Querying Functional Brain Connectomics to Discover Consistent Subgraph Patterns

Nantia D. Iakovidou, Stavros I. Dimitriadis, Nikos A. Laskaris and Kostas Tsichlas

***Abstract*— Dynamic recordings of functional activity maps can naturally and efficiently be represented in the form of functional/effective connectivity networks. New methods for mapping synaptic connections and recording neural signals generate rich and complex data about the structure and dynamics of brain networks. To study the most complex network in nature, the brain, there is need to integrate a huge amount of brain networks collected from laboratories over the world in large databases. Human Brain Project (Europe and USA) aims to explore brain functionality in various ways. Brain networks are central to achieving the goals of this ambitious plan. However, the immense amount of thousands of brain networks, prevent an easy way to utilizable knowledge.**

In this paper, we demonstrate a data-driven approach that discovers consistent patterns from a collection of brain networks via a querying approach: formulating a query of "finding an increasing or a decreasing consistent subgraph over an amount of subjects" after taking the difference between two sets of graphs referred as two conditions (an active and a baseline). Experiments demonstrated that our data-driven approach allows identifying frequency-dependent selective spatial pattern changes of the EEG functional connectivity network during a mental task. This is the first time that a method fully exploits the connectivity weights of a brain network to discover consistent subgraph patterns.

I. INTRODUCTION

With the arrival of new methodologies for deriving different modes of connectivity (anatomical, functional and effective) from neuroimaging data, across multiple spatiotemporal scales, neuroscience appears to initiate its own revolution. The Human Genome Project gave the baton to Human Brain Project (Europe [1] & USA [2]) which aims through connectomics to map the brain architecture and also to answer how structural networks evolved in time shaping brain functionality [3]. To address these challenges, both continental projects must include intense efforts directed at collecting consistent recording sets, data analysis and modeling employing network science supported from brain theory. Catalogs of functional activation maps recorded with noninvasive neuroimaging methods in the human brain are available as public repositories that can be data-mined and analyzed online [4, 5]. These databases have enabled important insights regarding the link between functional activations, networks of coactivation patterns and functional connectivity. Additionally, these functional network maps have been linked to various domains of behavior and cognition. The question is how we can detect consistent patterns related to a cognitive task or aberrant patterns related to a brain disease/disorder.

Several methods to analyze brain function from EEG recordings have been proposed in the past years [6, 7]. They range from traditional linear methods, such as Fourier transforms and spectral analysis [8], to nonlinear methods derived from the theory of nonlinear dynamical systems, also called chaos theory [9, 10]. In general, chaos-based approaches outperform the traditional linear methodologies [11], which assume that the signal is stationary and originates from a low dimensional linear system. However, in reality this is not always the case, because a real EEG is a non-stationary signal [12]. Moreover, given that EEGs are recorded continuously and many times over long periods of time, this often leads to the production of vast amounts of data and makes the task of knowledge discovery from them quite difficult but at the same time challenging, as well. It is thus very important to develop new methods in order to study EEG signals in different physiological and pathological states.

For these reasons, there has been a focus on studying EEG signals as graphs (networks) during the last few years [13, 14, 15, 16]. In particular, there is a growing interest in the theoretical aspects of network analysis in an attempt to model [17, 18], describe [13] and propose new measures [19] for a better understanding of complex systems, such as the human brain. An important approach though to exploring the rules that governs the structure of complex networks is to study their characteristic building blocks, called "motifs", which is in fact the focus of the present study.

In our previous work [16] we introduced a method to detect brain sub-network patterns using binary static and time-varying graphs of functional connectivity. Our target was to discover group-consistent motifs so as to characterize a particular brain state or a brain state with respect to a reference state. In the current paper, we consider the full

N. D. Iakovidou (corresponding author) is with the Department of Informatics, Data Engineering Laboratory (DELAB), Aristotle University of Thessaloniki, 54124, Greece, phone: +302310991924; fax: +302310991913; e-mail: niakovid@ csd.auth.gr.

S. I. Dimitriadis and N. A. Laskaris are with the Department of Informatics, Artificial Intelligence & Information Analysis Laboratory (AIIA), Aristotle University of Thessaloniki, 54124, Greece.

Kostas Tsichlas is with the Department of Informatics, Data Engineering Laboratory (DELAB), Aristotle University of Thessaloniki, 54124, Greece.

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

exploitation of connectivity weights by introducing a new approach for detecting motifs that indicate increases or decreases of connections between two particular brain states. The well-known gSpan (graph-based Substructure pattern mining) algorithm [20] is also used for mining frequent subgraphs that showed an increasing or decreasing trend. Our study (multivariate) overcomes the limited approach of well-known statistical techniques (e.g. statistical parametric mapping (SPM)), which investigates the change of each voxel/node independently (univariate).

The rest of the paper is organized as follows. Section II presents our novel methodology after a short description of the employed data. Section III describes the adopted experimental procedure and the obtained results. Conclusions are drawn in section IV.

II. Methodology

A. Functional Connectivity Data

The functional connectivity graph dataset was created as an intermediate result in previous studies [21, 22, 14] and hence a detailed description can be found therein. It consisted of two types of graphs that corresponded to two different brain states (resting state and mental calculations) and reflected neural synchrony in each of the five standard frequency bands (θ, α1, α2, β and γ). The original EEG signals had been recorded from 18 healthy adults while they were performing multiplications (active condition) or doing nothing (baseline condition). Starting from the signals recorded (at sampling frequency of 500Hz with 30 sensors covering the head according to International 10-20 system) and filtered within a particular frequency band, we derived Functional Connectivity Graphs (FCGs) by means of a phase synchrony estimator (in particular Phase Locking Value, PLV, [23]), that quantifies functional dependence in pairwise fashion. Each FCG has a tabular format, i.e. a [30 × 30] weighted-adjacency matrix W description, with elements $w_{ij} \in [0, 1]$ denoting the strength of functional association between the i-th and the j-th sensor (and implicitly between the cortical areas underneath). Hence, each graph represents connectivity estimates corresponding to subject's brain activity for long recording periods (8 sec.).

B. Proposed Approach

Assume two sets of weighted, undirected and labeled graphs $\boldsymbol{G_A} = \{G_{A1}, G_{A2}, \ldots, G_{A18}\}$ that corresponds to the active condition and $\boldsymbol{G_B} = \{G_{B1}, G_{B2}, \ldots, G_{B18}\}$ that corresponds to the baseline condition. Henceforth, for simplicity we represent by G_i both the graph and its corresponding weighted-adjacency matrix. Then, based on $\boldsymbol{G_A}$ and $\boldsymbol{G_B}$, we create two new sets of graphs $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$. In particular, $\boldsymbol{G}^+=\{G_{A1}-G_{B1}, G_{A2}-G_{B2}, \ldots, G_{A18}-G_{B18},\}$ where by $G_{Ai}-G_{Bi}$ ($i\in[1,18]$) we represent the subtraction of their corresponding weighted adjacency matrices, by zeroing all negative values. Similarly, we define $G^-=\{G_{A1}-G_{B1}, G_{A2}-G_{B2}, \ldots, G_{A18}-G_{B18}\}$, by zeroing all positive values. Then, the algorithm zeroes all cells of the matrices of the graphs in $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$ whose absolute value is less than a prespecified threshold and sets to one all other cells, whose absolute values are larger than threshold. Our next step is to identify frequently occurring patterns in these sets of graphs with minimum support *x*, where $x\in(0,1]$. That is, these sub-graphs must occur in at least x·N graphs in each set $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$. To provide a solution to this problem we use a robust tool named gSpan [20], which has been experimentally verified to be one of the most effective and efficient algorithms for the graph mining problem.

For this problem, most of the known algorithms follow the brute force approach of candidate generation which is expressed through a breadth-first search of the graphs. gSpan on the other hand, uses a depth-first search approach. Based on the DFS trees of the graphs in each set $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$ we map each graph to a unique minimal sequence of symbols, which is called DFS code. In particular, each node in the DFS trees is assigned a symbol (in fact a tuple of symbols). Sub-graphs to be tested are generated by extending smaller graphs in a DFS manner. The enumeration of sub-graphs is implemented indirectly by generating the corresponding sequences in a lexicographic order. The DFS code is constructed in such a way so that two sub-graphs are isomorphic if and only if the corresponding sequences are equal. In this way, the extraction of frequent sub-graphs from each set $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$ is reduced to extracting frequent patterns from strings.

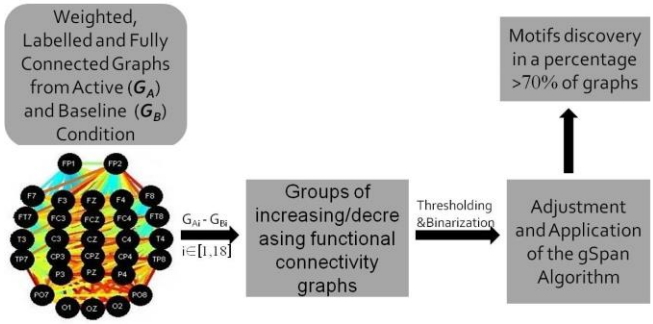

Fig. 1. The Procedure for formulating a query. Each query is referred to consistent changes in connectivity data, while pairs of graphs associated with particular brain states are compared for each subject separately.

III. Experiments and Results

Figure 1 provides a schematic outline of the steps followed in order to lead us to the discovery of consistent subgraph patterns. In our experiments, we used only static

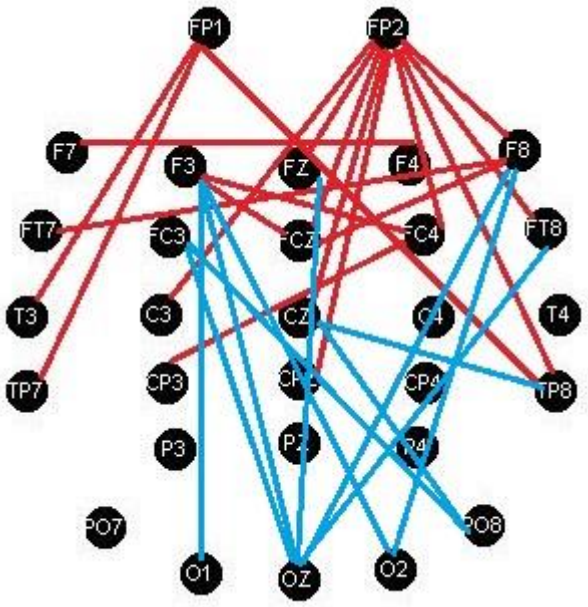

Fig. 2. Motifs (single edges) detected from a1 band in a percentage > 70% of subjects. Red color indicates decrease and blue color indicates increase.

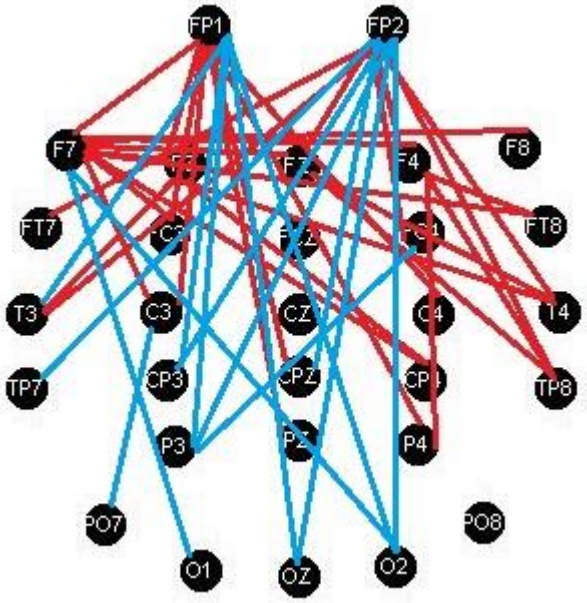

Fig. 3. Motifs (single edges) detected from a2 band in a percentage > 70% of subjects. Red color indicates decrease and blue color indicates increase.

connectivity graphs and each frequency band (defined, conventionally, as θ (4–8 Hz), α1 (8–10 Hz), α2 (10–13 Hz), β (13–30 Hz) and γ (30–45 Hz)), was treated separately. A common thresholding scheme had been applied in all cases after the creation of $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$, for keeping only the highest values of differences in the resulted graphs. After exhaustive experimentation, this threshold was set equal to the highest possible value so that enough edges are retained in the thresholded graphs (threshold=0.05). Then, we continue with the detection of the most frequent motifs in $\boldsymbol{G}^+$ and $\boldsymbol{G}^-$, which could be single edges or larger subgraphs. We call all motifs found in $\boldsymbol{G}^+$ "increasing motifs", since the functional association within them is higher when a person performs multiplications than when doing nothing. In the same way, all motifs found in $\boldsymbol{G}^-$ are called "decreasing motifs".

Task specific motifs that appear to a percentage > 70% of subjects related to the cognitive processes involved in the multiplication were detected in α_1, α_2 and γ bands. In α_1, α_2 bands, increasing motifs were extended in the fronto-parieto-occipital brain regions and decreasing motifs were located over bilateral frontal and temporal brain areas. Finally, in γ-band, characteristic increasing motifs were located between bilateral parieto-occipital and fronto-central brain areas.

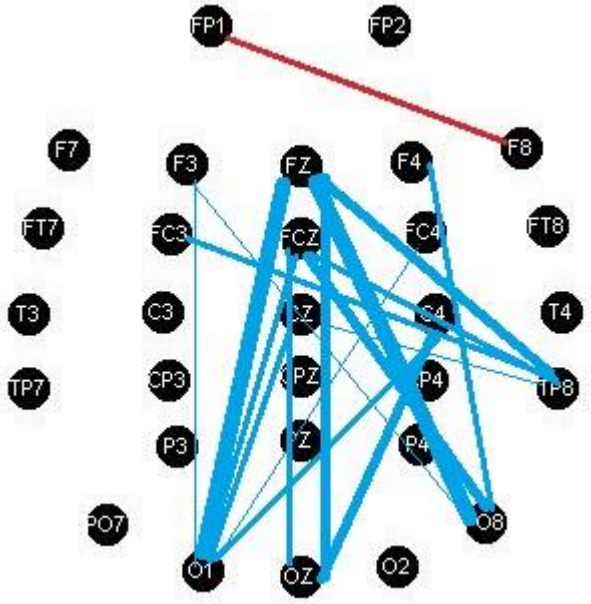

Fig. 4. Results from gamma band in a percentage > 70% of subjects. Red color indicates decrease and blue color indicates increase. The thickness of blue lines indicates frequency of appearance of these edges in the discovered motifs. (The larger the thickness, the bigger the frequency.)

IV. Conclusions

An important methodological contribution in this paper is the contrastive learning scheme for motif extraction, which includes subtraction of two sets of weighted, undirected and labeled graphs (baseline and cognitive condition). Based on this model, we attempted to detect increasing and decreasing related consistent subgraphs, with the use of robust graph mining tools.

Our analysis showed, in $\alpha_{1,2}$-band, increasing motifs that were extended in the fronto-parieto-occipital brain regions and decreasing motifs located over bilateral frontal and temporal brain areas while in γ-band increasing motifs were located between bilateral parieto-occipital and fronto-central

brain areas. These results are in line with earlier EEG findings that reported long-range (fronto-parietal) connections in lower frequency bands and short-range connections in the γ band during a WM task [24]. Finally, the presence of both increasing and decreasing motifs in $\alpha_{1,2}$-band probably means that there is a difference in attentional [24] and memory processes developed via α band activity [25]. A recent study [26], analyzing a two-digit multiplication task revealed a widespread pattern in four lobes of the brain reflecting demanding cognitive functions (working memory, retrieval). Increasing motifs related with γ activity, most likely reflect retrieval of information independent of sensory modality but related to the cognitive mental task and memory processes [27, 28].

Finally, only static graphs were included in the present study, since our primary goal was to ascertain that the introduced technique can provide promising and interesting results. The next goal is to apply the introduced technique to time-varying functional/effective connectivity graphs in order to track the appearance of a motif along time and also to unmask any important dynamic pattern that was covered by employing static FCGs. Additionally, the presented technique can be employed to detect anatomical difference between healthy and diseased subjects based on structural data (magnetic resonance imaging - **MRI**, diffusion spectrum imaging - **DSI**). The above directions will be considered in the near future.

REFERENCES

[1] http://www.humanbrainproject.eu/

[2] http://www.whitehouse.gov/infographics/brain-initiative

[3] O. Sporns. "Making sense of brain network data", Nature Methods, vol.10, no.6, June 2013.

[4] P. T. Fox, A. R. Laird, S. P. Fox, P. M. Fox, A. M. Uecker, M. Crank, et al. "BrainMap Taxonomy of Experimental Design: Description and Evaluation", Human Brain Mapping, vol. 25, pp. 185–198, 2005.

[5] T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, and T. D. Wager. "Large-scale automated synthesis of human functional neuroimaging data", Nature Methods, vol. 8, no. 8, pp. 665–670, June 2011.

[6] C. J. Stam. "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field", Clin. Neurophysiology, vol. 116, no. 10, pp. 2266–2301, 2005.

[7] F. Mormann, R.G. Andrzejak, C.E. Elger and K. Lehnertz. "Seizure prediction: the long and winding road", Brain, vol.130, pp.314–333, 2007.

[8] Z. Rogowski, I. Gath and E. Bental, "On the prediction of epileptic seizures", Biol. Cybern., vol.42, pp. 9–15, 1981.

[9] A. Wolf, J. B. Swift, H. L. Swinney and J. A. Vastano. "Determining Lyapunov exponents from a time series", Physica D., vol. 16, pp. 285–317, 1985.

[10] W. S. Pritchard and D. W. Duke, "Measuring 'chaos' in the brain: a tutorial review of EEG dimension estimation", Brain Cogn. Vol. 27, pp. 353–397, 1995.

[11] M. I. Rabinovich, P. Varona, A. I. Selverston and H. D. Abarbanel. "Dynamical principles in neuroscience", Rev. Mod. Phys., vol. 78, pp. 1213–1265, 2006.

[12] D. Gribkov and V. Gribkova. "Learning dynamics from non-stationary time series: analysis of electroencephalograms", Phys. Rev. E., vol. 61, pp. 6538–6545, 2000.

[13] M. Rubinov and O. Sporns. "Complex network measures of brain connectivity: uses and interpretations", Neuroimage, vol. 52, pp. 1059–1069, 2010.

[14] S.I. Dimitriadis, N.A. Laskaris, V. Tsirka, M. Vourkas, S. Micheloyannis. "An EEG study of brain connectivity dynamics at the resting state", Nonlinear Dynamics, Psychology and Life Sciences, vol. 16, pp. 5-22, January 2012.

[15] C. Schmidt, T. Weiss, C. Komusiewicz, H. Witte and L. Leistritz. "An Analytical Approach to Network Motif Detection in Samples of Networks with Pairwise Different Vertex Labels", Computational and Mathematical Methods in Medicine, vol. 12, 2012.

[16] N. D. Iakovidou, S. I. Dimitriadis, N. A. Laskaris, K. Tsichlas and Y. Manolopoulos. "On the discovery of group-consistent graph substructure patterns from brain networks", Journal of Neuroscience Methods, vol. 213, pp. 204-213, 2013.

[17] S.L. Simpson, M.N. Moussa, P.J. Laurienti. "An exponential random graph modeling approach to creating group-based representative whole-brain connectivity networks", Neuroimage, vol. 60, no. 2, pp.1117-1126, 2012.

[18] P. E.Vértes, A. F. Alexander-Bloch, N. Gogtay, J. N. Giedd, J. L. Rapoport, and E. T. Bullmore. "Simple models of human brain functional networks", Proc. Nat. Acad. Sci., 2012.

[19] K. E. Joyce, P.J. Laurienti, J. H. Burdette, S. Hayasaka. "A New Measure of Centrality for Brain Networks", PLoS ONE, vol. 5, no. 8, 2010.

[20] X. Yan and J. Han. "gSpan: Graph-Based Substructure Pattern Mining", ICDM 2002, pp. 721-724, 2002.

[21] S. I. Dimitriadis, N. A. Laskaris, V. Tsirka, M. Vourkas, S. Micheloyannis and S. Fotopoulos. "Tracking brain dynamics via time-dependent network analysis", Journal of Neuroscience Methods, vol. 193, pp. 145-155, 2010.

[22] S. I. Dimitriadis, N. A. Laskaris, V. Tsirka, M. Vourkas and S. Micheloyannis. "What does delta band tell us about cognitive Processes: a mental calculation study", Neuroscience Letters, vol. 483, pp. 11-15, 2010.

[23] J. P. Lachaux, E. Rodriguez, J. Martinerie and F. J. Varela. "Measuring phase synchrony in brain signals", Hum. Brain Mapp. vol. 8, pp. 194–208, 1999.

[24] A. von Stein and J. Sarnthein. "Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization", Int. J. Psychophysiol., vol. 38, pp. 301–313, 2000.

[25] W. Klimesch, B. Schack and P. Sauseng. "The functional significance of theta and upper alpha oscillations for working memory: a review", Exp. Psychol., vol. 52, pp. 99–108, 2005.

[26] S. Micheloyannis, V. Sakkalis, M. Vourkas, V. Tsirka, E. Karakonstantali, K. Kanatsouli, et al. "The influence of ageing on complex brain networks: a graph theoretical analysis", Hum. Brain Mapp., vol. 30, pp. 200–208, 2009.

[27] C. Tallon-Baudry and O. Bertrand. "Oscillatory gamma activity in humans and its role in object representation", Trends Cogn. Sci., vol. 3, pp. 151–162, 1999.

[28] P. Sauseng and W. Klimesch. "What does phase information of oscillatory brain activity tell us about cognitive processes?" Neurosci. Biobehav. Rev., vol. 32, pp. 1001–1013, 2008.