PROBABILISTIC BOOLEAN NETWORK FOR INFERRING BRAIN CONNECTIVITY USING FMRI DATA

Zheng Ma, Z. Jane Wang *

Department of Electrical and Computer Engineering
University of British Columbia
Canada

Martin J. McKeown

Pacific Parkinson's Research Centre
Brain Research Centre
University of British Columbia
Canada

ABSTRACT

Recent research has suggested disrupted interactions between brain regions may contribute to some of the symptoms of Parkinson Disease (PD). It is therefore important to develop models for inferring brain functional connectivity from non-invasive imaging data, such as functional magnetic resonance imaging (fMRI). In this paper, we propose applying Probabilistic Boolean Network (PBN) for modeling brain connectivity due to its solid stochastic properties, computational simplicity, robustness to uncertainty, and capability to deal with small-size data, typical for fMRI data sets. Applying the proposed PBN framework to real fMRI data recorded from PD subjects, we noticed that the PBN method detected statistically significant brain connectivity between region-of-interest (ROIs) in PD and normal subjects. In addition, the PBN results suggest a mechanism of the effectiveness of L-dopa, the principal treatment for PD.

Index Terms— Probabilistic Boolean Network, fMRI, Brain Connectivity, Group Analysis

1. INTRODUCTION

The introduction of non-invasive medical technologies such as fMRI and quantitative EEG allows researchers to gain more insightful understanding about human brain functioning in disease states. Understanding brain connectivity, the neural influence that one brain region exerts over another, is increasingly recognized as important for brain function, and its impairment may be associated with neurodegenerative diseases such as Parkinson Disease (PD).

Learning effective connections between brain regions of interest (ROIs) requires a mathematical model suitable for complex, large-scale, and dynamical system computation, and there have been several models proposed in the literature, such as structural equation modeling (SEM) [1], multivariate autoregressive models (MAR) [2], dynamic causal modeling (DCM)[3], and dynamic Bayesian network (DBN)[4]. Both SEM and MAR are conventional linear models which take the blood-oxygenation-level-dependent (BOLD) signals as direct inputs. As the functioning of brain has been widely believed to act in a nonlinear way, DCM, a nonlinear input-state-output model, has been actively applied for brain connectivity learning as a supplement to linear modeling methods. Recently, (dynamic) Bayesian networks (BN) have been applied to discover brain connectivity in fMRI. BN approach does not require a prior structure based on anatomical connections, making it suitable for pathological conditions like PD, where functional connectivity may not be known before hand.

*Thanks to NSERC agency for funding.

DBN's recent success for learning brain connectivity [4] has inspired us to explore the Probabilistic Boolean Network (PBN). PBN has proved successful in several areas, such as in gene expression analysis. PBN describes the dynamic of the studied system in the probabilistic context of markov chain [5]. The interaction among different nodes of the system is represented based on logical rules with stochastic properties. PBN has many features that makes it very attractive in fMRI data analysis: it has computational simplicity, robustness to uncertainty, and is capable of dealing with small-size data.

The paper is organized as follows. We describe the PBN framework for brain effective connectivity and group analysis in Section 2. A case study using fMRI data from normal and PD subjects performing a motor task at three progressive levels of difficulty is discussed in Section 3.

2. METHODS

In this section, we present a PBN-based framework for inferring connectivity between all interested brain regions. This framework not only visualizes, but also quantifies the connections between brain regions. Those quantitative inter-region connections will offer insightful understanding of the underlying mechanism of brain activities. The PBN framework primarily consists of three components, including the preprocessing of fMRI data, the inference of PBN, as well as group analysis in fMRI using PBNs.

2.1. Preprocessing of fMRI data

The raw fMRI data first goes through 3D motion correction and slice timing correction upon being obtained, and the data are then further motion corrected with Motion Corrected Independent Component Analysis (MCIMCA), a computationally expensive, but highly accurate method. We then manually draw all 18 ROIs on each structural scan that has been aligned with the functional data using Amira software. The resulted fMRI data is a matrix containing time courses of voxels belonging to the 18 ROIs. Since PBN is a discrete-valued probabilistic model, its input data is confined to be binary vectors, unlike the popular Bayesian Network normally dealing with continuous-valued data. In addition, we focus on an integrated analysis on the ROI-level, instead of studying the individual behavior of each voxel. Therefore, the success of our PBN analysis depends on the well-being of three important preprocessing steps: denoising, voxel selection, and binarization.

fMRI Denoising: Most current denoising methods focus on building a reasonable model for serially correlated fMRI data. A tradi-

tional model for fMRI time series y of each voxel is linear regression model [6]. Some new approaches for fMRI denoising have been proposed as efficient ways to extract the underlying components that are task-related. Among them, exploratory analysis such as Independent Component Analysis (ICA) and Canonical Correlation Analysis (CCA) have received the most attention. CCA focuses on finding independent components $s_i(t)$ with maximum autocorrelation, and under certain circumstances may compare favorably with ICA and PCA in extracting interested signals from fMRI [7] since it is more computationally efficient, and more robust to low sample size. However, both linear regression modeling and exploratory analysis of fMRI data have their own shortcomings that may affect the statistical results. Though we have included CCA as part of a complete denoising process in our study, for simplicity, here we simply remove the low-frequency drift from raw data through linear detrending as denoising. We leave other tasks such as activation detection to the following processing steps in PBN modeling.

Voxel Selection: Many previous fMRI modeling methods pay special attention to those activated regions indicated by high significant values (t-value) obtained through software such as SPM. In our study, we take into account all 18 ROIs since we aim at obtaining a complete PBN. Given the fact that our study focuses on region-specific interactions, choosing appropriate voxels in each region to represent the overall behavior of that region will significantly contribute to the success of our PBN analysis. There are various ways for voxel selection including peak selection, eigenvariate selection, and average selection. There exists no universal standard on how to choose a subset for optimal performance. Our approach for voxel selection combines several previous methods: we first use SPM to obtain t-values of all voxels in the brain, and then find the two-tailed critical value c at a 95% level. Those voxels with t-values above the critical value are divided into two groups, positively activated group with t-values higher than the critical value and the negatively activated group with t-values lower than the threshold. For voxels in the positively activated group, we keep their time courses unchanged, while we flip the time courses of voxels in the negatively activated group. To distinguish activated voxels from un-activated voxels, we further reconstruct the data by zeroing the time courses of all un-activated voxels. For each ROI, we do not simply take the mean of all activated voxels inside to represent the activities for that region, instead, we build a cubic with unit length centering at each activated voxel, and replace its original time course with the mean of time courses of voxels inside that cubic.

Binarization: One big challenge we face for PBN inference is to develop a feasible binarization method which could quantize real-valued fMRI data into binary values (on-state and off-state). General data-driven binarization methods usually aim at finding a threshold, such as the "big jump" method in [5], and the approach described in [9]. The basic idea of "big jump" method is to find the largest difference between two consecutive values in the sorted data as the threshold. Since different ROIs reveal different levels of measurements, our binarization method is conducted on a ROI-wise basis. However, our preliminary results from exploring the "big jump" method do not seem optimistic to us given the fact that the largest difference often happens in the beginning of our sorted data, which makes the resulted binary data almost have no off states. [9] proposed a new binarization method based on extended linear model. It involves pre-whitening unknown noise by estimating the serial correlations of noise components, and then comparing the noise-removed data with a prespecified threshold, which is defined as a constant mutiplied by the variance of the whitened noise. The results for each voxel are stored in a binary vector to indicate whether elevated activity due to the paradigm occurs at each measurement time. As for this project, we adopt the approach proposed in [8], which focuses on capturing the first-order statistical trend in the fMRI data because of its robostness against parameter change. Since the measurements of fMRI, known as the blood oxygen level-depend (BOLD) response, directly reflects the tested behavior, usually described by the task paradigm, it is reasonable to model the BOLD response as a hidden markov model (HMM) process with parameters determined by the paradigm and itself. The key elements necessary to form a complete HMM include a set of observations O(t), which, in our case, are the time courses of the ROIs, a two-state (1: off, 2: on) transition matrix A, and observation probability distribution B_i for each state ($i = 1, 2$), which is simplified as a Gaussian Distribution, and an initial state π. Since we assume that the task starts from the rest condition, the initial state π could be set as $\pi_1 = 1$, and $\pi_2 = 0$. The mean μ_i and variance σ_i of observation probability distribution for state i are determined by the structure of the paradigm p and the observations O_t. Take the off-state for example,

$$\mu_1 = \frac{1}{|p_{off}|} \sum_{t \in p_{off}} O_t, \sigma_1 = \sqrt{\frac{1}{|p_{off}|} \sum_{t \in p_{off}} (O_t - \mu_1)^2}, \quad (1)$$

where p_{off} is the off-state fraction of the task paradigm. For the transition matrix A, we try a couple of values within an acceptable range, and especially pay attention to those results shared by different choices of A. We then employ Viterbi Algorithm to obtain the most likely state sequence, which is used to represent the fMRI voxel time course in the binary domain.

2.2. Learning PBN

Since PBN is an extension of boolean network (BN), we first give a brief introduction of BN. A BN consists of two parts: a set of nodes $(x_1, ..., x_n)$, and a list of binary functions $(f_1, ..., f_n)$. Each f_i contains a number of input nodes $(x_1^i, ..., x_k^i)$, where k varies according to different nodes. Values of input nodes at the time t keep updating the value of target node x_i at the next time $(t + 1)$:

$$f_i(x_1^i(t), ..., x_k^i(t)) = x_i(t + 1). \quad (2)$$

As we can see, the dynamic of a BN is deterministic, and the evolvement of system variable Z(t) $(x_1(t), ..., x_n(t))$ completely depends on the initial condition $(x_1(0), ..., x_n(0))$. To avoid the potential risk caused by the deterministic rigidity of BN, Shmulevich proposed to add probabilistic setting to BN in one of his papers [5]. The basic of PBN is to accommodate more than one functions for each node. To be more specific, for each node x_i, there is a set of binary functions f_j^i where each one is a possible function to determine the value of at the next time point with certain probability p_j^i. For simplicity, we denote x_i as the target, and f_j^i as predictors,

$$F = \left\{ f_j^i \right\} j = 1, ..., n, Pr(x_i(t + 1) = f_j^i(t)) = p_j^i, \quad (3)$$

where n is the number of possible functions of x_i.

There are 18 ROIs under our analysis with each one being denoted as a node x_i (i = 1, 2, ...18). In order to find predictors for each node x_i, we use PBN toolbox implemented by Lhdesmki, and Shmulevich, which aims at finding best-fit predictors. We assume that the maximum number of k to be five, and then search boolean functions with all possible input variable combination for those with minimal error size (the number of mismatches between predicted sequence and data sequence). Since each optimal predictor f_j^i for

target x_i has the same error size, therefore producing the same Coefficient of Determination θ_j^i. The corresponding probability p_j^i is calculated as follows:

$$p_j^i = \frac{\theta_j^i}{\sum_{j=1}^{n} \theta_j^i}. \quad (4)$$

It is noted that predictors always share the same probability in our case. Upon obtaining predictors, we need to identify the impact of one node on another, which is measured in terms of influence value. A large influence value means that the corresponding node play a very important role in determining the target value. Consider node x_i with a function set F_i including all predictors f_j^i and corresponding probabilities p_j^i (j = 1...n).The influence of x_k upon x_i - $I_k(x_i)$ is defined as:

$$I_k(x_i) = \sum_{j=1}^{n} I_k(f_j^i) * p_j^i, \quad (5)$$

where $I_k(f_j^i)$ is the influence of x_k on the predictor f_j^i:

$$I_k(f_j^i) = Pr\left\{\frac{\partial f_j^i}{\partial x_k} = 1\right\}. \quad (6)$$

This influence calculation is applied all 18 ROIs, thus forming a 18x18 matrix G with each entry $G_{i,j}$ being the influence of x_i on x_j. The influence matrix G enables us to obtain a complete connectivity network. This network can be demonstrated as a weighted directed graph, with each arrow representing the influence of one ROI to another. The value for each arrow is the magnitude of the influence.

2.3. Group Analysis for Inter-Subject Variability

So far we have learned PBN for each individual subject. However, to meaningfully extrapolate PBN results from one subject to an entire population (e.g. PD patient group) first requires methods to meaningfully integrate results from individual subjects and rigorously compare PBNs across group. To address this inter-subject variability issue, a common and critical challenging problem in many biomedical studies, we do group analysis for each influence connection contained in G individually through the Analysis of Variance (ANOVA).

ANOVA could test the effects of multiple factors on the data. Identifying the PD-induced connection impairment is doubtlessly our first priority. Nevertheless, given the fact that our experiments involve three-frequency tasks, we also need to consider the effects of different frequencies. As a result, we incorporate two factors in the ANOVA analysis: frequency and group (e.g. normal, PD on medicine, and PD off medication), and consider these two factors independently. The ANOVA will return a p-value for each factor, and a small p-value below certain threshold (0.05 in our case) indicates the existence of significant difference caused among data by the corresponding factor.

However, the ANOVA test only offers us general information that the means of data are significantly different so that they could reject the hypothesis that they are all the same. We need to know exactly which pairs of means are significantly different, and which are not. The multiple comparison method provides us with graphs indicating the estimated mean of each group from the current factor with a predefined confidence interval. Two means are significantly different if their intervals are disjoint, and are not significantly different if their intervals overlap. We should be able to tell which pair causes the significant difference by looking at those graphs. To further investigate the effectiveness of current medicine on the PD, we use multiple comparison method to estimate the means of each connection in different groups (M_{NP}, M_{PA}, M_{PB}), and evaluate the performance of medicine by comparing the difference between NP and PA with the difference between PB and PA. If ($M_{NP} - M_{PA}$) shares the same sign as ($M_{NB} - M_{PA}$), then we can say that the medicine does alleviate the PD symptoms.

3. FMRI CASE STUDY

3.1. fMRI Data Collection

The study was approved by the University of British Columbia ethics board and all subjects were gave written informed consent prior to participating. By using a pressure-responsive bulb which was electronically connected to a computer, subjects incluing ten normal and ten PD patients were asked to squeeze the bulb to control the height of a vertical bar to match a target bar moving up and down in a sinusoidal fashion. They lay down and squeezed a rubber bulb at four frequencies (0.00Hz, 0.25Hz, 0.5Hz and 0.75Hz) for 30-s blocks as instructed in a pseudo-random order. The patients took the experiment twice, once before medication and the other after medication. fMRI data of their brain activities during performing the task was collected with a Philips Achieva 3.0 T scanner. The following eighteen brain regions were selected as the ROIs in the study, both the left and right: primary motor cortex (M1), supplementary motor cortex (SMA), lateral cerebellar hemisphere (CER), putamen (PUT), caudate (CAU), thalamus (THA), prefrontal cortex (PFC), anterior cingulate cortex (ACC), and globus pallidus (GLP).

3.2. Results and Discussions

First, to illustrate the learned PBN for individual subjects, an example of PBN for a normal subject performing High frequency squeezing using left hand was shown in Fig. 1. From this figure, we are able to visually identify important core nodes (ROIs) which play important roles on the corresponding fMRI task performing. For the purpose of backbone ROI identification, we employ a simple way here by detecting the ROIs which represent the largest topology change if the node is removed from the network. It is noticed that ROI R_PUT, R_THA, R_CER, L_M1, and R_PFC are identified as core regions which is consistent with medical knowledge. The R_CER, L_M1 regions represent a ipsilateral motor network that appears recruited at higher frequencies. The subcortical regions of R_PUT, R_THA have been in many prior studies to be important for the scaling of movements.

We further examine the group analysis results. Based on the ANOVA test of individual connections, we first identify all influence connections whose values are significantly different due to either group or frequency factor. Here due to the space limitation, we only report the detail results due to the group factor. Table I lists all the connections which are identified significantly different due to the group factor, in which there are three groups including NP(normal people), PA(PD patients off medicine), and PB(PD patients on medicine). Some connections in Table I match up with the results obtained through DBN, such as the connection from R_M1 to R_SMA. The SMA is considered "more upstream" in the motor control and projects to M1.

Another purpose of group analysis here was to determine the overall group effect of L-dopa medication on PD patients and to identify the specific connections between ROIs that were affected

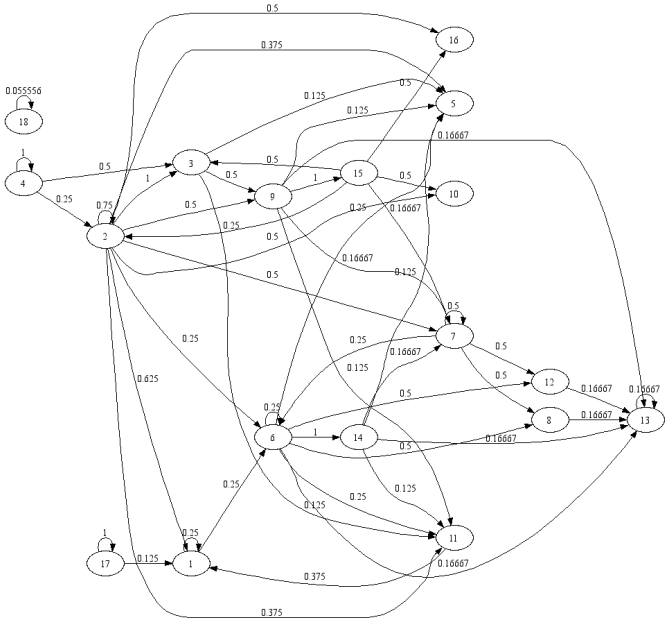

Fig. 1. An example of the learned PBN from the fMRI data for a normal subject performing a High frequency squeezing using the left hand.

by medication. Table II indicates that L-dopa does have effects on some of the connections ($R\_PUT \rightarrow L\_PUT$, $L\_THA \rightarrow L\_THA$, $R\_THA \rightarrow L\_ACC$, $L\_CER \rightarrow R\_M1$, $R\_M1 \rightarrow R\_SMA$, $R\_GLP \rightarrow R\_M1$). However, among those connections, there are several cases where the medicine seems to "overshot" PD ($R\_THA \rightarrow L\_ACC$, $L\_CER \rightarrow R\_M1$, $R\_GLP \rightarrow R\_M1$). In contrast, L-dopa seems to worsen the situation under some cases such as $R\_THA \rightarrow R\_M1$, $L\_PFC \rightarrow R\_ACC$, $L\_ACC \rightarrow L\_PUT$, $L\_GLP \rightarrow R\_CER$, which requires further clinical research.

Table 1. Significantly different connections identified across groups when performing the left hand squeezing tasks.

Connection	P-value
$R\_PUT \rightarrow L\_PUT$	0.0258
$L\_THA \rightarrow L\_THA$	0.0181
$R\_THA \rightarrow R\_M1$	0.0106
$R\_THA \rightarrow L\_ACC$	0.0169
$L\_CER \rightarrow R\_M1$	0.0438
$R\_M1 \rightarrow R\_SMA$	0.0365
$L\_PFC \rightarrow R\_ACC$	0.0095
$L\_ACC \rightarrow L\_PUT$	0.0427
$L\_GLP \rightarrow R\_CER$	0.0197
$R\_GLP \rightarrow R\_M1$	0.0054

We also examine the effects of the frequency factor. In total, 15 connections are identified as statistically significant difference due to various frequency levels (H(0.75Hz), M(0.5Hz), and L(0.25Hz)). We omit the detail here due to the space limitation. Among them, connection $L\_M1 \rightarrow L\_SMA$ is particularly interesting because of its monotonous trend along frequency levels. We will further employ multiple comparison method to find out whether different groups within the same frequency also indicate significant difference between NP and PA, and such difference is somehow improved by medicine.

Table 2. Effects of medicine on brain connections given in Table 1

Connection	$M_{NP} - M_{PA}$	$M_{NP} - M_{PA}$	$M_{NP} - M_{PB}$
$R\_PUT \rightarrow L\_PUT$	0.1076	0.0241	0.0835
$L\_THA \rightarrow L\_THA$	-0.2596	-0.1935	-0.0661
$R\_THA \rightarrow R\_M1$	-0.0583	0.083	-0.1413
$R\_THA \rightarrow L\_ACC$	-0.0959	-0.1334	0.0375
$L\_CER \rightarrow R\_M1$	-0.078	-0.0836	0.0056
$R\_M1 \rightarrow R\_SMA$	-0.1521	-0.1146	-0.0375
$L\_PFC \rightarrow R\_ACC$	-0.0037	0.0727	-0.0764
$L\_ACC \rightarrow L\_PUT$	-0.0065	0.0491	-0.0556
$L\_GLP \rightarrow R\_CER$	-0.0046	0.1275	-0.1321
$R\_GLP \rightarrow R\_M1$	-0.083	-0.1011	0.0181

4. CONCLUSION

We have presented a Probabilistic Boolean Network (PBN)-based framework for inferring brain functional connectivity and for revealing brain region-region interaction profiles observed in PD. The feasibility of the proposed method was demonstrated by applying it to a real fMRI study in PD. The detected connectivity differences between the PD group and the control group suggest that brain connectivity can be a sensitive marker for PD development and may provide a better understanding of the underlying mechanisms of PD. Future work will focus on improving the group analysis in fMRI using PBNs. More specifically, different group structure models (e.g. individual structure model, common structure model) will be investigated for studying inter-subject variability, and more rigorous comparisons between different group analysis approaches will be done statistically.

5. REFERENCES

[1] M. Goncalves, D. Hull, "Connectivity analysis with structural equation modeling: An example of the effects of voxel selection", *NeuroImage*,vol. 20, pp.1455-1467, 2003.

[2] S. Kim, E. Dougherty, M. Bittner et al., "General nonlinear framework for the analysis of gene interaction via multivariate expression arrays", *J.Biomed.Opt*, vol. 5, pp. 411-424, 2000.

[3] K. Friston, L. Harrison,and W. Penny, "Dynamic causal modeling", *NeuroImage*, vol. 19, pp.1273-1302, 2003.

[4] J. Li, Z. Wang et al., "Dynamic beyesian networks (DBNS) demonstrate impaired brain connectivity during performance of simultaneous movements in parkinson's disease", 2005.

[5] I. Shmulevich, E. Dougherty, S. Kim, W. Zhang, "Probabilistic boolean networks: A rule-based uncertainty model for gene regulatory networks", *Bioinformatics*, vol. 18, pp.261-274, 2002.

[6] K. Worsley, K. Friston, "Analysis of fMRI time-series revisited - again." *NeuroImage*, vol. 2, pp.173-181, 1995.

[7] M. Borga, O. Friman, P. Lundberg et al., "A canonical correlation approach to exploratory data analysis in fMRI", 2002.

[8] R. Duan, H. Man, W. Jiang, W. Liu, "Activation detection on fMRI time series using hidden markov model", 2005.

[9] R. Patel, F. Bowman, J. Rilling, "Determining Hierarchical Functional Networks from Auditory Stimuli fMRI." *Human Brain Mapping*, vol. 27, pp.462-470, 2006.