSE to extract the weighted subgraph, its time complexity analysis, efficient computational methods and complex network metrics to detect and quantify the cognitive activities is discussed in Section 3. The data collection and preprocessing methods are described in Section 4. A detailed discussion on the results of the proposed framework is presented in Section 5. We conclude with summary of findings and future directions in Section 6.

2. Functional brain network analysis

Cognition is a mental process that results from the dynamic interactions of distributed brain regions. The knowledge of structural and functional principles of complex networks facilitates to understand how cognition takes place in the human brain [19]. The electrical activities among the brain regions can be captured from multi-channel EEG data using various linear/non-linear measures such as Pearson's Correlation Coefficient (r), Magnitude Squared Coherence (MSC), Approximate Entropy (AE), Mutual Information (MI), and Synchronization Likelihood (SL) [20–22]. One of the widely used linear metric Pearson's correlation coefficients measures the linear relationship between two signals in terms of the ratio of covariance of the variables to their standard deviations [20]. It measures only the linear dependency between two signals.

The application of non-linear dynamics to EEG data has opened up a wide range of new perspectives for the study of normal and disturbed brain function [22]. MI takes into account both linear and non-linear associations between two random variables X and Y and quantitatively. MI expresses the inherent dependence of X and Y as a non-negative value and is symmetric. It measures the information that X and Y share and is zero if they are independent. High MI value of X and Y indicates the less uncertainty in X knowing Y and vice versa. In this case, MI is the same as the entropy of Y (or X) [23]. Since MI ranges from 0 to $\min\{H(X), H(Y)\}$, to interpret and compare the results across different conditions, the minimum of the entropies of X and Y is used for normalization. The authors have already performed extensive analyses on various linear and nonlinear statistical metrics such as Pearson's Correlation Coefficient, Mean Squared Coherence, Mutual Information, and Transfer Entropy to construct the graph databases of multichannel EEG data. Since r and NMI yielded promising results, these measures have been chosen to test the performance of WeSE algorithm in this paper [36,37].

2.1. Need for thresholding of functional brain networks

Analyzing the characteristics of the fully connected weighted networks called the functional brain networks is much complicated. This is due to various factors such as (i) overhead of exponential number of computationally expensive comparisons, (ii) limited number of analyses and (iii) need for careful interpretation of the analysis results [24]. To avoid these problems, removing the noisy/weak/insignificant connections from the fully connected weighted networks is essential while retaining the significant or relatively strong connections. Various thresholding methods for the neurophysiological/neuroimaging data have been proposed in the literature.

The *fixed thresholding method* selects a threshold to retain only the strong connections based on a significance level computed using some statistical methods, constant threshold across networks, or a threshold that minimizes the number of connections. In this method, the resulting small/large threshold values might generate almost fully connected/disconnected subnetworks. Hence, such subnetworks are not suitable for further analysis in identifying significant and useful patterns.

The *fixed average degree method* uses the value of fixed average degree as threshold which might result in a network that includes many insignificant connections while pruning significant connections. In another method called the *fixed edge density* the threshold is decided based upon the desired density (ratio of the number of actual edges to the number of possible edges) of the connections to be retained for analysis. This is equivalent to fixed average degree method when the networks under study have the same number of nodes [14,25].

These thresholding methods suffer from problems such as user intervention to fix the threshold and the possibility of obtaining disconnected subnetworks. This property of disconnectedness affects the analysis of various network metrics and results in incorrect inferences as well. For example, the characteristic path length diverges to infinity if the network has disconnected nodes. The comparison of the results computed using network metrics for various subjects may be biased when the resulting networks