

Research article

Shortest path based network analysis to characterize cognitive load states of human brain using EEG based functional brain networks

M. Thilaga¹, Vijayalakshmi Ramasamy^{2,3,*}, R. Nadarajan¹, D. Nandagopal³

¹ Computational Neuroscience Laboratory, Department of Applied Mathematics and Computational Sciences, PSG College of Technology, Coimbatore, Tamil Nadu, 641004, India

² Department of Computer Science and Software Engineering, Miami University, Oxford, Ohio, 45056, USA

³ Cognitive Neuroengineering Laboratory, School of Information Technology and Mathematical Sciences, Division of IT, Engineering and the Environments, University of South Australia, Adelaide, 5005, South Australia

*Correspondence: ramasav@miamioh.edu (Vijayalakshmi Ramasamy)

<https://doi.org/10.31083/JIN-170049>

Abstract

Understanding and analyzing the dynamic interactions among millions of spatially distributed and functionally connected regions in the human brain constituting a massively parallel communication system is one of the major challenges in computational neuroscience. Many studies in the recent past have employed graph theory to efficiently model, quantitatively analyze, and understand the brain's electrical activity. Since, the human brain is believed to broadcast information with reduced material and metabolic costs, identifying various brain regions in the shortest pathways of information dissemination becomes essential to understand the intricacies of brain function. This paper proposes a graph theoretic approach using the concept of shortest communication paths between various brain regions (electrode sites) to identify the significant communication pathways of information exchange between various nodes in the functional brain networks constructed from multi-channel electroencephalograph data. A special weighted network called the Shortest Path Network is constructed from a functional brain network where the edge weight is computed as the sum of frequency of occurrence of that edge in all possible shortest paths between every pair of electrodes. The weighted Shortest Path Networks thus constructed portray information on the number of times the edges are used in information propagation during different cognitive states. This network is further analyzed by computing the influential communication paths to characterize the information dissemination among brain regions during different cognitive load conditions. The experimental results presented demonstrate the efficacy of this novel graph theoretic approach in identifying, quantifying, and comparing the significant shortest pathways of information exchange during mild and heavy cognitive load conditions. Analysis also suggests that future research should consider the biological inferences of the ability of the human brain to use reduced material and metabolic cost during the instantaneous transmission of information.

Keywords

Electroencephalography; functional brain networks; cognition; graph theory; shortest paths; brain's electrical activity

Submitted: March 30, 2017; Accepted: September 18, 2017

1. Introduction

It has been one of the major themes of research in the field of computational neuroscience in the recent past to understand the complex functioning of the human brain during different cognitive load states. Cognition is a mental process of transforming sensed information into action. Besides an active research in unveiling the abnormality of brain functioning in the literature, understanding of normal brain functioning has gained significant attention in addressing various mental health issues [1–6]. Characterization of the neuronal interaction patterns of the brain during different activities is a challenging problem and many studies have used graph theory as a tool to model and analyze the organizational principles of the brain to thereby understand the complexity of cognitive function. Graph, a renowned data structure, models the relationships/interactions among entities as networks to facilitate better understanding of the functioning of complex systems [7–11].

There has been an increasing interest in human neuroscientific research not only due to advances both in various non-invasive data acquisition techniques such as neuroimaging (e.g., magnetic resonance imaging (*MRI*), functional magnetic resonance imaging (*fMRI*)) and neurophysiological recording (e.g., electroencephalography (*EEG*), magnetoencephalography (*MEG*)) that help to acquire the electrical activity of the human brain but also due to technological developments in methods of data analysis. Despite the advent of advanced neuroimaging technologies, for several decades the wide use of *EEG* in many areas of clinical work and research to investigate the brain's behavior has been promising due to its non-invasive nature, superior temporal resolution, and low hardware cost etc [12]. The electrical activity of the brain recorded using multi-channel *EEG* has enabled acquisition of dynamic and non-trivial information related to the patterns of interaction of various brain regions. Analysis of such massive spatiotemporal databases requires the use of extensive and

efficient computational knowledge acquired from many disciplines, such as graph theory, inferential statistics, signal processing, and complex network theory, which help in the understanding of human brain behavior during various mental tasks.

Recent studies in neuroscience have revealed the notion of shortest paths for the transfer of information within the brain. The human brain exchanges information with reduced material and metabolic costs [13]. The fact that brain networks use the shortest paths to communicate information during different brain activities has been analyzed by investigators to understand functional connectivity with regard to structural connectivity [14]. Some of the network analysis measures used in Functional Brain Network (*FBN*) analysis such as average path length, diameter, and global efficiency are based on the concept of shortest paths, thus demonstrating the fact that the most effective information dissemination takes place through shorter paths. A short average path length and high clustering coefficient are the characteristics of small-world networks exhibited by the brain networks that allow segregated and cost-effective information processing with regard to wiring and metabolic cost [15]. Hence, analyzing the shortest information communication pathways between brain regions provides a better understanding of brain function during various cognitive activities in normal brain and when networks are altered by brain diseases. Many investigators have used the concept of shortest path to study the topological properties of brain networks in the recent past. The current study emphasizes the analysis of the shortest path networks constructed for mild and heavy cognitive load states to characterize the information communication patterns among brain regions during these states. It uses the thresholded subnetwork of the *FBN* derived from a special thresholding algorithm (proposed by the authors) for studying the shortest paths between brain regions used for information dissemination.

This paper proposes a shortest path based network analysis approach to characterize mild and heavy cognitive load conditions of the human brain by computing the number of times data transmission occurs through the shortest paths in the *FBNs*. The *FBN* is constructed using a non-linear statistical metric on preprocessed *EEG* data results in a fully connected weighted *FBN*. To retain and analyze only those strong connections in *FBN*, a Branch-and-Bound based thresholding algorithm called Weighted Subgraph Extraction algorithm (*WeSE*) is used as a measure of data reduction [16]. A special weighted network called the Shortest Path Network (*SPN*) is constructed from the thresholded *FBN* by computing the weight of each edge as the number of times it occurs for all possible shortest paths of data transmission between all pairs of electrode sites. Thus, the weights of the edges in *SPN* contain the essential information about the number of times each edge lies in the shortest communication path between brain regions. The weighted *SPNs* constructed for mild and heavy cognitive load states are further analyzed using graph theoretical methods and statistical tests to characterize cognitive activity of the brain.

2. Functional connectivity analysis of brain using network theory

Neuroscience is one of the most predominant areas of research to study the human brain activity ranging from macro to micro scales. The massive volumes of data generated by various neuroimaging/neurophysiological techniques have posed many challenges to the computational neuroscience research community including mod-

elling, storing, and mining of the data. To address these challenges, many investigators have analyzed the complex and dynamic behavior obtained from the human brain by use of advanced methods derived from graph theory, signal processing, statistics, and information theory. In particular, connectivity based methods derived from graph theory have played a significant role in characterizing the brain's behavior during different activities. Graph based analytical techniques model brain regions and the complex relationships that exist between them as nodes and edges, respectively. The intrinsic neuronal connectivity patterns attributed to various normal and disturbed brain functions are explored using neuroimaging/neurophysiological data that is modelled as a weighted graph/network. These networks are constructed by different brain regions (electrode sites) as nodes/vertices, the physical connections (synapses or axonal projections) or functional association among brain regions as edges (links) between them, and the connectivity strength (correlations between the nodes) as weights of the edges [8].

The magnitude of temporal dependency measured among n brain regions during a specific activity using linear/non-linear statistical measures results in an $n \times n$ functional connectivity matrix that portrays the multivariate relationships between them. The temporal dependencies computed between all pairs of brain regions are represented as a network referred to as a Functional Brain Network where electrode sites and the statistical dependencies computed between them are considered as its nodes and edge weights, respectively [17]. A range of 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*) have been used by various researchers to construct *FBN* from neuroimaging/neurophysiological data [18–22]. While the linear measures capture only the linear dependency between two variables, many investigators have employed non-linear measures (such as *MI*, *SL*) to analyze underlying non-linear relationships, as such measures take into account both linear and non-linear associations between two variables. Here, *MI* is employed to construct *FBNs* as a quantitative measure of the relation between pairs of random variables. The *MI* between two discrete random variables X and Y is computed as

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \left(\frac{p(x,y)}{p(x)p(y)} \right) \quad (1)$$

where $p(x,y)$ is the joint probability distribution function of X and Y . $p(x)$ and $p(y)$ are the marginal probability distribution functions of X and Y , respectively. The *MI* of X and Y is zero if they are independent and it is a symmetric measure. *MI* is the same as the entropy of Y (or X). Since the *MI* of X and Y lies in the range 0 to $\min H(X), H(Y)$, where $H(X)$ and $H(Y)$ are the entropies of X and Y respectively, it needs to be normalized before comparing the results across different activities. In analyses reported here, a minimum of the entropies of X and Y is used for normalization [20]. To facilitate analysis of a fully connected weighted *FBN*, a process called thresholding, that eliminates the noisy/ weak connections, is used thereby retaining only the relatively strong connections. Various thresholding methods such as fixed thresholding, fixed average degree, and fixed edge density are used to analyze *FBNs* [23–25]. To overcome a number of limitations in these techniques, a Branch-and-Bound based *WeSE* algorithm which does not use any arbitrarily chosen threshold value is used. The novelty of this algorithm is that, it dynamically selects

only the strong connections from a fully connected weighted network while preserving the connectedness of the network [16].

Graph theory based complex network analysis techniques are extensively used to examine various topological properties of the functional connectivity of networks. Investigating the shortest paths between the network elements of the thresholded *FBN* opens a useful perspective on the influential communication pathways. Such an analysis provides a comprehensive understanding of the small-world properties of brain network organization, nodes that are hubs during a particular brain activity, community structures, and the local and global communication abilities [26].

2.1. Role of shortest path analysis in complex networks

Many topological properties of real-world complex networks such as biological, social, and technological networks are formalized based upon the assumption that the information flows in a network along the shortest (or geodesic) paths so as to optimize transportation cost. Given a weighted undirected graph $G = (V, E)$ with positive weight function $w: E \rightarrow \mathbb{R}$, the weight w of a path $p = (v_0, v_1, \dots, v_n)$ is computed as

$$w(p) = \sum_{i=1}^n w(v_{i-1}, v_i) \quad (2)$$

where $w(p)$ is the sum of the weights of the edges of the path P . For any two nodes s (source) and t (destination), the shortest path weight (represented as $\delta(s, t)$) is computed as

$$\delta(s, t) = \begin{cases} \min\{w(p)\} & \text{if there is a path } p \text{ from } s \text{ to } t \\ \infty & \text{otherwise} \end{cases} \quad (3)$$

The shortest path between the nodes s and t is defined as any path p with weight $w(p) = \delta(s, t)$ [27]. In other words, the shortest path between two nodes s and t is defined as an ordered set of edges linking the nodes s and t where the sum of the weights of the edges of the path is minimal. A network may contain multiple shortest paths between a pair of nodes. The concept of shortest paths has been widely used in many application areas including social networks (to identify influential persons and communities), metabolic networks (to find the optimal pathways between compounds), transportation networks (to identify efficient routes between source and destination) and communication networks (to efficiently manage various resources) [28–31].

Shortest paths in *FBNs* have been considered the principal routes/functional pathways through which information propagates most of the time to various brain regions [27]. In particular, it has already been proved that the human brain exhibits small-world properties characterized by a high clustering coefficient (densely connected local clusters) and short path length between any pair of nodes. This property confirms both the segregated and distributed information processing abilities of the brain [15]. The small-worldliness further ensures that information is disseminated along the shortest paths thereby minimizing conduction delays and the energy required to propagate information between various brain regions. Shortest paths in *FBNs* have also demonstrated their role in the prediction of the cognitive abilities of individuals [32]. Moreover, the concept of shortest path has been used as an underlying metric in defining a number of popular network measures such as closeness, betweenness,

characteristic path length, local and global efficiencies, eccentricity, and diameter [10, 33].

Various investigators have studied and characterized different types of networks using shortest paths based on the assumption that information dissemination over shortest paths is faster and less noisy [8]. Shortest-path based network analysis has been used to identify the genes that modulate longevity by construction of a shortest path longevity network from a protein-protein interaction dataset. Analysis of this network enabled not only the identification of longevity-associated genes but also the genes that cause multiple age-associated diseases such as cancer, heart disease, and neurodegenerative disorders [34]. The shortest paths between the nodes in a network have also served as a sampling method for online social networks. Each edge occurring in the shortest paths between many pairs of selective nodes are ranked based on the number of times they appear in various shortest paths. The subnetwork (sampled network) that includes a specified percentage of the highly ranked edges is then extracted and used for further analysis [35].

The Union of Shortest Path Trees (*USPT*) rooted at each node in the network has been used as a sampling method to analyze *FBNs*. The *USPTs* of a *FBN* represent the most important connections. These are interpreted as the functional highways of brain networks. Analyzing these shortest paths facilitates both the identification of pathological network changes that result in certain brain diseases, such as multiple sclerosis, and the significant differences in network topology between patients with mental disorders and healthy controls [27]. Thus, the study of shortest paths in *FBNs* helps to identify frequently used communication paths and influential nodes that occur repeatedly in shortest paths and to characterize brain functions during different activities.

2.2. Data collection and pre-processing

The *EEG* data used in this study was collected in the Cognitive Neuroengineering & Computational Neuroscience Lab (*CNEL*) at the University of South Australia, HREC Approval (30855), from nine participants (P1 through P9). Participants were asked to drive a US standard Simuride driving simulator. A 30 channel Nuamps *EEG* amplifier is used to acquire the *EEG* data using Curry V7 software. *EEG* data is collected for two distinct cognitive load conditions: a simple driving along a curved track maintaining constant speed (*Drive*) and driving with audio distraction (*DriveAdo*). Detailed description about cognitive load experiments given to participants during data collection and the data preprocessing methods used are given in Thilaga *et al.* [16].

3. Characterizing cognitive load states using shortest path based approach

The proposed Shortest Path based Network Analysis Framework (*SPNA Framework*) is shown in Fig. 1 with the following four-fold objectives.

- (i) Construct Graph Database (*GD*) containing *FBNs* for the two cognitive load conditions mentioned earlier from *EEG* data pre-processed using *NMI*.
- (ii) Extract the thresholded (influential) subnetworks from the fully connected *FBNs* using a dynamic Branch-and-Bound based Weighted Subgraph Extraction (*WeSE*) thresholding algorithm.

- (iii) Construct Shortest Path Networks from thresholded *FBNs* using Dijkstra's shortest path algorithm, and
- (iv) Characterize the brain functioning during mild and heavy cognitive activities using *SPNs*.

The following subsections describe the steps involved in modelling and analyzing *FBNs* to characterize brain function during the various cognitive activities.

3.1. Construction of *FBNs*

The preprocessed *EEG* data (time domain-amplitude) of mild (*Drive*) and heavy (*DriveAdo*) cognitive load states are partitioned into a number of chunks (each of length two seconds). It is well known that the human brain is a complex non-linear system showing emergent properties. The *EEG* signal produced by the brain is non-stationary as its statistical properties are time variant [36]. However, in time and frequency domain analysis, an *EEG* signal is considered stationary over a short time interval for efficient estimation of different average statistical functions [37, 38]. Non-stationarity in *EEG* signals and methods to overcome this through segmentation, including 2 second segmentation, has been described well in literature [39–43]. Any choice of *EEG* time segment (epoch length) for analysis is always associated with a compromise between time and frequency resolution. It is not possible to choose a segment length which will give both high temporal and frequency resolution [44]. Here, two second *EEG* segments are used, following previously reported literature. Each chunk of data is modelled as a *FBN* by considering the n electrode sites (in this study, $n = 30$) as nodes and the pair-wise associations of these sites measured using a non-linear statistical measure *NMI* as weights on the edges. These *NMI* graphs computed for forty chunks of *EEG* data are averaged into a single graph. The fully connected weighted undirected *NMI* graphs (*FBNs*) of different cognitive load states are stored into *GD*. To retain only the influential/strong connectivity patterns that contribute to the study of cognitive activities, the weak/false positive connections are eliminated from the undirected complete *FBNs*. To achieve this, a dynamic Branch-and-Bound based *WeSE* algorithm is used to obtain the thresholded subnetworks. The significance of the *WeSE* is that it dynamically extracts the influential connections from the fully connected *FBNs* without any user specified threshold value while simultaneously ensuring network connectivity [16]. These weighted undirected thresholded graphs/networks are used to study the shortest communication paths of information exchange in the *FBNs* during mild and heavy cognitive load states.

3.2. Construction of Shortest Path Networks

To identify and analyze the frequently used shortest communication paths of information transfer in the thresholded *FBNs* stored in *GD*, all the shortest paths between all pairs of nodes are computed using Dijkstra's algorithm [45]. A special network called Shortest Path Network (*SPN*) is constructed using Algorithm 1.

The *SPN* is defined as a weighted undirected network/graph $G = (V, E, w)$, where V and E are sets of vertices (electrode sites) and edges (links between the electrode sites) respectively and w represents the weights on the edges E computed as the frequencies of occurrences of the edges in all possible shortest paths between all pairs of nodes. Algorithm 1 computes all the shortest paths between all pairs of nodes in the thresholded *FBNs*. For each edge (u, v) in the shortest paths, it computes the number of times (frequency) it lies in all possible shortest paths (steps 5 to 9) between all pairs of nodes.

Algorithm 1 Construction of *SPN*

Input: Thresholded *FBNs* in *GD*

Output: *SPNs*

1. $frequency(u, v) \leftarrow 0$, $Paths = \{\emptyset\}$, $SPN \leftarrow \{\emptyset\}$, (u, v) is the edge between the nodes u and v
2. for each *FBN*, FBN_i , $1 \leq i \leq N$, $N = \text{number of } FBNs \text{ in } GD$
3. for each node, $Node_j$ in FBN_i , $1 \leq j \leq n$
4. $Paths \leftarrow$ Set of all possible shortest paths from $Node_j$ to $Node_k$, $1 \leq k \leq n, j \neq k$
5. for each $PathP_a$ in $Paths$
6. for each edge (u_j, v_k) in P_a
7. $frequency(u_j, v_k) \leftarrow frequency(u_j, v_k) + 1$;
8. end for
9. end for
10. Update SPN_i with all the edges in P_a with $weight(u_j, v_k) = weight(u_j, v_k) + frequency(u_j, v_k)$
11. $frequency(u_j, v_k) \leftarrow 0$, $1 \leq j, k \leq n$
12. end for
13. Append SPN_i into *GD*
14. end for

The frequency values computed for all the edges are then updated as edge weights in the *SPN* (step 10). For an edge (u, v) in *SPN*, the high edge weight indicates that this edge is used many times in the shortest path of information transfer and hence, it is influential and dominant. Therefore, by analyzing the *SPNs* of different cognitive load states, it is possible to identify the influential functional pathways that are used for information dissemination most of the times during a specific brain activity.

3.3. Analysis of *SPNs*

Analyzing the topological characteristics of *SPN* using efficient computational methods helps in the study of information communication patterns in the brain. The edge weights in the *SPN* of a specific brain activity contain significant information about how many times a particular edge has been used in the shortest path of information exchange to other brain regions. Hence, by comparing the edge weights of *SPNs* of mild and heavy cognitive load states, the Dominant neuronal Connections (*DCs*) due to specific cognitive activity can be identified. The *DCs* in the *SPN* of mild cognitive load state (SPN_{DCmild}) are defined as those connections with high edge weights when compared to heavy cognitive load state and are computed as

$$SPN_{DCmild} = \text{Edges with weight in } SPN_{Mild} > SPN_{Heavy} \quad (4)$$

Similarly, the *DCs* of heavy cognitive load state ($SPN_{DCHeavy}$) are computed by including only those edges with high edge weights during this state and is given as:

$$SPN_{DCHeavy} = \text{Edges with weight in } SPN_{Heavy} > SPN_{Mild} \quad (5)$$

This connectivity analysis provides useful information at the macro level based on the number of dominant connections present in *SPNs* of mild and heavy cognitive load states. A two-tailed *t*-test is carried out by comparing the number of *DCs* in each brain region (electrode sites) in the SPN_{DCmild} and $SPN_{DCHeavy}$ networks to ensure that the group means of these networks are significantly different. To further investigate only those high weighted edges that contribute more to the study of frequently used shortest communica-

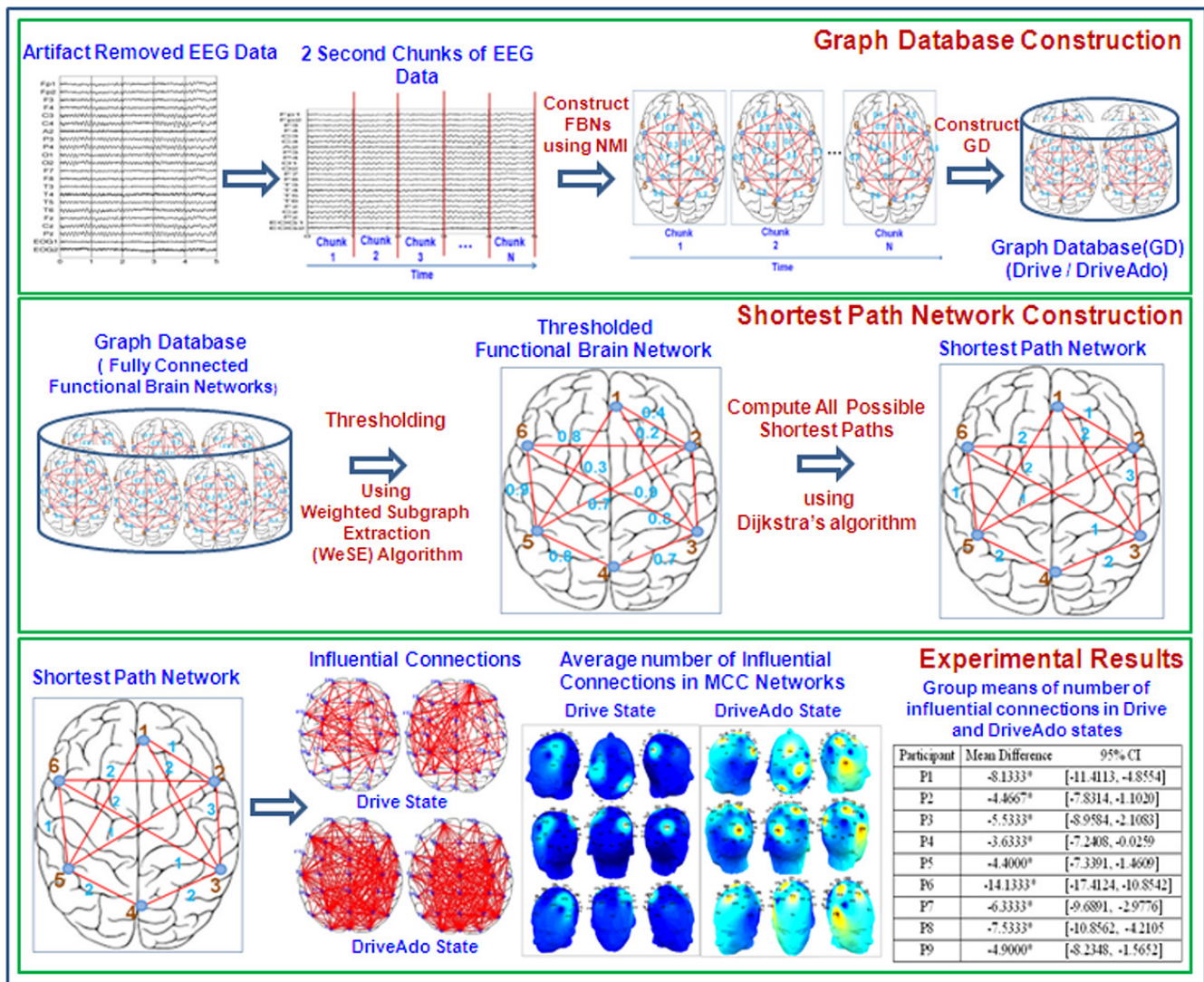


Fig. 1. Shortest Path based Network Analysis Framework.

tion paths among the brain regions, the highly dominant connectivity patterns in the *SPNs* are extracted using a graph theoretic approach called Minimum Connected Component (*MCC*) [3]. *MCC* is a spanning subgraph of the original graph which extracts significant connected components of a given weighted graph without using any specified parameter. The *MCCs* of *SPNs* constructed for mild and heavy cognitive load states are further analysed to uncover the brain regions that are involved in transmitting information through the shortest paths most of the time.

The construction of a *SPN* is illustrated using an undirected fully connected weighted graph $G(FBN)$ with seven nodes (edge weights represent the statistical dependencies between the nodes) and its adjacency matrix as shown in Fig. 2a and 2b.

To reduce the computational complexity involved in analyzing the *FBNs*, the weak/insignificant connections are removed from G using a *WeSE* thresholding algorithm. The thresholded subnetwork of G (G') constructed using the *WeSE* algorithm is shown in Fig. 3a.

Fornito et al. [46] have stated that the edges with high weights in a functional connectivity matrix are the strongest and the most

reliable connections among brain regions. Performing a shortest path analysis using the weighted *FBN* is anomalous because the shortest path between any two nodes s and t is computed as the sum of weights of the edges between them in the paths that are minimal. As a consequence, such an analysis includes less influential edges and avoids the most important communication routes of information transfer (i.e., high weighted edges). To overcome this problem, a monotonic remapping of the edge weights is performed before computing the shortest paths. This remapping converts the largest weights into smallest weights and vice versa. If the edge weights w_{ij} , $1 \leq i, j \leq n$, are between 0 and 1, either $w_{ij} \leftarrow -\log w_{ij}$ or $w_{ij} \leftarrow 1 - w_{ij}$ can be considered as remapping functions [14]. Since the weights of the edges in the *FBN* are positive values in the range 0 to 1, the edge weights are remapped by subtracting them from 1. This enables the shortest path algorithm to choose high weighted edge at each step. In the current study, Dijkstra's shortest path algorithm [45] is used to compute all possible shortest paths between all pairs of nodes from the given *FBN*. Graph G' after remapping of edge weights is shown in Fig. 3b. All possible shortest paths computed between all

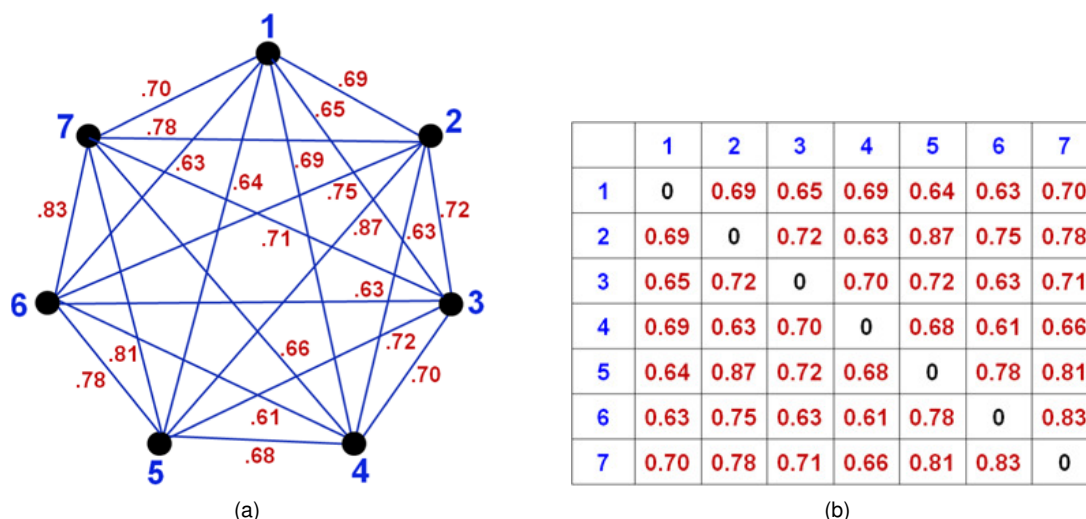


Fig. 2. (a) Undirected fully connected weighted Graph G ; (b) Adjacency Matrix of G .

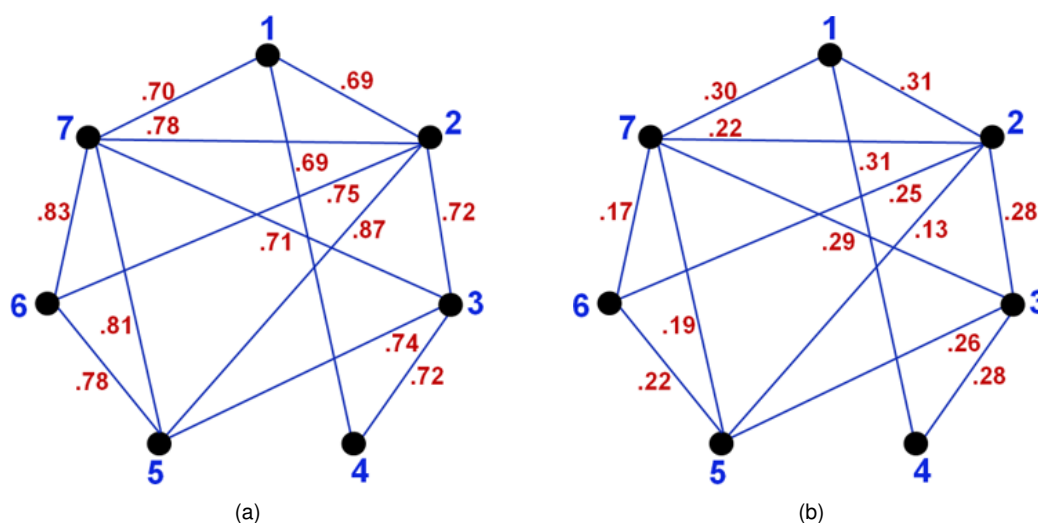


Fig. 3. (a) Thresholded subnetwork G' ; (b) G' after remapping the edge weights.

pairs of nodes in G' (Fig. 3b) are shown in Table 1.

Table 1. All possible shortest paths from each node to every other node in G' after remapping the edge weights

Nodes	1	2	3	4	5	6	7
All possible shortest paths	1-2	2-1	3-2-1	4-1	5-2-1	6-7-1	7-1
	1-2-3	2-3	3-4-1	4-3-2	5-2	6-2	7-2
	1-4-3	2-3-4	3-7-1	4-3	5-3	6-7-3	7-3
	1-7-3	2-5	3-2	4-3-5	5-3-4	6-7-3-4	7-3-4
	1-4	2-6	3-4	4-3-7-6	5-6	6-5	7-5
	1-2-5	2-7	3-5	4-3-7	5-7	6-7	7-6
	1-7-6		3-7-6				
	1-7		3-7				

From the shortest paths list, the frequency of occurrence of each edge is calculated by counting its presence in all possible shortest paths. This gives useful information about which edges are promi-

nent in propagating information through the shortest paths. For instance, in the illustrated example, the edge (3, 4) occurs in six different shortest paths (2-3-4, 3-4-1, 3-4, 5-3-4, 6-7-3-4 and 7-3-4) and hence the edge weight of this edge in SPN is 6. The frequencies of occurrence of all the edges in G' are computed to represent the edge weights of the SPN . The SPN of G' (assuming that this network represents the heavy cognitive load state) constructed using Algorithm 1 and its adjacency matrix are shown in Fig. 4a and 4b respectively.

Using the $SPNs$ of mild and heavy cognitive load states, the DCs in each of these states are extracted using equations (4) and (5). The extraction of DCs from the $SPNs$ of these cognitive load states is illustrated using the two networks shown in Fig. 5a and 5b as SPN_{Mild} and SPN_{Heavy} cognitive load networks, respectively.

To accomplish the objective of characterizing the information communication patterns among various brain regions during mild and heavy cognitive load conditions using a shortest path based network analysis approach, the following heuristic is employed. If an edge weight (representing the frequency of occurrence of that edge

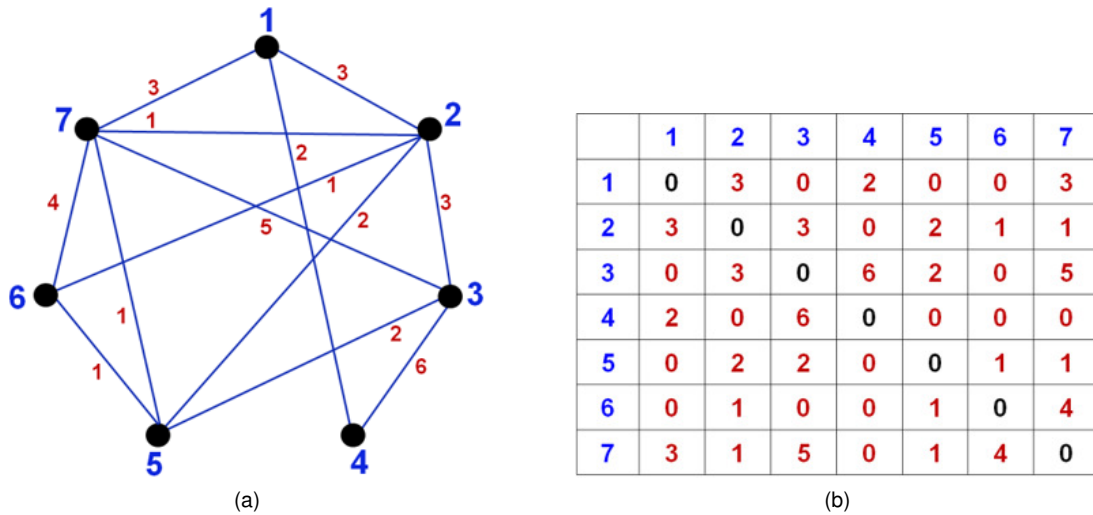


Fig. 4. (a) SPN_{Heavy} constructed from G' ; (b) Adjacency Matrix of SPN_{Heavy} .

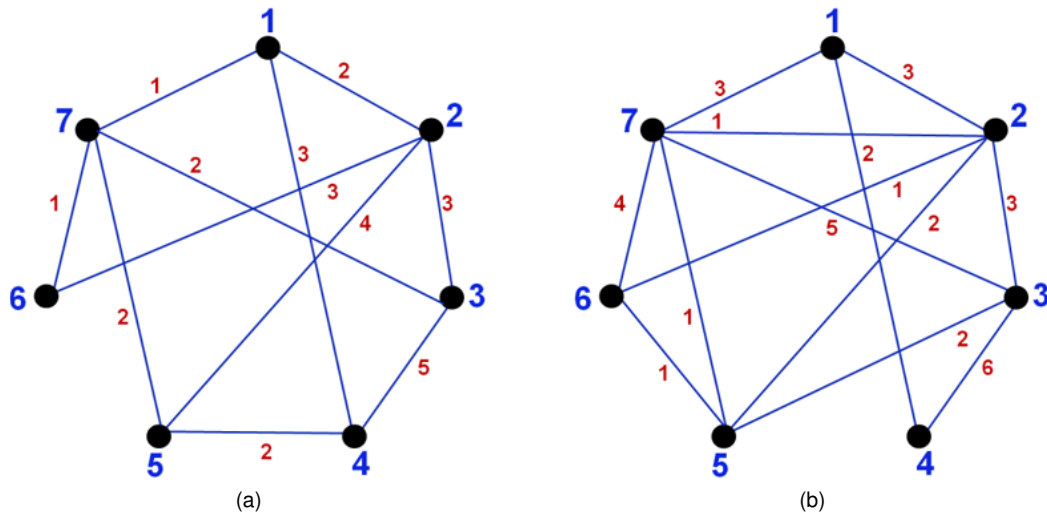


Fig. 5. (a) SPN_{Mild} ; (b) SPN_{Heavy} .

in all possible shortest paths between all pairs of nodes) is more, for instance, during heavy cognitive load state than that of mild cognitive load state, it is retained only in the SPN of the heavy cognitive load state. In case of equal frequencies, the contribution of that edge in disseminating information during different cognitive load states is equal. Hence, such an edge will neither help expose the differences nor a change in the communication patterns of information exchange during mild and heavy cognitive load states.

The heuristic used to compare the SPN networks highlights the most significant shortest functional pathways of information transmission during a specific cognitive activity. For instance, the edge weight of the edge (1, 2) in SPN_{Mild} and SPN_{Heavy} are two and three, respectively, as shown in Fig. 5a and 5b. So, this DC is retained in SPN_{Heavy} but removed from SPN_{Mild} . Similarly, the edge (2, 3) has edge weight three in both the SPN_{Mild} and SPN_{Heavy} networks and hence it is removed from both these networks. This process is repeated for all the edges present in the $SPNs$ of mild and heavy cognitive load states. The resulting SPN_{DCMild} and $SPN_{DCHheavy}$

networks are shown in Fig. 6a and 6b, respectively. The DCs in the SPN_{DCMild} and $SPN_{DCHheavy}$ networks show the influential and frequently used connectivity patterns employed for dissemination of information to various other brain regions during mild and heavy cognitive load conditions respectively.

Further analyzing these edges with efficient computational techniques, such as graph theory based methods and statistical tests, would facilitate detection of the shortest communication patterns of brain networks during different brain states.

4. Results and discussion

The $SPNs$ constructed from the thresholded states of the $FBNs$ of *Drive* and *DriveAdo* for participants were analyzed using graph theoretic concepts and the results are statistically validated using a two-tailed statistical t -test. The results and discussion of various analyses of $SPNs$ are included in the following subsections.

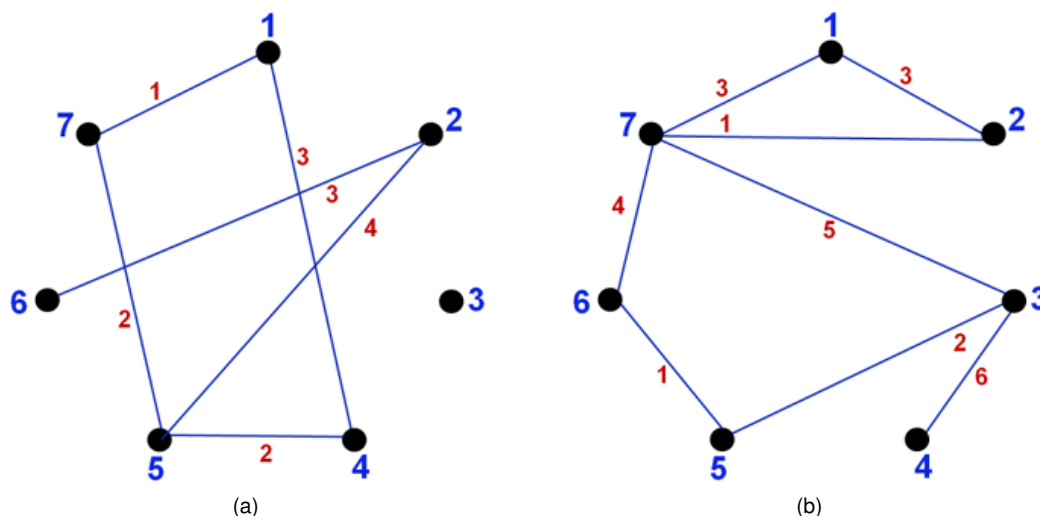


Fig. 6. SPNs with Dominant Connections (a) *SPNDCMild*; (b) *SPNDCHeavy*.

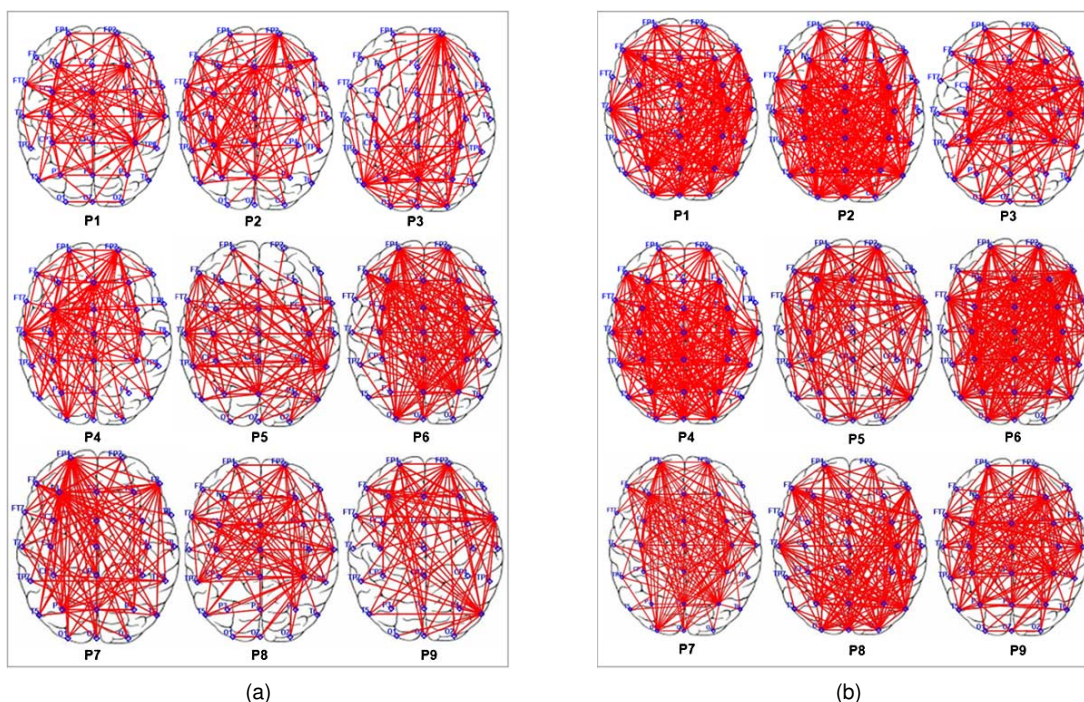


Fig. 7. SPNs of all the participants (a) *Drive* state; (b) *DriveAdo* State.

4.1. Analysis of connectivity patterns of SPN

The connectivity patterns (edges) between the electrode sites of SPNs of *Drive* and *DriveAdo* states of all participants are shown in Fig. 7a and 7b, respectively.

It is important to note that the edges in the SPNs of *Drive* and *DriveAdo* states are computed from the thresholded FBNs of the respective states. The thresholded FBN (obtained using the *WeSE* thresholding algorithm) includes only the influential edges (edges with high edge weights) which are determined by the topology of the fully connected FBN and associated edge weights. Hence, the

shortest paths computed using the thresholded FBN provides only high-probability paths, such that information is disseminated over them to various other brain regions. These shortest paths (computed using Dijkstra's algorithm) are used to construct the SPNs of *Drive* and *DriveAdo* states as shown in Fig. 7a and 7b, respectively. Since, the SPNs of these states are dense, it is very difficult to identify the connections employed by the shortest paths most of the time for information exchange during a particular brain activity. To understand only the relatively strong (dominant) connectivity patterns among the electrode sites which might characterize the *Drive* and *DriveAdo* states, the edge weights in the SPNs of these states are compared

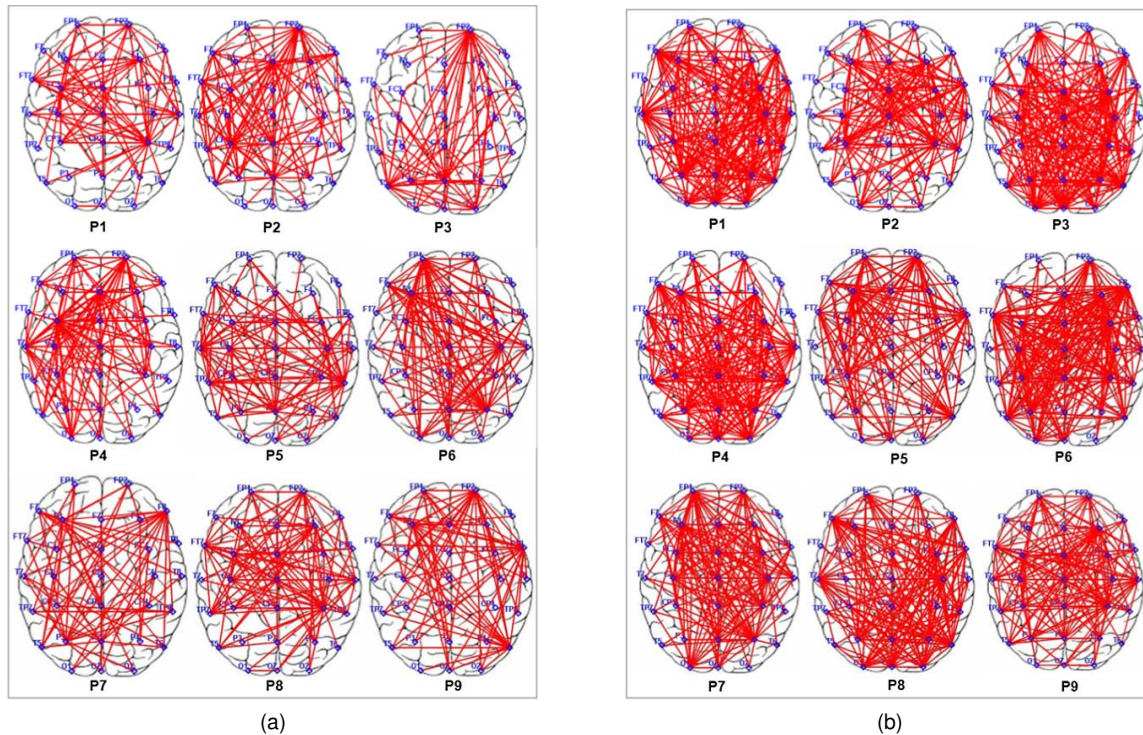


Fig. 8. Dominant Connections of all the participants (a) *SPNDCDrive*; (b) *SPNDCDriveAdo*.

using the heuristic as discussed in Section 4.3. Fig. 8a and 8b illustrate the DCs of *Drive* (*SPNDCDrive*) and *DriveAdo* (*SPNDCDriveAdo*) states, respectively, for all participants. These dominant connections are used in information transfer many times during the respective cognitive load state.

The DCs present in the SPNs of *Drive* and *DriveAdo* states are analyzed using an edge density measure. From Fig. 8a and 8b, it is interesting to note that the number of DCs (edge density) during the *DriveAdo* state is relatively high when compared to the *Drive* state for all participants. The high edge density during *SPNDCDriveAdo* indicates that many edges are utilized in information exchange using shortest paths. Moreover, the high edge density results of the short average path length of the network enable effective communication among the various regions active during the *DriveAdo* state.

The presence of many DCs between brain regions during the *DriveAdo* state further indicates that information flow to other brain regions takes place over many shortest paths enabling higher functional integration during heavy cognitive load states. To statistically validate that the number of DCs of electrode sites between the *SPNDCDrive* and *SPNDCDriveAdo* networks are significantly different, group means are computed using a post-hoc *t*-test (two-tailed) at $\alpha = 0.05$ with Bonferroni adjustment. Results for all participants are shown in Fig. 9.

The results of the *t*-tests reveal that the group means of the number of DCs of electrode sites computed for *Drive* and *DriveAdo* states are significantly different for all participants and that the number of DCs of electrode sites in the SPN of *DriveAdo* state is relatively high. This indicates that with an increasing cognitive load, the number of DCs connections employed to transfer information between brain regions also increases. Experimental results further illustrate that

the number of DCs present in the *SPNDCDriveAdo* network is greater when compared to the *SPNDCDrive* state for all participants.

Table 2. Group means of the number of DCs of electrode sites in *SPNDCDrive* and *SPNDCDriveAdo* networks of all the participants

Participant	Mean Difference of DCs	95% CI	<i>p</i> -value
P2	−4.4667*	[−7.8314, −1.1020]	0.0102
P3	−5.5333*	[−8.9584, −2.1083]	0.002
P4	−3.6333*	[−7.2408, −0.0259]	0.0447
P5	−4.4000*	[−7.3391, −1.4609]	0.038
P6	−7.2667*	[−10.9129, −3.6205]	0.00061
P7	−6.3333*	[−9.6891, −2.9776]	0.00573
P8	−7.5333*	[−10.8562, −4.2105]	0.0003145
P9	−4.9000*	[−8.2348, −1.5652]	0.01356

*Mean difference is significant at $\alpha < 0.05$ level.

To quantitatively compare differences in the number of DCs between the electrode sites of the *SPNDCDrive* and *SPNDCDriveAdo* networks, the mean differences between the number of DCs of these states are computed using post-hoc *t*-tests (two-tailed) at $\alpha = 0.05$ with Bonferroni adjustment and the results are presented in Table 2. After Bonferroni adjustment, *p* (new α -values) less than 0.05 are considered statistically significant. It can be seen that *p* values for all the participants are less than 0.05, hence the group means of all participants are significantly different. Therefore, the null hypothesis that the mean difference of the number of DCs of the electrode sites in the *SPNDCDrive* and *SPNDCDriveAdo* networks is not significantly different is rejected for a significance level of 5%.

The results shown in Fig. 8a and 8b give only limited information about the number of DCs present in the SPNs of *Drive* and

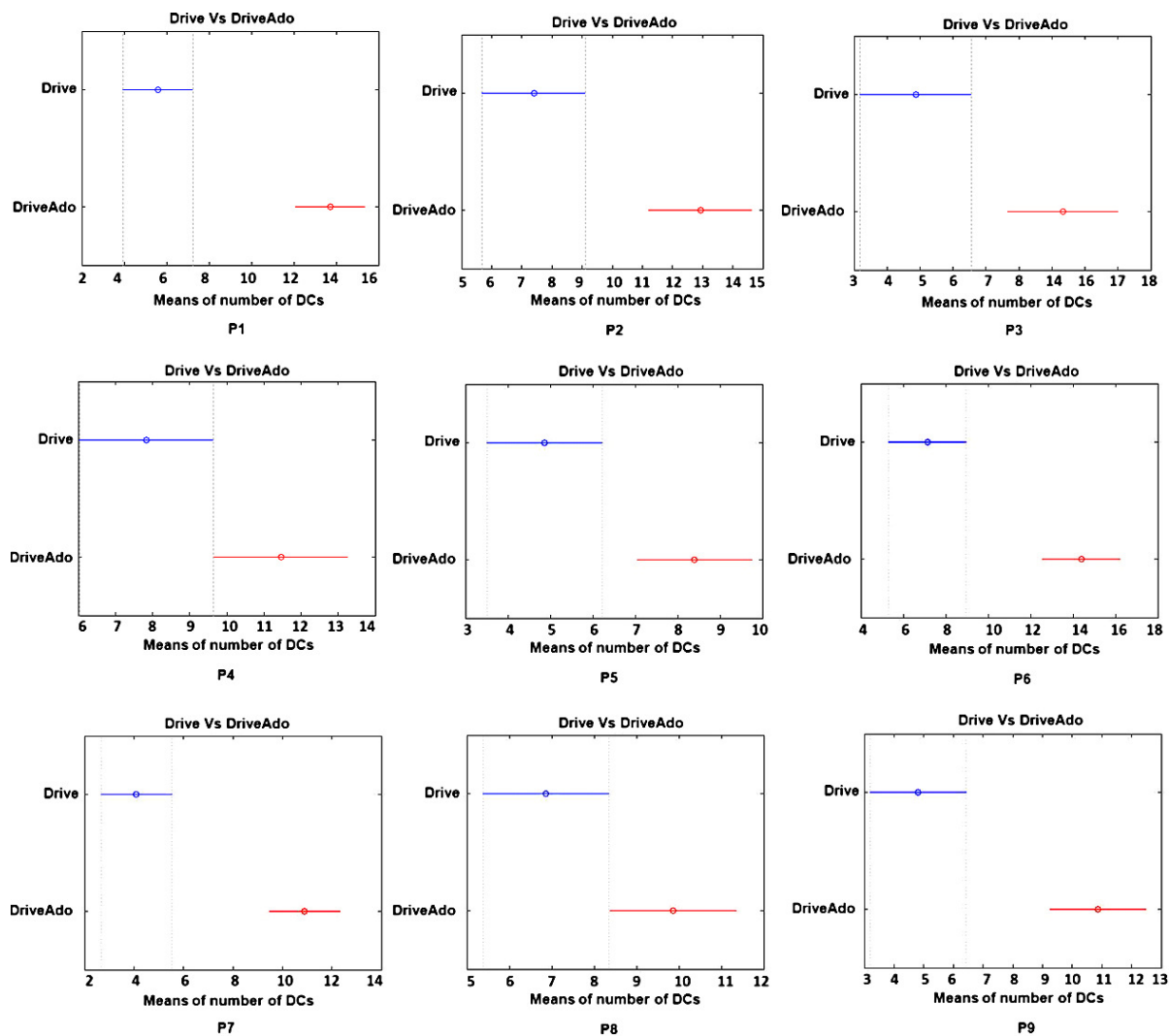


Fig. 9. Statistical *t*-test results of all the participants.

DriveAdo states. To further investigate and extract the highly dominant connections which occurred most of the time in the shortest paths of information exchange, the graph theoretic concept of *MCC* was used. It extracted only the highly dominant connections (i.e., the edges with higher weights) from the $SPN_{DCDrive}$ and $SPN_{DCDriveAdo}$ networks. Using the *MCC* networks of mild and heavy cognitive load states, the average number of connections of all the electrode sites normalized in the range of values 0 to 1 (low to high), represented using the color bar (blue: low to red: high), is plotted in the 3D headplot as shown in Fig. 10 to visualize and identify the electrode sites which are many times found within the shortest paths of information exchange between other electrode sites.

The empirical results of headplots show that only a few electrode sites in the frontal, central, and central-parietal regions during the *Drive* state used many shortest paths for exchanging information. On the other hand, during *DriveAdo*, transmission of information occurred through many of the frontal electrode sites using highly dominant shortest path connections. This can be interpreted to mean that the frontal region is much involved in faster information propagation to various other brain regions during the *DriveAdo*, which assists

in promoting various significant cognitive skills such as problem solving, reasoning, creativity, and judgment.

5. Conclusion

The shortest path based network analysis approach is employed in this study to identify information communication patterns in the brain during mild and heavy cognitive load states. The uniqueness of this study is that the shortest paths are computed using a thresholded network of fully connected *FBN* that includes only influential connections (connections over which the amount of information exchange is high). Using these influential connections, all possible shortest paths between all pairs of nodes were computed and the number of occurrences of each edge was calculated to construct a *SPN* network. The *SPNs* of mild and heavy cognitive load states are analyzed using a heuristic that highlights the *DCs* (i.e., edges occurring in the shortest paths many times) of the respective state. The connection density of the networks of different brain states informs the spread of information to various brain regions. The empirical analysis performed using the *DCs* revealed that the number of *DCs* (connection density) during *DriveAdo* is relatively higher than that of

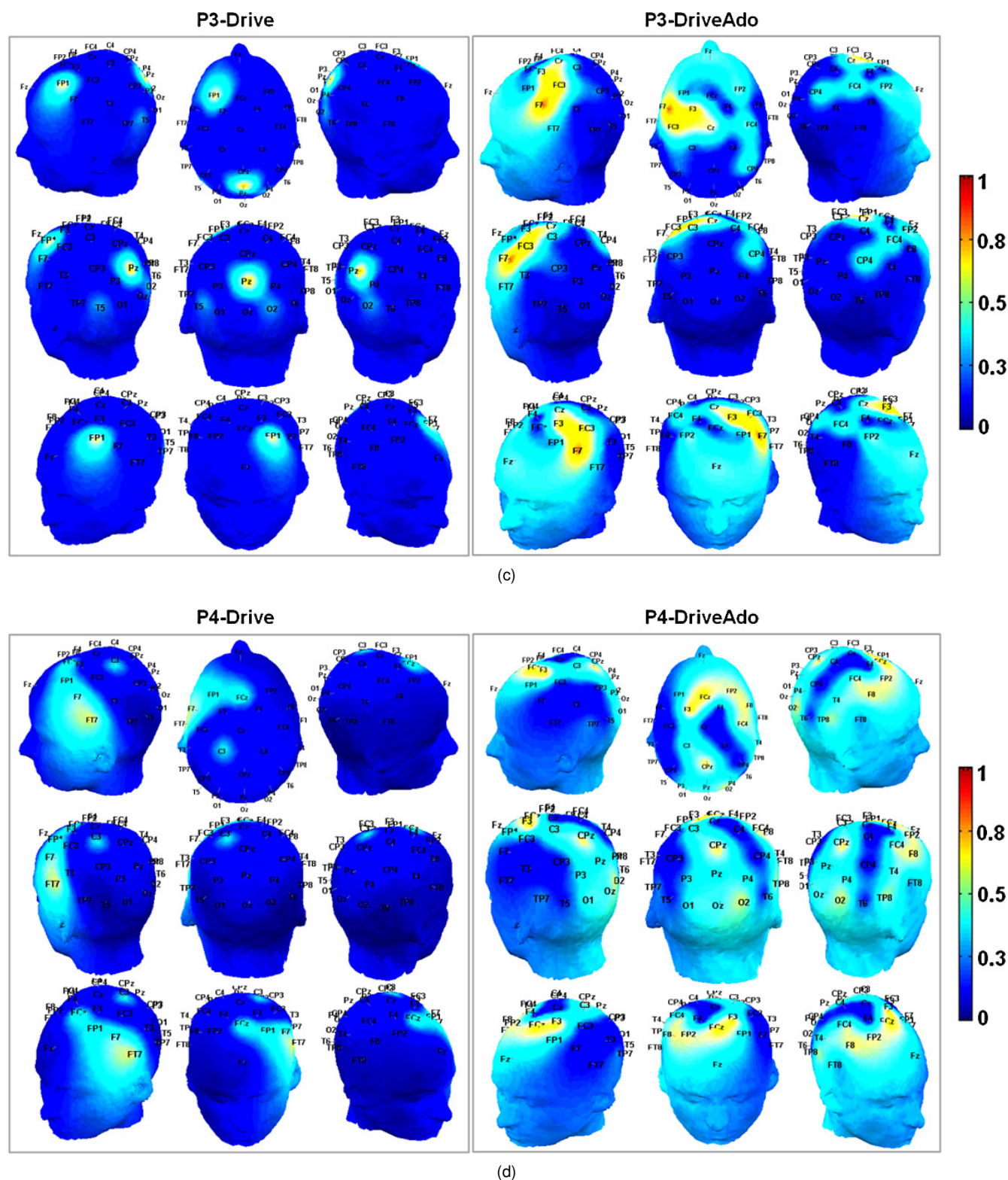
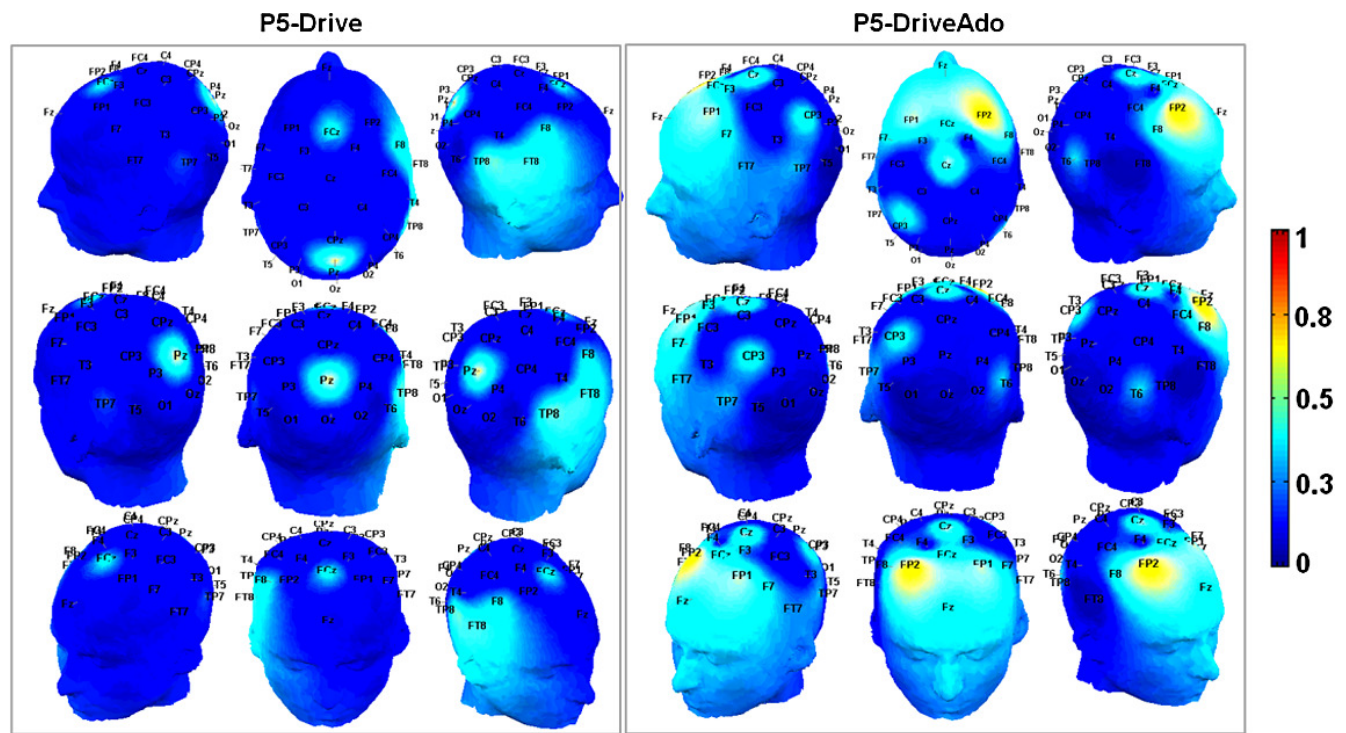


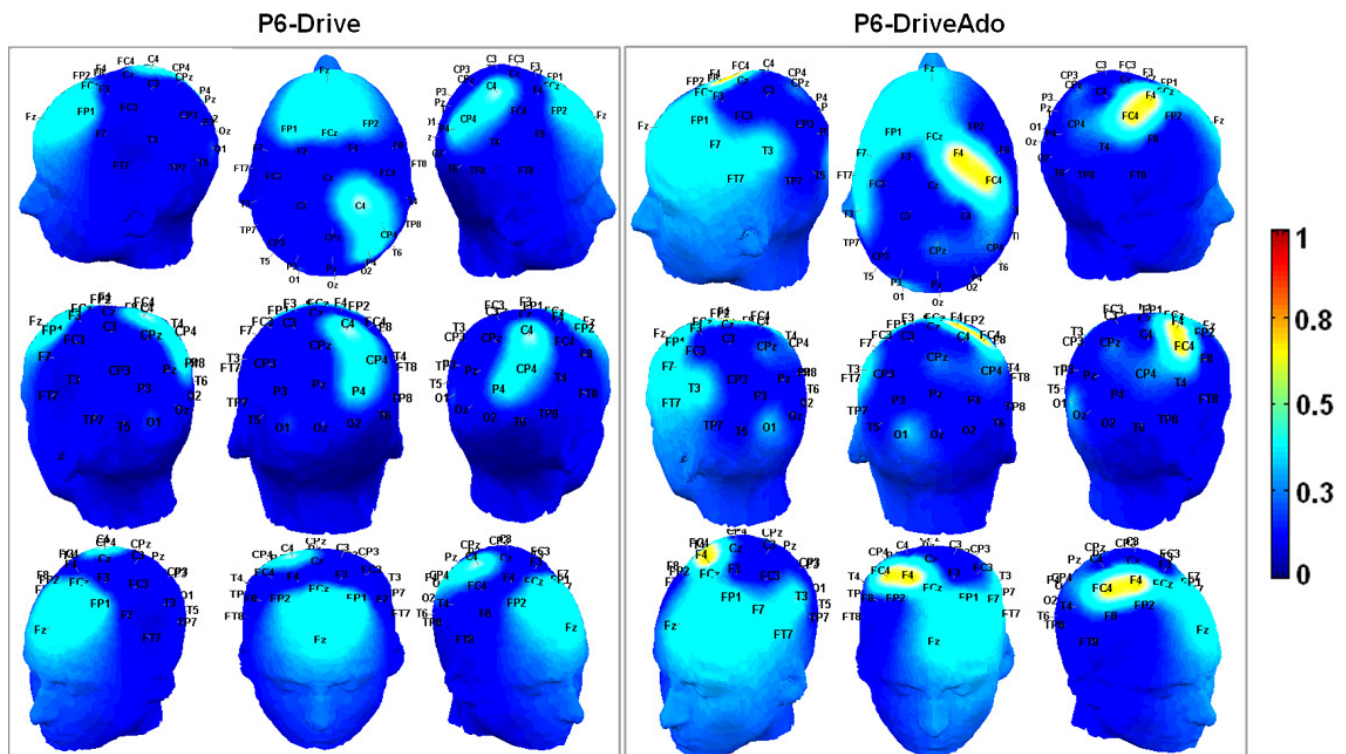
Fig. 10. (Continued.)

of the electrode sites in the frontal regions have a greater number of *DCs*. This highlights the fact that the frontal electrode sites contributed more effective information integration during a heavy cog-

nitive load state, which is essential for cognitive activities. Overall, the *SPN* based *FBN* analytic technique demonstrated its efficacy in detecting cognitive load changes in brain activity. The techniques



(e)



(f)

Fig. 10. (Continued.)

described here have potential for the clinical diagnosis of cognitive impairments. Such research could be further extended to identify connected components (community structures) for the analysis and

understand of the topology of *SPN* networks of different cognitive load states.

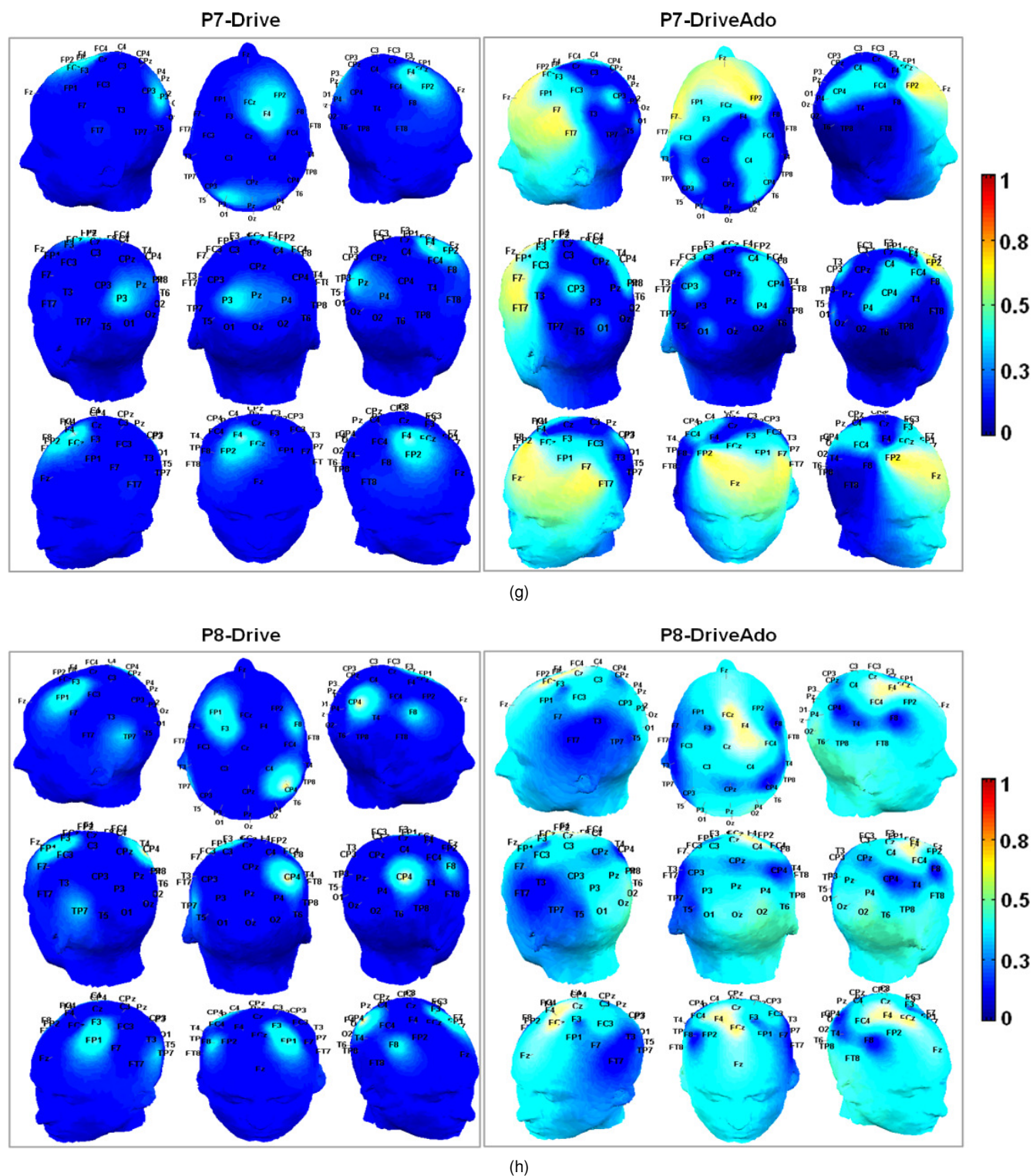


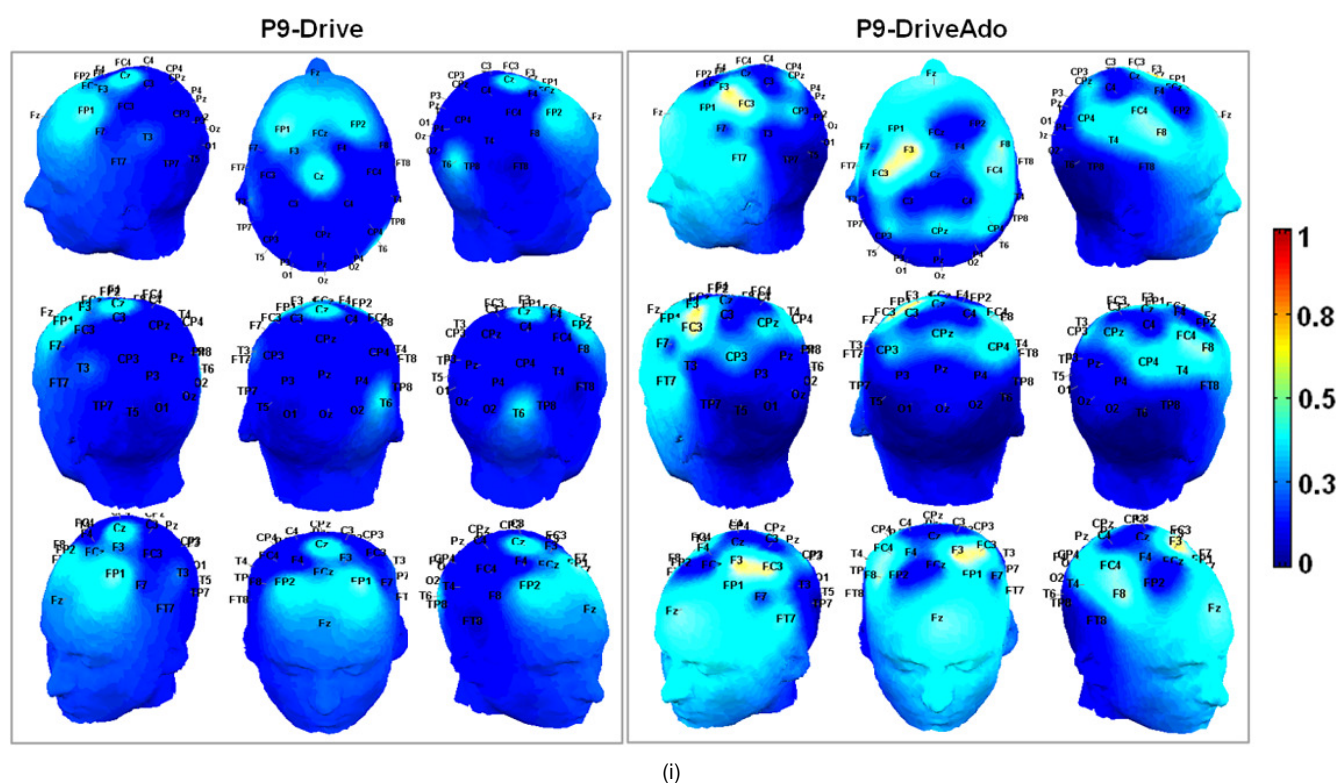
Fig. 10. (Continued.)

Acknowledgments

This work has been carried out in collaboration with the Cognitive Neuro Engineering & Computational Neuroscience Laboratory (CNeL), University of South Australia, Australia.

Conflict of Interest

All authors declare no conflicts of interest.



References

- ## References
- [1] Cocks B, Nandagopal N, Vijayalakshmi R, Thilaga M, Dasari N, Dahal N (2013) Breaking the camel's back: can cognitive overload be quantified in the human brain? *Procedia-Social and Behavioral Sciences* **97**, 21-29.
 - [2] Nandagopal D, Vijayalakshmi R, Cocks B, Dahal N, Dasari N, Thilaga M (2015) Computational neuroengineering approaches to characterise cognitive activity in EEG data. *Knowledge-Based Information Systems in Practice*, 115-137.
 - [3] Vijayalakshmi R, Nandagopal D, Dasari N, Cocks B, Dahal N, Thilaga M (2015) Minimum connected component—a novel approach to detection of cognitive load induced changes in functional brain networks. *Neurocomputing* **170**, 15-31.
 - [4] Dasari NM, Nandagopal ND, Ramasamy V, Cocks B, Thomas BH, Dahal N, Gaertner P (2015) Moment to moment variability in functional brain networks during cognitive activity in EEG data. *Journal of Integrative Neuroscience* **14**(3), 383-402.
 - [5] Thilaga M, Vijayalakshmi R, Nadarajan R, Nandagopal D (2016) A novel pattern mining approach for identifying cognitive activity in EEG based functional brain networks. *Journal of Integrative Neuroscience* **15**(2), 223-245.
 - [6] Vijayalakshmi R, Nandagopal DN, Thilaga M, Cocks B (2015) Characterisation of cognitive activity using minimum connected component. *International Conference on Neural Information Processing*, 531-539.
 - [7] Stam CJ, Reijneveld JC (2007) Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomedical Physics* **1**(1), 3.
 - [8] Bullmore E, Sporns O (2009) Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* **10**(3), 186-198.
 - [9] Sporns O (2009) *The Human Brain Network*. In, S. Boccaletti, V. Latora and Y. Moreno (eds.) *Handbook on Biological Networks* (pp. 199-216). World Scientific Publishing.
 - [10] Rubinov M, Sporns O (2010) Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* **52**(3), 1059-1069.
 - [11] Demuru M, Fara F, Fraschini M (2013) Brain network analysis of EEG functional connectivity during imagery hand movements. *Journal of Integrative Neuroscience* **12**(4), 441-447.
 - [12] Ramachandran VS (2002) *Encyclopedia of the Human Brain, Four-Volume Set*. San Diego, Calif., Academic Press.
 - [13] Bullmore E, Sporns O (2012) The economy of brain network organization. *Nature Reviews Neuroscience* **13**(5), 336-349.
 - [14] Goñi J, van den Heuvel MP, Avena-Koenigsberger A, de Mendizabal NV, Betzel RF, Griffa A, Hagmann P, Corominas-Murtra B, Thiran JP, Sporns O (2014) Resting-brain functional connectivity predicted by analytic measures of network communication. *Proceedings of the National Academy of Sciences* **111**(2), 833-838.
 - [15] Bassett DS, Bullmore E (2006) Small-world brain networks. *The Neuroscientist* **12**(6), 512-523.
 - [16] Thilaga M, Vijayalakshmi R, Nadarajan R, Nandagopal D, Cocks B, Archana C, Dahal N (2015) A heuristic branch-and-bound based thresholding algorithm for unveiling cognitive activity from EEG data. *Neurocomputing* **170**, 32-46.
 - [17] De Vico FF, Richiardi J, Chavez M, Achard S (2014) Graph analysis of functional brain networks: practical issues in translational neuroscience. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **369**(1653), 20130521.

- [18] Kornbrot D (2014) Pearson Product Moment Correlation. In: Michalos AC (eds) *Encyclopedia of Quality of Life and Well-Being Research* (pp. 202-310). Brandon, Canada.
- [19] Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M (2004) Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology* **115**(10), 2292-2307.
- [20] Estévez PA, Tesmer M, Perez CA, Zurada JM (2009) Normalized mutual information feature selection. *IEEE Transactions on Neural Networks* **20**(2), 189-201.
- [21] Steuer R, Kurths J, Daub CO, Weise J, Selbig J (2002) The mutual information: detecting and evaluating dependencies between variables. *Bioinformatics* **18**(2), S231-S240.
- [22] Jalili M, Knyazeva MG (2011) Constructing brain functional networks from EEG: partial and unpartial correlations. *Journal of Integrative Neuroscience* **10**(2), 213-232.
- [23] Simpson SL, Bowman FD, Laurienti PJ (2013) Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Statistics Surveys* **7**, 1-36.
- [24] Van Wijk BC, Stam CJ, Daffertshofer A (2010) Comparing brain networks of different size and connectivity density using graph theory. *Plos One* **5**(10), e13701.
- [25] Boersma M, Smit DJ, Boomsma DI, De Geus EJ, Deleamarre-van de Waal HA, Stam CJ (2013) Growing trees in child brains: graph theoretical analysis of electroencephalography-derived minimum spanning tree in 5- and 7-year-old children reflects brain maturation. *Brain Connectivity* **3**(1), 50-60.
- [26] Bassett DS, Bullmore ET (2017) Small-world brain networks revisited. *The Neuroscientist* **23**(5), 499-516.
- [27] Meier J, Tewarie P, Van Mieghem P (2015) The union of shortest path trees of functional brain networks. *Brain Connectivity* **5**(9), 575-581.
- [28] Kempe D, Kleinberg J, Tardos É (2003) Maximizing the spread of influence through a social network. In: KDD-2003 Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (pp. 137-146), ACM.
- [29] Rahman SA, Schomburg D (2006) Observing local and global properties of metabolic pathways: 'load points' and 'choke points' in the metabolic networks. *Bioinformatics* **22**(14), 1767-1774.
- [30] Bast H, Funke S, Sanders P, Schultes D (2007) Fast routing in road networks with transit nodes. *Science* **316**(5824), 566.
- [31] Pastor-Satorras R, Vespignani A (2007) *Evolution and structure of the Internet: A statistical physics approach*. Cambridge, U.K., Cambridge University Press.
- [32] Bassett DS, Bullmore ET, Meyer-Lindenberg A, Apud JA, Weinberger DR, Coppola R (2009) Cognitive fitness of cost-efficient brain functional networks. *Proceedings of the National Academy of Sciences* **106**(28), 11747-11752.
- [33] Freeman LC (1978) Centrality in social networks conceptual clarification. *Social Networks* **1**(3), 215-239.
- [34] Managbanag J, Witten TM, Bonchev D, Fox LA, Tsuchiya M, Kennedy BK, Kaerberlein M (2008) Shortest-path network analysis is a useful approach toward identifying genetic determinants of longevity. *Plos One* **3**(11), e3802.
- [35] Rezvani A, Meybodi MR (2015) Sampling social networks using shortest paths. *Physica A: Statistical Mechanics and its Applications* **424**, 254-268.
- [36] Kaplan AY, Fingelkurts AA, Fingelkurts AA, Borisov SV, Darkhovsky BS (2005) Nonstationary nature of the brain activity as revealed by EEG/MEG: methodological, practical and conceptual challenges. *Signal Processing* **85**(11), 2190-2212.
- [37] Sanei S, Chambers J (2013) *EEG Signal processing*. Somerset, Wiley.
- [38] Kamel N, Malik AS (2017) *EEG/ERP Analysis: Methods and Applications*. CRC Press.
- [39] Hilfiker P, Egli M (1992) Detection and evolution of rhythmic components in ictal EEG using short segment spectra and discriminant analysis. *Electroencephalography and Clinical Neurophysiology* **82**(4), 255-265.
- [40] Kaplan AY, Shishkin SL (2000) Application of the change-point analysis to the investigation of the brain's electrical activity. *Non-parametric Statistical Diagnosis* **509**, 333-388.
- [41] Pardey J, Roberts S, Tarassenko L, Stradling J (1996) A new approach to the analysis of the human sleep/wakefulness continuum. *Journal of Sleep Research* **5**(4), 201-210.
- [42] Olbrich E, Achermann P, Meier P (2003) Dynamics of human sleep EEG. *Neurocomputing* **52**, 857-862.
- [43] Barrett AB, Murphy M, Bruno MA, Noirhomme Q, Boly M, Laureys S, Seth AK (2012) Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia. *Plos One* **7**(1), e29072.
- [44] Van Drongelen W (2011) *Signal Processing for Neuroscientists: An Introduction to the Analysis of Physiological Signal*. In: Van Drongelen W (Eds.) *Signal Processing for Neuroscientists*. Amsterdam, Elsevier.
- [45] Dijkstra EW (1959) A note on two problems in connexion with graphs. *Numerische Mathematik* **1**(1), 269-271.
- [46] Fornito A, Zalesky A, Bullmore E (2016) *Fundamentals of brain network analysis*. Amsterdam, Elsevier.