33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Community detection for directional neural networks inferred from EEG data

Ying Liu, Jason Moser and Selin Aviyente

Abstract— **One major challenge in neuroscience is to identify the functional modules from multichannel, multiple subjects recordings. Most research on community detection has focused on finding the association matrix based on functional connectivity, instead of effective connectivity, thus not capturing the causality in the network. In this paper, we propose a community detection algorithm suitable for weighted and asymmetric (directed) networks representing effective connectivity, and apply the algorithm to multichannel electroencephalogram (EEG) data. In addition, we extend the algorithm to find one common community structure from multiple subjects.**

I. INTRODUCTION

With the advance of neuroimaging technology, such as EEG, it is possible to record brain activity with higher temporal resolution and accuracy than ever before. In order to understand the functioning of the brain better, methods to identify the communities or functional modules from the observed multichannel, multiple subjects recordings have been developed. Functional and effective connectivity are two widely studied measures to quantify the connectivity patterns in the brain. Unlike functional connectivity which only quantifies the statistical dependencies between two processes, effective connectivity quantifies the influence one node exerts on another node. Traditionally, effective connectivity has been quantified using measures of causality, such as Granger causality and partial directed coherence (PDC) [1]. Granger causality based methods are model dependent and limited to detecting the linear relations, however, EEG recordings are known to have nonlinear dependencies between recordings from different sites. Therefore, in our previous work, directed information was proposed to quantify the information flow in the brain network [2]. Unlike Granger causality, directed information is model free and can quantify the nonlinear relations.

Although these measures are effective at quantifying the relationship between pairs of neuronal populations, they do not reveal the actual network structure. In recent years, graph theoretic methods, such as community detection, have been applied to association matrices defined by either functional connectivity or effective connectivity, i.e., to undirected or directed networks. For example, functional MRI data has been used to show that community structure changes with age [3]. However, most of the work so far has focused on undirected networks [4]. Therefore, in this paper, we focus on the graph representation of the brain networks using effective connectivity, which results in a weighted and asymmetric association matrix.

In order to discover the underlying organization of the network, traditional clustering algorithms such as Kernighan-Lin algorithm, agglomerative (or divisive) algorithm, k-means clustering, etc., have been used widely. However, these algorithms need to pre-determine the number of clusters [5]. Therefore, modularity based algorithms are widely used to choose the best partitions of a network by maximizing the modularity, which include greedy techniques such as simulated annealing and spectral optimization [5]. Recently, Blondel introduced a greedy approach for the modularity optimization of the weighted graph, which is proven to be efficient, multi-level and close to the optimal value obtained from slower methods [6]. Meunier applied this method to functional MRI and investigated the hierarchical structure of the functional brain network [7]. Here we extend this algorithm to weighted and directed networks to find the functional communities.

In this paper, we concentrate on a study focused on understanding the cognitive control networks in the brain, in particular those involved in error-processing across multiple subjects. The traditional way to deal with group analysis is to determine the community structure of each subject separately and average the results. However, this method ignores the between-subject variability and the results may be influenced by some outliers [7]. In this paper, instead of maximizing the modularity and identifying the communities of each subject separately, we detect the community structure of a group by maximizing the total modularity of the group.

II. BACKGROUND

A. Modularity

The concept of modularity is motivated by the idea that nodes in the same module have very dense connection and sparse inter-module connection [8]. Modularity is first proposed as a stopping criteria for the Girvan and Newman algorithm [9], but later widely used as a quality function to choose the best partitions of a network. A good partition of a network has high modularity Q, where Q=(fraction of edges within communities)-(expected fraction of such edges) [8]. The original expression of modularity for undirected binary

This work was in part supported by the National Science Foundation under Grants No. CCF-0728984, CAREER CCF-0746971.

Y. Liu and S. Aviyente are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA liuying5@egr.msu.edu, aviyente@egr.msu.edu

J. Moser is with the Department of Psychology, Michigan State University, East Lansing, MI 48824, USA jmoser@msu.edu

networks is given as,

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{ij} - \frac{k_i k_j}{2m} \right] \delta_{c_i, c_j}, \tag{1}$$

where A_{ij} is the adjacency matrix, k_i is the degree of vertex i, δ_{c_i,c_j} is equal to 1 when i and j are in the same community; is equal to 0 otherwise. For a directed network, the probability of a directed edge relies on the in-degree and out-degree of the vertex. Leicht extends the definition of modularity to directed binary networks as [8],

$$Q_d = \frac{1}{m} \sum_{i,j} \left[A_{ij} - \frac{k_i^{in} k_j^{out}}{m} \right] \delta_{c_i, c_j}, \tag{2}$$

where k_i^{in} (k_i^{out}) is the in-degree (out-degree) of vertex i. Arenas gives a more general expression of modularity to directed weighted networks [10],

$$Q_{gen} = \frac{1}{A_W} \sum_{i,j} \left[W_{ij} - \frac{s_i^{in} s_j^{out}}{A_W} \right] \delta_{c_i, c_j}, \tag{3}$$

where $W_{i,j}$ is the weight of edge $e_{i,j}$, s_i^{in} (s_i^{out}) is the in-degree (out-degree) of vertex i, $A_W = \sum W_{i,j}$. In this paper, $W_{i,j}$ is quantified by the directed information measure introduced in the following section.

B. Directed information

Different information measures have been proposed to quantify the causal relationship between two random processes. Directed information (DI) developed by Massey has been proven to be a suitable measure to study the information flow in networks of stochastic processes [11]. The definition of DI for two length N sequences $\mathbf{X} = X^N = X_1, \cdots, X_N$ and $\mathbf{Y} = Y^N = Y_1, \cdots, Y_N$ is as follows:

$$\begin{aligned} DI(X^N \to Y^N) &= \sum_{n=1}^{N} I(X^n; Y_n | Y^{n-1}), \\ &= \sum_{n=1}^{N} [H(X^n Y^{n-1}) - H(X^n Y^n)] + H(Y^N), \\ &= \sum_{n=1}^{N} [I(X^n; Y^n) - I(X^n; Y^{n-1})], \end{aligned} \tag{4}$$

where $X^n = (X_1, \ldots, X_n), Y^n = (Y_1, \ldots, Y_n)$ are length n random sequences.. $I(X;Y)$ is the mutual information between two random variables X and Y. Since $0 < DI(X^N \to Y^N) < I(X^N; Y^N) < \infty$, in practice a normalized version of DI, which maps DI to the $[0, 1]$ range is used for comparing different interactions [12]:

$$\rho_{DI} = \sqrt{1 - e^{-2 \sum_{n=1}^{N} I(X^n; Y_n | Y^{n-1})}}. \tag{5}$$

Based on the definition of DI, the computation of DI requires the estimation of joint probabilities of high dimensional random variables over time. In this paper, directed information estimation based on mutual information is used to estimate the DI directly from EEG data by using adaptive partitioning method [13].

III. Algorithm for community detection

A. Algorithm for community detection for a single graph

In this paper, we extend the method proposed by Blondel to weighted and directed networks [6]. The algorithm consists of two steps. First, for each node i, the gain of the modularity ΔQ_{gen} is computed when the node is assigned to the communities of all other nodes j, $j = 1, \cdots, N$ and $j \neq i$, and the community for which ΔQ_{gen} is highest is chosen. This process is applied sequentially for all nodes. ΔQ_{gen}, which partly determines the efficiency of the algorithm, can be computed as follows,

$$\begin{aligned} \Delta Q_{gen} = &\frac{1}{W} \sum_{p=1}^{N_j} \left(A_{i,j_p} - \frac{s_i^{out} s_{j_p}^{in}}{W} + A_{j_p,i} - \frac{s_i^{in} s_{j_p}^{out}}{W} \right) \\ &- \frac{1}{W} \sum_{p=1}^{N_i - 1} \left(A_{i,i_p} - \frac{s_i^{out} s_{i_p}^{in}}{W} + A_{i_p,i} - \frac{s_i^{in} s_{i_p}^{out}}{W} \right), \end{aligned} \tag{6}$$

where $j_p \in C_j$, $i_p \in C_i$ and $i_p \neq i$, N_j (N_i) is the number of nodes in community C_j (C_i) to which node j (i) belongs, s_i^{in} (s_i^{out}) is the in-degree (out-degree) of vertex i. The first term of the right hand side of above equation is the gain of modularity when node i moves to C_j, while the second term is the modularity gained when node i stays in its original community C_i. Next, the generated communities from the first step are used to form several meta-nodes. The weights between any two new nodes are given by the sum of the weights of edges between nodes in the corresponding communities [6].

$$A_{new}(i_{new}, j_{new}) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} A_{i,j}, \tag{7}$$

where $i_{new}, j_{new} = 1, \cdots, tN$, tN is the current number of meta-nodes, and $i = 1, \cdots, N$, N is the number of nodes. The two steps are iterated until the maximum modularity is achieved. This algorithm is shown in Algorithm 1.

B. Algorithm for community detection for multiple subjects

In neuroscience, one of the challenging problems is group analysis when analysis from multiple subjects need to be merged. There are two broadly used approaches for group analysis, i.e., 'virtual-typical-subject' (VTS) approach and 'individual structure' (IS) approach. The former approach averages the group data to obtain one community structure for the whole group. The latter one applies a clustering algorithm (e.g. Algorithm 1) to each individual subject and finds the common structures among them. These approaches do not consider either the inter-subject variability or the effect of outliers. For this reason, in this paper, we extend our community structure algorithm to multiple subjects by maximizing a common modularity function across subjects as shown in Algorithm 2. To be more specific, compared to Algorithm 1, instead of computing the gain of modularity when changing the community of node i for one subject, we consider the increase of the modularity of the whole group (line 8).

Algorithm 1 Community detection of weighted network

Input: Weighted adjacency matrix $A \in (0,1)^{N\times N}$, nodes $1,\cdots,N$, initial community structure $C = \{\{1\},\cdots,\{N\}\}$, $tN = N$.
Output: M Communities.

repeat
  $\Delta Q_{total} = 0$;
  Compute the modularity Q of the current network;
  for $h = 1$ to tN **do**
    for $j = 1$ to tN **do**
      The change of modularity ΔQ_{gen_j} when node h is assigned to C_j;
    end for
    $j* = \arg\max_j \Delta Q_{gen_j}$;
    if $\Delta Q_{gen_{j*}} > 0$ **then**
      $C_{j^*} = C_{j^*} \cup v_h$;
    end if
  end for
  Compute the change of the modularity of the current network $\Delta Q_{total} = Q_{new} - Q$;
  Nodes in the same community form new meta-nodes;
  tN is equal to the current number of communities;
  Recompute the weighted matrix $A \in (0,1)^{tN\times tN}$;
until $\Delta Q_{total} \leq 0$.

IV. Results

In this section, we test the effectiveness of the proposed community detection algorithm on both synthetic network and real EEG data.

A. Synthetic data

In this subsection, we test our algorithm on a directed network with 64 nodes and 4 clusters, with each cluster having 16 nodes. Each entry of the association matrix is uniformly distributed between $[0,1]$, which resembles the normalized DI value. The means of intra-cluster connectivity strength in the four clusters are 0.3, 0.5, 0.7, and 0.9, respectively. The mean of inter-cluster connectivity is 0.15. To demonstrate the robustness of the algorithm, the standard deviation of these distributions is modified from 0.1 to 0.5 with a step size of 0.1. The community detection results are evaluated by computing the percentage of false discoveries F,

$$F = \sum_{i,j}^{N} \frac{O_{i,j} - M_{i,j}}{N^2} \tag{8}$$

where N is the number of nodes, $O_{i,j}$ is a binary matrix with entries equal to 1 if nodes i and j are in the same cluster. If nodes i and j are identified in the same cluster by the algorithm, then $M_{i,j} = 1$; otherwise, $M_{i,j} = 0$. Without loss of generality, we generated the network 10 times and the average F is obtained for different standard deviations. The result is shown in Fig.1. We can observe that the false discovery rate grows with increasing standard deviation. Even so, the maximum false discovery rate of our algorithm is around 0.16, which is low and acceptable.

Algorithm 2 Community detection of multiple weighted networks

Input: Weighted adjacency matrix $A_i \in (0,1)^{N\times N}$, $i = 1,\cdots,L$, nodes $1,\cdots,N$, initial community structure $C = \{\{1\},\cdots,\{N\}\}$, $tN = N$.
Output: M Communities.

repeat
  $\Delta Q_{total} = 0$;
  Compute the modularity Q_i of the current network of subject i, $i = 1,\cdots,L$;
  The modularity of the group is $Q = \sum_{i=1}^{L} Q_i$;
  for $h = 1$ to tN **do**
    for $j = 1$ to tN **do**
      The change of modularity $\Delta Q^i_{gen_j}$ when node h is assigned to C_j for subject i, $i = 1,\cdots,L$;
      The change of the whole group is $\Delta Q_{gen_j} = \sum_{i=1}^{L} \Delta Q^i_{gen_j}$;
    end for
    $j* = \arg\max_j \Delta Q_{gen_j}$;
    if $\Delta Q_{gen_{j*}} > 0$ **then**
      $C_{j^*} = C_{j^*} \cup v_h$;
    end if
  end for
  Compute the modularity of the whole group $Q_{new} = \sum_{i=1}^{L} Q^i_{new}$, Q^i_{new} is the modularity of subject i;
  Compute the change of the modularity $\Delta Q_{total} = Q_{new} - Q$;
  Nodes in the same community form new meta-nodes;
  tN is equal to the current number of communities;
  Recompute the weighted matrix $A_i \in (0,1)^{tN\times tN}$ of subject i, $i = 1,\cdots,L$;
until $\Delta Q_{total} \leq 0$.

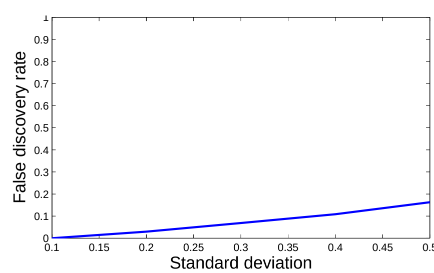


Fig. 1. False discovery rate of the community detection algorithm.

Moreover, to test the effectiveness of our group analysis method, we test Algorithm 2 on a group of six directed networks with the same community structure. Each network has 64 nodes and 4 clusters, with each cluster having 16 nodes. Each entry of the association matrix is uniformly distributed between $[0,1]$. The means of intra-cluster connectivity strength in the four clusters are 0.3, 0.5, 0.7, and 0.9, respectively. The mean of inter-cluster connectivity is 0.15. For each network, the standard deviation of all the edge values are randomly chosen from $[0.1, 0.2, 0.3, 0.4, 0.5]$, which leads to the variation across the six networks. Without loss of generality, we generate 10 simulations of networks to

TABLE I

AVERAGE FALSE DISCOVERY RATE

Approaches	Algorithm 2	VTS	IS
F	0.0125	0.0250	0.3143

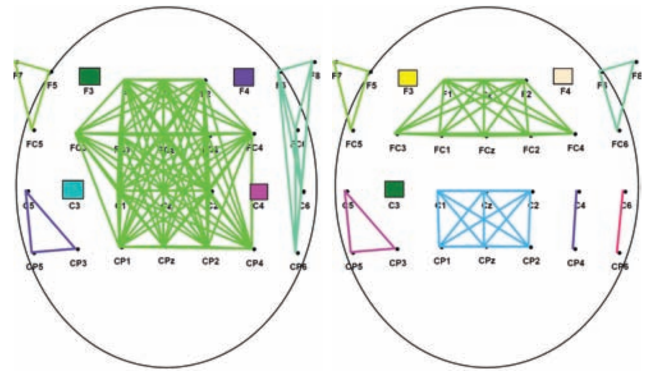
get the averaged false discovery rate. In addition, we compare our method with the two standard approach, i.e., VTS and IS. The averaged false discovery rate F for each method is shown in Table I. We can observe that the proposed algorithm has the lowest false discovery rate.

B. EEG Data

In this paper, we examined EEG data from a study containing the error-related negativity (ERN). The ERN is a brain potential response that occurs following performance errors in a speeded reaction time task. Previous work indicates that there is increased information flow associated with ERN for the theta frequency band ($4-8$ Hz) and ERN time window $25-75$ ms for Error responses (ERN) compared to Correct responses (CRN) [14]. We analyze data from 10 subjects for both the CRN and ERN. EEG data are preprocessed by the spherical spline current source density (CSD) waveforms to sharpen ERP scalp topographies and eliminate volume conduction [15]. In addition, bandpass filter is used to obtain signals in the theta band. The effective connectivity quantified by DI is computed over a window corresponding to the ERN response ($0-100$ ms after the response), for all trials between 30 electrode pairs in the theta band. Once the connectivity matrices for each response type of each subject is obtained, we use Algorithm 2 to identify the community structure of the group. Since it is a multi-level clustering algorithm, here we give the clustering result for the first level of clustering (highest resolution). The results are shown in Fig. 2. We can observe that the clusters of ERN are more localized. In addition, the frontal and central-parietal regions are not in the same cluster for ERN, which shows the functional specialization of the frontal and central-parietal regions whereas for CRN that specialization does not exist. The results are aligned with previous work in [16], which shows that error processing is controlled by the communication between the lateral prefrontal cortex and medial prefrontal cortex.

V. CONCLUSIONS

In this paper, we introduced a hierarchical community detection algorithm to identify the modules in the effective brain network. The association matrix of the network is obtained by applying the DI measure to EEG data involving a study of error-related negativity. In addition, we proposed a group analysis method to obtain a common community structure across subjects to address the problem of variability across subjects. The proposed method is applied to both synthetic data and EEG data and is shown to discriminate between erroneous and correct responses in terms of the community structures obtained.



(a) (b)

Fig. 2. Applying the algorithm to 10 subjects. (a) Corrected responses, 8 clusters. (b) Error related responses, 10 clusters.

REFERENCES

[1] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, pp. 1–37, 2005.

[2] Y. Liu and S. Aviyente, "Directed information measure for quantifying the information flow in the brain," in *International Conference of the IEEE, Engineering in Medicine and Biology Society*, 2009, pp. 2188–2191.

[3] D. A. Fair, A. L. Cohen, J. D. Power, N. U. Dosenbach, J. A. Church, F. M. Miezin, B. L. Schlaggar, and S. E. Petersen, "Functional brain networks develop from a "local to distributed" organization," *PLoS computational biology*, vol. 5, no. 5, pp. 1–14, 2009.

[4] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[5] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3-5, pp. 75–174, 2010.

[6] V. D. Blondel, J. L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, p. P10008, 2008.

[7] D. Meunier, R. Lambiotte, A. Fornito, K. D. Ersche, and E. T. Bullmore, "Hierarchical modularity in human brain functional networks," *Frontiers in neuroinformatics*, vol. 3, 2009.

[8] E. A. Leicht and M. E. J. Newman, "Community structure in directed networks," *Physical Review Letters*, vol. 100, no. 11, p. 118703, 2008.

[9] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, p. 7821, 2002.

[10] A. Arenas, J. Duch, A. Fernández, and S. Gómez, "Size reduction of complex networks preserving modularity," *New Journal of Physics*, vol. 9, p. 176, 2007.

[11] P. O. Amblard and O. J. J. Michel, "Measuring information flow in networks of stochastic processes," *Arxiv preprint arXiv:0911.2873*, 2009.

[12] A. Rao, A. O. Hero III, D. J. States, and J. D. Engel, "Using directed information to build biologically relevant influence networks," in *Proc. of Computational Systems Bioinformatics*, 2007, pp. 145–156.

[13] G. A. Darbellay and I. Vajda, "Estimation of the information by an adaptive partitioning of the observation space," *IEEE Transactions on Information Theory*, vol. 45, no. 4, pp. 1315–1321, 1999.

[14] S. Aviyente, E. M. Bernat, W. S. Evans, C. J. Patrick, and S. R. Sponheim, "A phase synchrony measure for quantifying dynamic functional integration in the brain," *Human Brain Mapping*, vol. 32, no. 1, pp. 80–93, 2010.

[15] J. Kayser and C. E. Tenke, "Principal components analysis of laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks," *Clinical Neurophysiology*, vol. 117, no. 2, pp. 348–368, 2006.

[16] J. F. Cavanagh, M. X. Cohen, and J. J. B. Allen, "Prelude to and resolution of an error: Eeg phase synchrony reveals cognitive control dynamics during action monitoring," *The Journal of Neuroscience*, vol. 29, no. 1, p. 98, 2009.