35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Functional Dependence in the Human Brain: A Graph Theoretical Analysis†

Bilal H. Fadlallah[1], Andreas Keil[2] and José C. Príncipe[1]

Abstract— **In this paper, we propose a graph-theoretical approach to reveal patterns of functional dependencies between different scalp regions. We start by computing pairwise measures of dependence from dense-array scalp electroencephalographic (EEG) recordings. The obtained dependence matrices are then averaged over trials and further statistically processed to provide more reliability. Graph structure information is subsequently extracted using several graph theoretical measures. Simple measures of node degree and clustering strength are shown to be useful to describe the global properties of the analyzed networks. More sophisticated measures, such as betweenness centrality and subgraph centrality tend to provide additional insight into the network structure, and therefore robustly discriminate two cognitive states. We further examine the connected components of the graph to identify the dependent functional regions. The approach supports dynamicity in that all suggested computations can be easily extended to different points in time, thus enabling to monitor dependence evolution and variability with time.**

I. INTRODUCTION

It has long been known that the human brain represents an immensely complicated network. Hence, it is only natural to consider applying network and graph analysis methodologies as means to bring better understanding of its underlying mechanisms. A good recent review outlining research efforts done on structural and functional brain networks can be found in [1].

In this paper, we show how graph theoretical concepts can be useful in assessing functional brain connectivity within a context using measures of statistical dependence. Measures of dependence have been suggested in the literature to quantify dependencies in neural data. Examples include correlation [2], [3], mutual information [4], [5] and Granger causality [6], [7]. Recently in our lab, we have explored the generalized measure of association or GMA as a measure that quantifies dependence between realizations of random variables. The concept was further developed to handle temporal dependence using an approach we termed the "time series generalized measure of association" or TGMA [8]. TGMA was applied on EEG time series to extract relational information between different recordings [9], [10]. In these studies, dependence was computed with respect to a reference channel. This assumption is relaxed in this paper, where all pairwise links are considered.

The rest of the paper is organized as follows. In Section II, we briefly outline the experimental methodology and the procedure followed to generate the dependence graph from processed EEG data, and in Section III, we detail different graph theoretical measures that can be used to characterize the structure of the obtained graphs. The experimental results are presented in Section IV and Section V offers discussion and concluding remarks.

II. EXPERIMENTAL PROCEDURE

A. Setting

The experimental setting is similar to the one described in [10] and included one participant, a male graduate student in his early twenties. Neuroelectric activity was recorded from 129 channel locations of a Hydro-Cell Geodesic Sensor Net (HCGSN) data acquisition system [11]. Each recording consisted of 4.2 seconds after stimulus onset, with a sampling rate $F_s = 1000$ Hz, and 15 trials were performed in which two different types of pictures were presented to the participant. The first shows a neutral human face (F) and the second a Gabor patch (G) with matched luminance, contrast and mean spatial frequency. The subject was explicitly asked not to blink during the recordings. Offline, the effect of volume conduction was minimized using surface Laplacian methods. Fig. 1 schematically illustrates the experimental design.

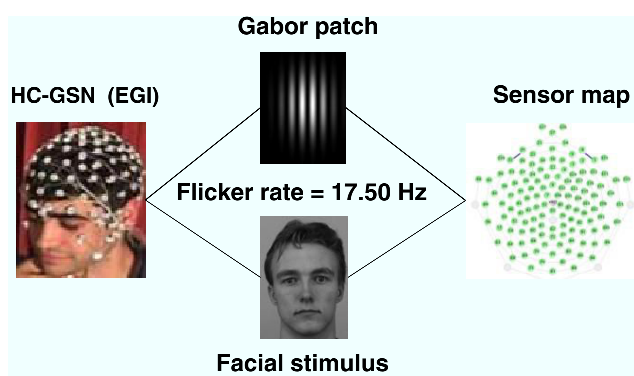

Fig. 1: Experimental setting using a HydroCell Geodesics Sensor Networks system from Electrical Geodesics, Inc. (Riverfront Research Park, 1600 Millrace Drive, Suite 307, Eugene, OR 97403, USA).

† This work was supported by the National Science Foundation under Grant IIS-0964197 and the Lebanese National Council for Scientific Research (CNRS).

1 B. Fadlallah and J. Príncipe are with the Department of Electrical and Computer Engineering (Computational NeuroEngineering Laboratory), University of Florida, Gainesville, FL 32611, U.S.A. E-mail: {bhf, principe}@cnel.ufl.edu

2 A. Keil is with the Department of Psychology (NIMH Center for the Study of Emotion & Attention), University of Florida, Gainesville, FL 32611, U.S.A. E-mail: akeil@ufl.edu

B. The Dependence Graph

Instrumentation noise is first removed by placing notch filters at the locations of the spikes in the recorded signal spectrum. Using adequate bandpass filtering, we concentrate on the relatively narrow band centered around 17.5 Hz as per the thorough analysis in [9]. Prior to computing TGMA values, we embed each processed time series in $\tau = 8$ dimensions to account for the propagation delay among neighboring channels. A dependence graph can be then constructed using the computed dependence values. The motivation behind using such dependence graphs is their proven usefulness in describing dependence relations between random variables [12], and neural sources [13]. We model the electrodes network as a complete undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of vertices or nodes and $\mathcal{E}$ is the set of edges. We assume that the graph is undirected since our goal is to quantify dependencies between brain regions, which does not account for direction. For each edge e_{ij} between two vertices i and j, we assign a value m_{ij} representing the dependence between the processed signals recorded at the electrodes locations. Note that for correlation, m_{ij} exists in the interval $[-1, 1]$, whereas for TGMA it lies in the interval $[0.5, 1]$. As a result, and after averaging over trials, a $129 \times 129 \times 38$ matrix of pairwise dependence values $\mathbf{M}_d^c$ is generated for each dependence measure d and condition c, where 38 refers to the number of frames or time windows (Note that 38 windows correspond to $38 \times 114 = 4332$ samples, which exceeds the initial 4200 samples due to the convolution with the bandpass filter term). The graph $\mathcal{G}$ is further transformed into an incomplete graph by discarding values falling below a predefined statistical threshold (more details in Section IV), which results in a thresholded adjacency matrix ${}^t\mathbf{M}_d^c$.

III. GRAPH THEORETICAL CONCEPTS

A. Basic Notations

The neighborhood of a vertex $v \in \mathcal{V}$ is the set $\mathcal{N}_v$ of all vertices (or neighbors) connected to v, i.e $\mathcal{N}_v = \{r : e_{rv} \in \mathcal{E} \ or \ m_{rv} \neq 0\}$. The degree of v consists of the number of vertices that are incident to v, i.e. $n_v = |\mathcal{N}_v|$, where $|.|$ denotes the cardinality of a set. A path from vertex r to vertex s is a sequence of vertices and edges that begins with r and ends with s, with an edge connecting each vertex with the succeeding one. The distance $d_{r,s}$ between r and s is the minimum length of any path connecting the two vertices. We hereby present some graph theoretical measures of interest for our analysis.

B. Node Clustering Coefficient

The clustering coefficient is a frequently used measure to characterize the local and global structure of unweighted graphs [14]. The clustering coefficient measures the extent to which nodes in a graph tend to cluster together. It is defined for a node i in the graph as:

$$CC_i = \frac{|\{e_{jk} : v_j, v_k \in \mathcal{N}_i, e_{jk} \in \mathcal{E}\}|}{n_i(n_i - 1)} \tag{1}$$

In Eq. 1, $n_i = |\mathcal{N}_i|$. The clustering coefficient of a vertex can be interpreted in terms of its tendency to promote connections among its neighborhood, and can therefore be considered as an indicator of information flow in dynamic networks [19]. It assumes values between 0 and 1.

C. Measures of Centrality

A vertex's importance in a graph can be quantified using several measures, including measures of centrality.

Betweenness Centrality: The betweenness centrality of a vertex i measures the number of shortest paths traversing that vertex. It was first proposed by Anthonisse [15] and has been later used in several contexts [16], [18], [17]. It can be defined as:

$$B_i = \sum_{j \neq i \neq k} \frac{\text{No. shortest paths from } j \text{ to } k \text{ via } i}{\text{No. shortest paths from } j \text{ to } k} \tag{2}$$

When the betweenness centrality of a vertex is high, the vertex is more likely to be an intermediate communication node in the graph. Such vertices can be seen as occupying the "structural holes" in the network [19], [20].

Subgraph Centrality: The subgraph centrality of a vertex i can be defined as the weighted sum of closed walks having different lengths, starting and ending at i.

$$S_i = \sum_{k=0}^{\infty} \frac{\mu_i^k}{k!} \tag{3}$$

where in Eq. 3, μ_i^k refers to the k^{th} local spectral moment, which defines the number of closed walks of length k, starting and ending on i. μ_i^k is computed using the i^{th} diagonal entry of the k^{th} power of the graph adjacency matrix.

$$\mu_i^k = \left({}^t\mathbf{M}_d^c\right)_{ii}^k \tag{4}$$

Due of space limitations, the computational details pertaining to subgraph centrality are omitted and can be found in [21].

Closeness Centrality: The closeness centrality of a vertex i is defined as the inverse of the sum of distances between i and all other vertices.

$$C_i = \frac{1}{\sum_{j \in \mathcal{V}} d_{i,j}} \tag{5}$$

D. Local Efficiency

The local efficiency of a vertex i can be defined in terms of the sum of inverse distances between the vertices in $\mathcal{N}_i$. The local efficiency mainly measures the efficiency in communication between the direct neighbors of i, when the node itself is removed.

$$E_i^{loc} = \frac{1}{n_i(n_i - 1)} \sum_{j,k \in \mathcal{N}_i} \frac{1}{d_{j,k}} \tag{6}$$

E. Connected Components

A connected component of an undirected graph is a maximal subgraph in which any two vertices are connected to each other by paths. The concept is graphically illustrated in Fig. 2.

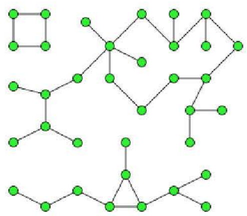

Fig. 2: An undirected graph consisting of three connected components.

IV. RESULTS

For each measure of dependence, we define $\mathbf{M}_d$ as the difference between matrices pertaining to each condition, i.e. $\mathbf{M}_d = \mathbf{M}_d^F - \mathbf{M}_d^G$. It is then possible to visualize how the differences in pairwise dependence values vary in sensor space across time. The resulting matrices are first processed for statistical significance by discarding all values falling within 2 standard deviations of their means, i.e. an edge is drawn between a given pair of vertices if the computed dependence between the two is significantly different than the mean dependence, at a confidence level of 2 standard deviations. Two representative measures of dependence are used in our simulations: the first is the time series generalized measure of association (TGMA) [8], and the second is Spearman's rho, a nonparametric measure of correlation. For each measure of dependence, we compute five node quantities, namely the degree, clustering coefficient, betweenness centrality, subgraph centrality and local efficiency. The absolute value of Spearman's rho is considered since anticorrelation also implies dependence. Computations are made per windows of time corresponding to 114 ms, to achieve a better time resolution and allow tracking any time-varying activity. Alternatively, the whole time series can be used to obtain mean assessment values. The obtained values are visualized in sensor space to identify the active regions involved. Fig. 3 shows the resulting plots (averaged over the time windows) for TGMA and Fig. 4 those for Spearman's rho.

In Fig. 5, we show how betweenness centrality changes with time. Each of the displayed subplots corresponds to computations extracted from 8 subsequent windows or approximately 1 second of data. We can observe a consistency in the active regions across time, especially towards the later time windows, which suggests the reinforcement of communication between sources as time passes.

Fig. 6 shows the connected components corresponding to each condition. The size of the corresponding connected component was plotted for each channel. The size of the main connected component for the "Face" condition is substantially higher than that of the "Gabor patch" condition. Fig. 7 maps the different in active regions to the sensor space.

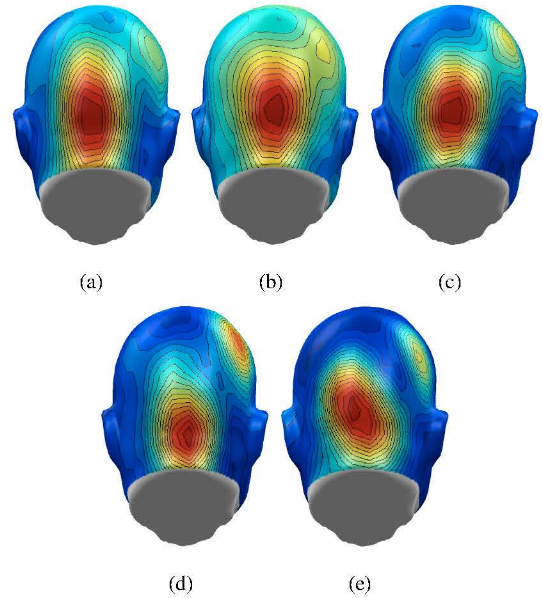

Fig. 3: Several graph theoretical measures extracted from two dependence graphs constructed using TGMA. First, we compute the difference between the two weighted adjacency matrices corresponding to the face and Gabor patch conditions (averaged over trials), and discard non-significant values. We then use graph theoretical measures to characterize the importance of each node. The measures were visualized in a heat map format where red denotes regions of more pronounced differences, and include from left to right: (a) the node degree, (b) the node clustering coefficient, (c) the node betweenness centrality, (d) the node subgraph centrality, and (e) the node local efficiency. Computations used windows of 114 samples and plots were averaged over 38 windows.

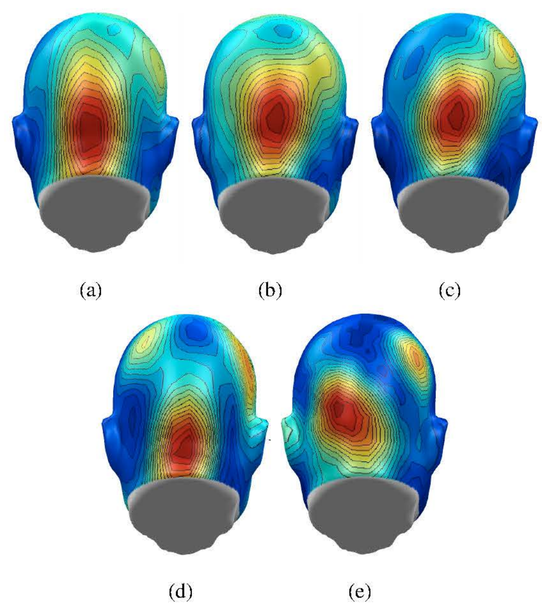

Fig. 4: Several graph theoretical measures extracted from two dependence graphs constructed using Spearman's rho, and displayed in heat map format. The same steps and graph theoretical measures described in Fig. 3 are used.

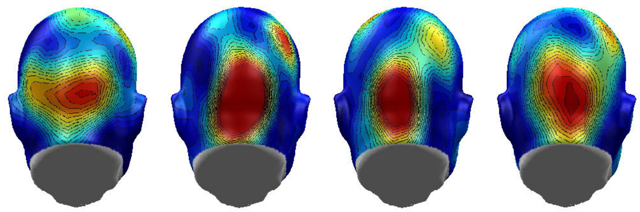

Fig. 5: Betweenness centrality for a graph constructed using TGMA. The subplots are sequential from left to right and each subplot represents approximately 1 sec of data or 8 windows of time. This corresponds to 32 time windows out of the 38 (alternatively the first 3.6 sec out of the 4.2 sec).

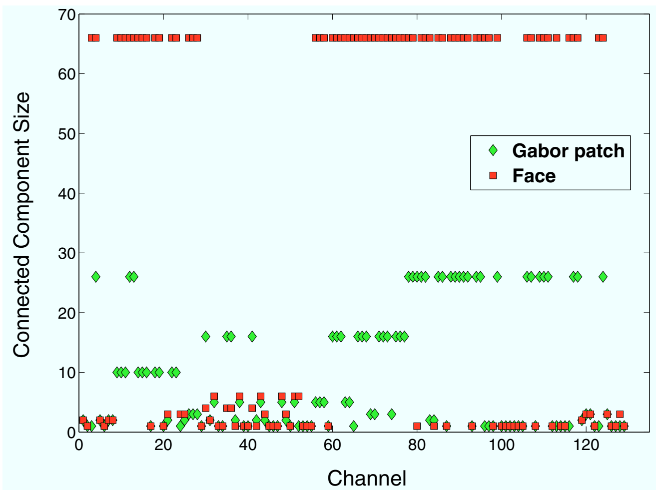

Fig. 6: The size of the connected components per channel for the face and Gabor patch conditions. The dependence measure used is TGMA, and the dependence matrices were averaged over trials and time windows, then statistically processed to discard values falling within two standard deviations of the mean. The connected components were then extracted from the resulting adjacency matrix and the size of the connected component at each node is displayed.

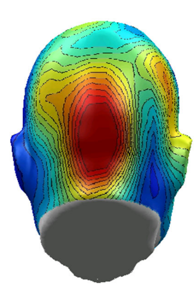

Fig. 7: The cardinality of the connected components shown in Fig. 6 mapped to sensor space.

V. CONCLUSION

This paper demonstrated using EEG data that it is possible to use graph theoretical concepts in order to identify the active recording sites when a human subject performs a cognitive task. Dependence graphs were generated from estimated dependence between pairwise channels, that were further statistically processed. Several notions from graph theory were applied to examine the network structure. They include the node degree, clustering coefficient, local efficiency, besides measures of centrality (betweenness and subgraph centrality), and an analysis of connected components. Results show that the extracted connected components differ based on the presented stimulus, and allow robust discrimination between the conditions of interest. Measures of centrality perform best among the considered strategies, in that their discriminatory schemes have more signal-to-noise ratio (SNR) and less spurious fluctuations. Both measures of dependence concurred on the identification of active regions. As future work, it would be interesting to validate the current methodology on more subjects to solidify the above conclusions.

REFERENCES

[1] E. Bullmore, "Complex brain networks: graph theoretical analysis of structural and functional systems", *Nat. Rev. Neurosci.*, vol. 10, no. 3, pp. 186-198, 2009.

[2] K. Schindler, H. Leung, C. E. Elger and K. Lehnertz, "Assessing seizure dynamics by analysing the correlation structure of multichannel intracranial eeg", *Brain*, vol. 130, no. 1, pp. 65-77, 2007.

[3] M. Guevaraa and M. Corsi-Cabrera, "EEG coherence or EEG correlation?", *Int. J. Psychophysiol.*, vol. 23, no. 3, pp. 145-153, 1966.

[4] J. Jeong, J. Gore and B. Peterson, "Mutual information analysis of the EEG in patients with Alzheimer's disease", *Clin. Neurophysiol.*, vol. 112, no. 5, pp. 827-835, 2001.

[5] S. Na, S. Jin, S. Kim and B. Ham, "EEG in schizophrenic patients: mutual information analysis", *Clin. Neurophysiol.*, vol. 113, no. 12, pp. 1954-1960, 2002.

[6] W. Hesse, E. Moller, M. Arnold and B. Schack, "The use of time-variant EEG Granger causality for inspecting directed interdependencies of neural assemblies", *J. Neurosci. Methods*, vol. 124, no. 1 pp. 27-44, 2003.

[7] M. Ding, Y. Chen and S. Bressler, "Granger causality: basic theory and application to neuroscience", *Biol. Cybern.*, vol. 85, no. 2, pp. 145-157, 2006.

[8] B. Fadlallah, A. Brockmeier, S. Seth, L. Li, A. Keil and J. Principe, "An association framework to analyze dependence structure in time series", *Proceedings of the* 34^{th} *International Conference of the IEEE EMBS*, pp. 6176-6179, 2012.

[9] B. Fadlallah, S. Seth, A. Keil and J. Principe, "Robust EEG preprocessing for dependence-based condition discrimination", *Proceedings of the* 33^{rd} *International Conference of the IEEE EMBS*, pp. 1407-1410, 2011.

[10] B. Fadlallah, S. Seth, A. Keil and J. Principe, "Quantifying cognitive state from EEG using dependence measures", *IEEE Trans. Biomed. Eng.*, vol. 59, no. 10, pp. 2773-2781, 2012.

[11] Electric Geodesics Inc., "Geodesic sensor net technical manual", Online: `http://www.egi.com`, pp. 29-30, 2007.

[12] S. Lauritzen, "Graphical Models", 6th edition, *Clarendon Press, Oxford*, 1996.

[13] C. Stam and J. Reijneveld, "Graph theoretical analysis of complex networks in the brain", *Nonlinear Biomed. Phys.*, vol. 1, no. 3, pp. 186-198, 2007.

[14] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D. Hwang, "Complex networks: structure and dynamics", *Physics Reports*, vol. 424, no. 4, pp. 175-308, 2006.

[15] J. Anthonisse, "The rush in a directed graph", *Stichting Mathematisch Centrum. Mathematische Besliskunde*, (BN 9/71), pp. 1-10, 1971.

[16] L. Freeman, "A set of measures of centrality based on betweenness", *Sociometry*, pp. 35-41, 2007.

[17] K. Goh, E. Oh, B. Kahng and D. Kim, "Betweenness centrality correlation in social networks", *Physical Review E*, vol. 67, no. 1, 017101, 2003.

[18] M. Barthelemy, "Betweenness centrality in large complex networks", *The European Physical Journal B-Condensed Matter and Complex Systems*, vol. 38, no. 2, pp. 163-168, 2004.

[19] L. Lopez-Fernandez, G. Robles, J. Gonzalez-Barahona and I. Herraiz, "Applying social network analysis techniques to community-driven libre software projects", *International Journal of Information Technology and Web Engineering (IJITWE)*, vol. 1, no. 3, pp. 27-48, 2006.

[20] R. Burt, "Structural holes: The social structure of competition", *Harvard University Press*, 1995.

[21] E. Estrada and J. Rodriguez-Velazquez, "Subgraph centrality in complex networks", *Physical Review E*, vol. 71, no. 5, 056103, 2005.