A Probabilistic Model of Functional Brain Connectivity Network for Discovering Novel Biomarkers

Jiang Bian, PhD[1], Mengjun Xie, PhD[2], Umit Topaloglu, PhD[1], Josh M. Cisler, PhD[1]
[1]University of Arkansas for Medical Sciences, Little Rock, AR; [2]University of Arkansas at Little Rock, Little Rock, AR

Abstract

Graph theoretical analyses of functional brain connectivity networks have been limited to a static view of brain activities over the entire timeseries. In this paper, we propose a new probabilistic model of the functional brain connectivity network, the strong-edge model, which incorporates the temporal fluctuation of neurodynamics. We also introduce a systematic approach to identifying biomarkers based on network characteristics that quantitatively describe the organization of the brain network. The evaluation results of the proposed strong-edge network model is quite promising. The biomarkers derived from the strong-edge model have achieved much higher prediction accuracy of 89% (ROC-AUC: 0.96) in distinguishing depression subjects from healthy controls in comparison with the conventional network model (accuracy: 76%, ROC-AUC: 0.87). These novel biomarkers have the high potential of being applied clinically in diagnosing neurological and psychiatric brain diseases with noninvasive neuroimaging technologies.

Introduction

Graph-theory based analyses of both structural and functional human brain networks using modern brain mapping techniques such as functional magnetic resonance imaging (fMRI) have been applied to understanding the topology of human brain organization. Previous research has revealed that the organization of human connectome (a connectome is a comprehensive map of neural connections in the brain) resembles features of other complex networks (e.g., social networks and the Internet) such as small-world topology, scale-free degree distribution, highly connected hubs with high degree or high centrality, and modularity[1,2]. However, conventional brain network analyses methods[1] assume temporal stability of brain activities over the entire fMRI timeseries when constructing functional brain connectivity networks. In[3], we introduced the strong-edge method that can generate a functional network with a higher temporal resolution, where entire timeseries are divided into equal length intervals and the stability of brain activities is only assumed within a small time interval. Extending our network generation method by considering the network dynamics influenced by the temporal fluctuations of the brain neurodynamics, we further propose a probabilistic model of functional connectivity network, namely the strong-edge model, which we believe can express the underlying brain states more accurately and provide a more informative representation of the functional brain connectivity network.

Based on neural connectivity patterns, researchers have been able to discriminate individuals with Schizophrenia[4], depression[5,6], and Alzheimer's disease[7] from healthy control individuals. However, differentiation approaches in previous research are disease specific and may not be generalized and applied to other neurological diseases. In this paper, we propose a novel approach to systematically discovering biomarkers. Our evaluation using a real depression dataset has shown that the biomarkers can be effectively obtained from the network characteristics of brain connectivity networks. More importantly, we find that the proposed probabilistic model can improve the classification accuracy dramatically compared to the conventional method. We expect that our probabilistic brain connectome model will help to not only shed light on the basic understanding of human brain, but also transform the theoretical analyses of human brain into discoveries of biomarkers for neurological and psychiatric brain diseases and lead to computer-aided diagnostic tools.

Materials

19 women with no history of Major Depression Disorder (MDD) and 19 women with a history of MDD (n=6 with current MDD) underwent a 7.2 minutes resting-state fMRI scan. MDD was characterized with the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID[8]). The neuroimaging dataset contains both anatomical and functional images that were acquired using a 3.0 T Siemens Magnetom Trio modality. For each subject in each session, 210 planar images depicting BOLD responses were acquired with a $2.02s$ TR, total duration of 7.2 minutes. During the *resting-state* scan, participants were instructed to lie passively in the scanner and to refrain from thinking about anything

specific. Image preprocessing followed standard steps and was completed using the AFNI software. Moreover, based on prior research demonstrating altered organization of the Default Mode Network (DMN) among depressed samples[9], we chose to discriminate the groups based on organization of the DMN. We selected 11 regions commonly identified with the DMN (i.e., PCC, rPG, lPG, SCC, rPC, lPC, rITG, lITG, rSFG, lSFG, MPC) that were derived from an independent component analysis (ICA) and verified by a neuroscience expert. The timecourses from each voxel comprising the region of interest (ROI) were extracted and singular value decomposition (SVD) was used to identify the time course component explaining the largest portion of variance in the voxels' fluctuations. This resulted in an 11×210 ($number\ of\ ROIs \times number\ of\ timepoints$) matrix for each individual.

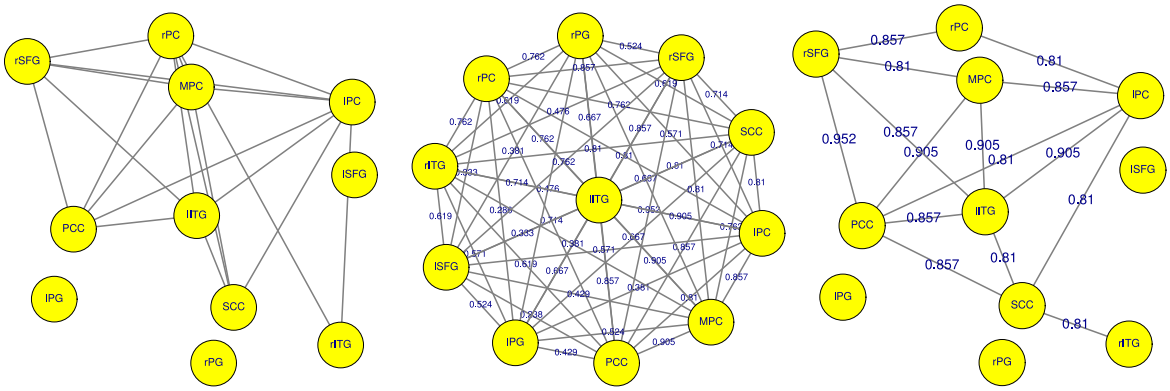
Methods

a) Defining neural processing networks with the strong-edge method:
In graph theoretical analyses of functional connectome, a brain network is modeled as an undirected graph, $G = (V, E)$ where a node/vertex (V) in the graph represents a brain region (i.e., ROI) and an edge (E) between two nodes indicates that the two brain regions are functionally connected. By calculating the correlation coefficient (e.g., Pearson product-moment correlation coefficient (PMCC)) for every pair of ROIs based on the brain activity timeseries, we generate a correlation coefficient matrix, from which we can derive a binary adjacency matrix. Two ROIs are regarded as adjacent and connected via an edge in the connectivity graph if they are highly correlated, that is, the correlation coefficient value is greater than the given threshold value T with a reasonable P-value. The threshold T is determined by a targeted density of the graph. The graph density is defined as the number of edges divided by the total number of possible edges. For an undirected graph $G = (V, E)$ with $|V|$ vertices and $|E|$ edges, $density = \frac{2|E|}{|V|(|V|-1)}$. For example, if the targeted density is 0.37, the threshold T is the value that eliminates 63% of the edges in the graph. The graph $density$ is usually set in the range of $[0.37, 0.50]$ in previous research[10]. A graph tends to be fragmented if $density$ is very small, and it may not represent a biological system if $density$ is very large. Clearly, a lower T incurs more correlations and thus more edges, resulting a higher $density$ of the graph. Once the connectivity graph is obtained, network characteristics can then be computed to describe the organization of the brain network.

A limitation of this traditional approach is that it assumes temporal stability of connectivity between two regions; that is, the correlation between two regions in the traditional approach is computed over the entire measured timeseries (often greater than 6 minutes). However, emerging research findings suggest that the connectivity strength between regions may be dynamic and temporally fluctuating[11,12]. Therefore, we propose to consider the frequency of functional connections between brain regions over time and regard the frequent connections as "strong" and important to the operation of the overall brain network[3]. By doing so, our method provides *a fine-grained probabilistic view of the functional brain network* that is *less susceptible to spurious and time-limited temporal covariance*. Furthermore, by emphasizing strong connections, we can effectively prune the network and reveal more significant functional neural connectivities.

The basic idea of constructing a strong-edge graph is to divide the entire timecourses of the scan into smaller time windows, and construct a snapshot graph for each time window. In other words, we take snapshots of the functional brain network of a human subject with a given frequency along the timecourses. We define the strength of an edge $e_{i,j}$ ($i \neq j$), denoted by $S_{e_{i,j}}$, as the frequency of its appearance in all the snapshot graphs generated from the entire timecourses. The total number of time points (TP) in a timecourse is denoted as $|TP|$, and the number of time points in each time window (TW) is denoted as n. For data smoothing purpose, two consecutive time windows have c number of TPs in common; that is, the last c time points in TW_{m-1} are the first c time points in TW_m. In the algorithm, we denote the data smoothing parameter as $step = |TP| - c$. For each TW_m, we construct a connectivity graph (G_m) using the conventional method (only consider the correlation coefficient with small P-value, e.g., 0.05) with a target $density$ value. The strength of an edge $e_{i,j}$ is defined as $S_{e_{i,j}} = |e_{i,j}|/|TW|$, where $|e_{i,j}|$ is the number of graphs that contain edge $e_{i,j}$ and $|TW|$ is the total number of graphs (i.e., the number of time windows). We then assign the statistical probability (i.e., strength) of each connection as the weight of the corresponding edge in the graph. In this manner, we turn an undirected binary graph in the conventional graph approach into an undirected weighted graph (i.e., a frequency graph). Further, we define "strong edges" as those edges whose strengths are no smaller than the minimal support S_{minsup} ($S_{minsup} \in [0, 1]$). Figure 1 shows the three brain connectivity graphs of the same depression subject, where graph (a) is constructed using the conventional graph approach; (b) is the frequency graph;

and (c) is the strong-edge graph.



(a) (b) (c)

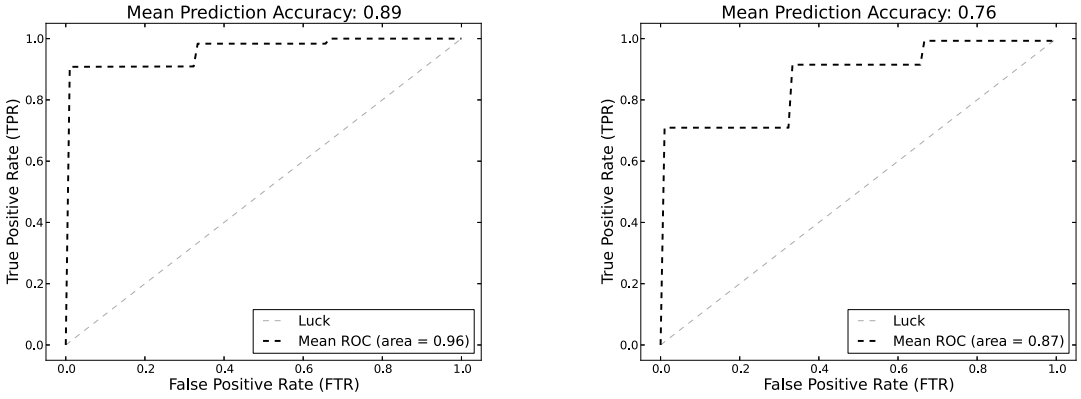
Figure 1: three brain connectivity graphs for the same depression subject, where graph (a) is constructed using the conventional graph approach ($density = 0.37$); (b) is the frequency graph ($density = 0.65$, $n = 20$, and $step = 10$); and (c) is the strong-edge graph where $S_{minsup} = 0.8$.

b) Mining biomarkers from graph-theory based network characteristics:
The research observation that neural network organization patterns differentiate disease states[4,5,6,7] demonstrates that neural network features are viable biomarkers of the disease. Moreover, identifying which neural organization patterns allow for accurate patient discrimination reveals the essential brain states that mediate the disease state. Elucidation of these key neural information processing patterns provides powerful targets with which to track symptom severity over time (e.g., to predict relapse or treatment response). Naturally, we think the differences in organization of brain networks shall be reflected by differences in the quantitative network characteristics. Therefore, we shall be able to build classification algorithms using the brain connectivity network characteristics as features to differentiate diseased subjects from healthy controls.

For each subject's brain connectivity network, network characteristics including each node's degree (or strength in a weighted graph), betweenness centrality[13], local and global clustering coefficient[14], closeness centrality[13], assortativity coefficient[15], and global characteristic path length[14] of the graph are calculated. Both graphs constructed by the conventional and the strong-edge approaches are undirected; however, the edges in a strong-edge graph are weighted using the occurrence probability of the edges. Therefore, network characteristics of strong-edge graphs are computed using their weighted definitions. These network characteristics are then used as features (i.e., 94 features extracted from the conventional model and 105 features from the strong-edge model) to build Support Vector Machine (SVM) classifiers. Various standard techniques such as scaling, grid-based kernel parameter search, feature selection using F-test, and etc.[16] for building an efficient SVM have been taken into account in this study.

Results

To reduce bias, we bootstrap the SVM classifier 1000 times. During each iteration, we randomly split the samples into two datasets, 2/3 (26 subjects) used for training and the other 1/3 (12 subjects) left as test samples. For simplicity, we maintain the two datasets balanced (e.g., the testing dataset contains 6 positive subjects and 6 negative subjects). We then train an SVM model using only the training set (including the F-test feature selection process), and measure the classifier's performance on the independent test samples. The final prediction accuracy and the ROC-AUC (a higher ROC-AUC value indicates the classifier has better sensitivity and specificity) value are averaged over all iterations. As shown in Figure 2, using the strong-edge model, the binary SVM classification of subjects with MDD yields a prediction accuracy of 89% and a ROC-AUC value of 0.96, while the classification model based on the conventional graph-theory approach yields a lower prediction accuracy of 76% and a ROC-AUC value of 0.87. Identifying and analyzing the features (i.e., network characteristics) that drive the multivariate predictor provides mechanistic information as to aspects of brain functional (dys)organization that characterize the depression state. The features derived from the strong-edge model that most strongly aided in group discrimination are the betweenness centrality of rPG, and the clustering coefficient of lSFG; while the best features learned from the conventional model are the betweenness



(a) (b)

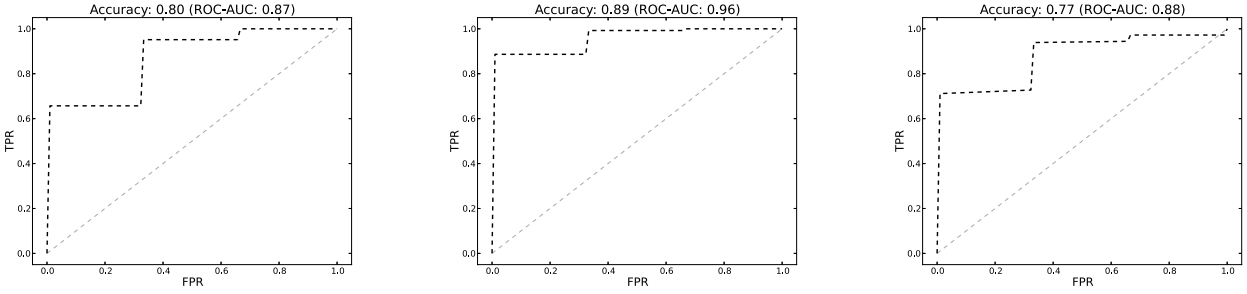
Figure 2: The mean accuracies and the ROC curves for classifiers using (a) the probabilistic strong-edge model ($density = 0.76$, $n = 10$, $step = 30$, and $S_{minsup} = 0.8$) V.S. (b) the conventional model.

centrality of rPG, lITG, and lSFG, the clustering coefficient of rPG, lITG, and rSFG, and the closeness centrality of rPC, lPC, rlTG, and llTG, indicating specific neural organizational patterns that differentiate MDD.

Discussion

There are a number of free variables that can influence the accuracy of brain network models and consequently affect the classifiers' performance. They are discussed as follows.

In conventional models, the choice of the correlation coefficient threshold T, which is used to determine connectivity (or edge), directly affects the structure of the graph. Normally, a target $density$ of the graph is first given to determine the threshold T. However, the density of each individual's brain connectivity network can vary, and is often unknown. Following previous studies [1 10], we vary the target density value ($0.37 \leq density \leq 0.50$) to construct a set of brain connectivity networks for each individual, measure the network characteristics of each graph, and then average the values across the density range.



(a) $density = 0.71$ (b) $density = 0.76$ (c) $density = 0.79$

Figure 3: The choice of the $density$ variable in the strong-edge model has a strong effect on the SVM classifiers' performance. Through experiments, $density = 0.76$ gives the best result.

In the strong-edge model, a statistical probability of occurrence is assigned to each edge as its weight. The weight can be interpreted as the importance of the neural pathway (edge) to the network organization efficiency. When constructing strong-edge graphs, we choose a high density value (accordingly, a low T value) to preserve all the necessary edges to prevent information loss. Because, if an edge exists in the connectivity graph using a low target density value, the same edge will be preserved with a higher density value. We find that the SVM classifier can achieve the highest prediction accuracy with density value 0.76. However, higher $density$ values reduce the classifier's performance, possibly because more noise is introduced while preserving more information, as shown in Figure 3. We also examine other variables (e.g., the size of time window n, the $step$ value for time window smoothing, and

S_{minsup} that defines the criteria for a strong edge) besides $density$ to understand their effects on constructing strong-edge graphs and subsequently the organizational properties of the brain network. Through extensive experiments we conclude that other variables have less expressive influence on the topologies of constructed networks and the classifiers' performance. Due to space limit, these results and potential neuroscience interpretations are not included in this paper.

Conclusion and future work

In this paper, we proposed a probabilistic brain network model—the strong-edge model—to provide a more realistic and informative representation of the brain organization with high temporal resolution. Moreover, we introduced a systematic approach to discovering novel biomarkers for neurological and psychiatric brain diseases based on network characteristics. Our evaluation results support the contention that the strong-edge approach to network analysis yields more reliable candidate biomarkers of brain diseases. A limitation of the current approach is that the selection of ROI network relies on prior knowledge in neuroscience and specific brain diseases. A promising alternative to ROI-based approaches is to conduct voxel-level network analyses in a data-driven (i.e., no prior hypotheses or knowledge) fashion.

Acknowledgements

The work describe in this manuscript is supported by award UL1RR029884 through the NIH National Center for Research Resources and National Center for Advancing Translational Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The work is also supported by the National Science Foundation EPSCoR Cyber infrastructure award #EPS-0918970.

References

1. Bullmore E. and Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci., 10: 186–198, Mar 2009.
2. van den Heuvel M. P., Stam C. J., Boersma M. and Hulshoff Pol H. E. Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain. Neuroimage, 43:528–539, Nov 2008.
3. J. Bian, M. Cisler J., M. Xie, A. James G., R. Seker and D. Kilts C. A methodology for empirical analysis of brain connectivity through graph mining. In Proceedings of the 2011 IEEE International Conference on Systems, Man, and Cybernetics, SMC '11, Anchorage, Alaska, USA, 2011. IEEE Systems, Man, and Cybernetics Society.
4. Calhoun V. D., Maciejewski P. K., Pearlson G. D. and Kiehl K. A. Temporal lobe and "default" hemodynamic brain modes discriminate between schizophrenia and bipolar disorder. Hum Brain Mapp, 29(11):1265–1275, Nov 2008.
5. Craddock R. C., Holtzheimer P. E., Hu X. P. and Mayberg H. S. Disease state prediction from resting state functional connectivity. Magn Reson Med, 62(6):1619–1628, Dec 2009.
6. Fu C. H., Mourao-Miranda J., Costafreda S. G., Khanna A., Marquand A. F., Williams S. C. et al. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. Biol. Psychiatry, 63(7):656–662, Apr 2008.
7. Chen G., Ward B. D., Xie C., Li W., Wu Z., Jones J. L. et al. Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging. Radiology, 259(1):213–221, Apr 2011.
8. First M.B., Spitzer R.L. and Williams J.B. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Non-Patient Edition. Structured Clinical Interview for DSM-IV Axis I Disorders: SCID-I. New York: Biometrics Research, New York State Psychiatric Institute, 2002.
9. Greicius M. D., Flores B. H., Menon V., Glover G. H., Solvason H. B., Kenna H. et al. Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol. Psychiatry, 62(5):429–437, Sep 2007.
10. Lynall M. E., Bassett D. S., Kerwin R., McKenna P. J., Kitzbichler M., Muller U. et al. Functional connectivity and brain networks in schizophrenia. J. Neurosci., 30:9477–9487, Jul 2010.
11. Whitlow C. T., Casanova R. and Maldjian J. A. Effect of Resting-State Functional MR Imaging Duration on Stability of Graph Theory Metrics of Brain Network Connectivity. Radiology, Mar 2011.
12. Stamoulis C., Gruber L. J. and Chang B. S. Network dynamics of the epileptic brain at rest. Conf Proc IEEE Eng Med Biol Soc, 2010:150–153, 2010.
13. Freeman L. Centrality in social networks: Conceptual clarification. Social Networks, 1(3):215–239, 1979. ISSN 03788733. doi: 10.1016/0378-8733(78)90021-7. URL http://dx.doi.org/10.1016/0378-8733(78)90021-7.
14. Watts D. J. and Strogatz S. H. Collective dynamics of 'small-world' networks. Nature, 393(6684):440–442, Jun 1998.
15. Newman M. E. Assortative mixing in networks. Phys. Rev. Lett., 89:208701, Nov 2002.
16. Hsu C. W., Chang C. C. and Lin C. J. A practical guide to support vector classification. Department of Computer Science and Information Engineering, National Taiwan University, Taipei, Taiwan, 2003. URL http://www.csie.ntu.edu.tw/\~{}cjlin/papers/guide/guide.pdf.