RESEARCH ARTICLE OPEN ACCESS

Clustering Brain Region from Functional MRI Using Interaction K-MEANS

Ms.P.Angalaparameswari (Me), Mrs.P.Rupa Ezhil Arasi Me (Ap/Cse),
Department of cse, Muthayammal Engineering College, Namakkal.
Department of cse, Muthayammal Engineering College, Namakkal.

ABSTRACT
Functional Magnetic Resonance Imaging (FMRI) is a technique for measuring brain activity that directly measures the blood flow in the brain there by providing information on brain activity. To understand the complex interaction patterns among brain regions we propose a novel clustering technique. We model each subject as multivariate time series, where the single dimensions represent the FMRI signal at different anatomical regions. Our objective is to assign objects exhibiting a similar intrinsic interaction pattern to a common cluster. Based on this novel cluster notion, we propose Interaction K-means (IKM), an efficient algorithm for partitioning clustering. The results on FMRI studies demonstrate the potential of IKM to contribute to a better understanding of normal brain function.

Keywords- Clustering, Interaction patterns, Multivariate time series.

I. Introduction

Functional magnetic resonance imaging (FMRI) opens up the opportunity to study human brain function in a noninvasive way. The basic signal of FMRI relies on the blood-oxygen-level-dependent(BOLD) effect, which allows indirectly imaging brain activity by changes in the blood flow related to the energy consumption of brain cells. Recently, resting-state FMRI has attracted considerable attention in the neuroscience community[1]. In resting state fMRI, subjects are instructed to just close their eyes and relax while in the scanner. fMRI data are time series of 3-dimensional volume images of the brain.

Data from FMRI experiments are massive in volume with more than hundred thousands of voxels and hundreds of time points. Since these data represent complex brain activity, also the information content can be expected to be highly complex. Recent findings suggest a modular organization of the brain into different functional modules. In Tracking cognitive processes with functional MRI mental chronometry it says that Functional magnetic resonance imaging (fMRI) is used widely to determine the spatial layout of brain activation associated with specific cognitive tasks at a spatial scale of millimeters. Recent methodological improvements have made it possible to determine the latency and temporal structure of the activation at a temporal scale of few hundreds of milliseconds. Despite the sluggishness of the hemodynamic response, fMRI can detect a cascade of neural activations — the signature of a sequence of cognitive processes. Decomposing the processing into stages is greatly aided by measuring intermediate responses. By combining event-related fMRI and behavioral measurement in experiment and analysis, trial-by-trial temporal links can be established between cognition and its neural substrate.

To obtain a better understanding of complex brain activity, it is essential to understand the complex interplay among brain regions during task and at rest. Inspired by this idea, we propose a novel technique for mining the different interaction patterns in healthy and diseased subjects by clustering. At the core of our method is a novel cluster notion: A cluster is defined as a set of subjects sharing a similar interaction pattern among their brain regions. A cluster analysis of motion stream data potentially identifies clusters with similar movements, usually performed by different persons. Time series data are frequently large and may contain outliers. In addition, time series are a special type of data set where elements have a temporal ordering. Therefore clustering of such data stream is an important issue in the data mining process.

Numerous techniques and clustering algorithm shave been proposed earlier to assist clustering of time series data streams. The clustering algorithms and its effective nesson various applications are compared to develop a new method to solve the existing problem. After standard pre-processing including parcellation into anatomical regions, we model each subject as a data object which is represented by a multivariate time series. Each of the dimensions is a time series corresponding to the fMRI signal of a specific anatomical brain region. Each of the dimensions is a time series corresponding to the FMRI signal of a specific anatomical brain region.

The volume of time series data generated by a fiber monitoring system can be huge. This limits the applicability of data mining algorithms to this problem domain. A widely used solution is to reduce the data size through feature extraction. Our approach Interaction K-means (IKM) simultaneously clusters the data and discovers the relevant cluster specific interaction patterns[2]. The algorithm IKM is a general technique for clustering multivariate time series and not limited to FMRI data. In this paper we explain IKM to demonstrate potential to discover interaction patterns among brain regions from FMRI data.

II. Imaging Techniques

To view activities or problems within the human brain without invasive surgery we use scanning techniques. CT(Computed Tomography) scan builds up a picture of the brain based on the differential absorption of X-rays but it reveal the gross features of the brain but do not resolve its structure well[3]. EEG(Electro Encephalography)[4] measure the electrical activity of the brain by according to the electrodes placed on the scalp but detecting changes in the brain is done on mille-sec level only.MEG(Magneto encephalography) measure the magnetic field produced by electrical activity in the brain e.g. speed of thought. It measures brain activity in 1/1000 of sec but it is more expensive[5].

ERP(Event related potentials) [6] provides the electrodes on the scalp measure voltage fluctuations resulting from electrical activity in the brain. But in ERP patients undergone brain surgery or to localize ERP sources is to place electrodes in the brain. SSPT (Steady State probe Topography) [7] measure steady state visually evoked potentials and records at the rate of 13 times per second from 64 electrodes in a light weight skull cap. PET (Positron Emission Tomography) [8] measures blood flow by injecting people with radioactive water and measure changes in radiation.

FMRI produce images at higher resolution than PET. It involves rapid scanning of the brain to see which area of brain becomes activated. It is highly sensitive so small changes also detected and multiple scans can be done on the same subject. By using concentration level we can identify the abnormality detection as shown below

1. Normal - 90%
2. Hypothermia – below 90%
3. Organ Function – below 80% Changes

III. Clustering

In this section, we elaborate our cluster notion based on characteristic interaction patterns. We want to find clusters of objects, which are represented by multivariate time series[9] sharing a common cluster-specific interaction pattern among the dimensions. Hidden Markov models[10] it measure only the similarity nerve cells only. Sequence Clustering Refinement[11] algorithm it acts as a K-means but it does not cluster different regions. Independent component Analysis [12] cluster it measures only the independent nerve cells. Nearest hyper plane distance neighbor clustering algorithm [13] that can be applied to high dimensional data, it cannot be applied to low dimensional data and it does not provide efficient clustering.

An Automatic Unsupervised Classification [14] observed that the grey matter distance can best separate the Alzheimer's disease patients from the cognitively normal control but it takes more time to diagnose the diseases. Fuzzy clustering[15] algorithms classify the tissues based on single channel MRI-Images but after clustering noise and misclassification error may present. Hierarchical clustering analyzes micro array data and makes it easier to interpret the results of a cluster analysis but it diagnoses only Alzheimer's disease only. Recently, compression-based similarity measures have been proposed.

We propose Interaction K-means (IKM), [2] a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns. Existing approaches to clustering multivariate time series do not consider interaction information. Most approaches rather extract features from each dimension and cluster the resulting feature vector.To demonstrate that the information on attribute interaction is valuable for clustering multivariate time series, we compare IKM to feature-based approaches. As a baseline, we consider classical K-Means clustering with Euclidean distance, which we term *Naive* in the following. The naive algorithm considers the concatenated dimensions of an object as a feature vector. Similar to classical K-means, IKM is an iterative algorithm, which efficiently converges towards a local minimum of the optimization space.

In brain, there are four regions namely Frontal lobe, Temporal lobe, Occipital lobe, Parietal lobe is shown in Figure 1

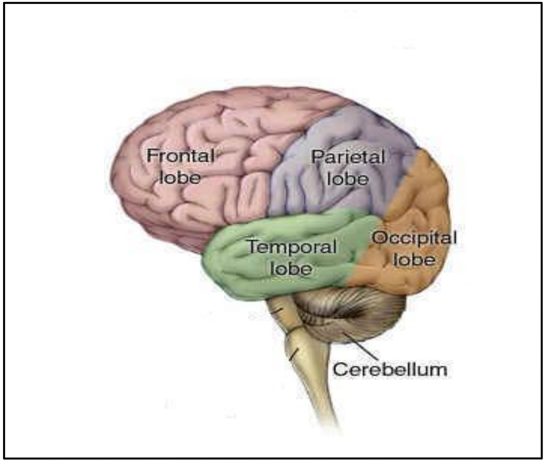

Figure 1. Brain Regions

While clustering first it takes time series from any regions after that it perform image segmentation then preprocessing. In pre-processing, we use low pass filter to remove the noise in that image. After that, we use Interactive k-means algorithm it clusters the normal nerves cells clustered in one group and abnormal nerve cells in another group. From that, abnormality is detected.

IV. Proposed System

In this section, we introduce the algorithm interaction K-means (IKM) which minimizes the clustering objective function. Similar to classical K-means,[16] IKM is an iterative algorithm, which efficiently converges towards a local minimum of the optimization space.

Algorithm IKM

Analogously to K-means, the first step of IKM is the initialization. As a common strategy for K-means, we propose to run IKM several times with different random initializations and keep the best overall result. For initialization, we randomly partition DS into K clusters. For IKM it is favorable that the initial clusters are balanced in size to avoid overfitting. Therefore, we partition the data set into K equally sized random clusters.

After assignment, in the update step, the models of all clusters are reformulated. Pseudo code of IKM is provided in Figure 2. As an iterative partitioning clustering algorithm, IKM follows a similar algorithmic paradigm as K-means. The similarity measure applied in IKM is the errors with respect to the set models of a cluster. This similarity measure is always evaluated between an object and a cluster, and not between two objects. In contrast to K-means or K-medoid algorithms, we can not state that a data object is the representative of a cluster. The cluster representative in IKM is a set of models describing a characteristic pattern of Interaction among the dimensions.

From Smoothened analysis of K-means it was one of the most widely used clustering algorithms, drawing its popularity from its speed in practice. Recently, however, it was shown to have exponential worst-case running time. In order to close the gap between practical performance and theoretical analysis, the k-means method has been studied in the model of smoothed analysis. But even the smoothed analyses so far are satisfactory as the bounds are still super-polynomial in the number n of data point.In this paper, we settle the smoothed running time of the k-means method. We show that the smoothed number of iterations is bounded by a polynomial in n and 1=_, where _ is the standard deviation of the Gaussian perturbations. This means that if an arbitrary input data set is randomly perturbed, then the k-means method will run in in expected polynomial time on that input set.

IKM converges as soon as no object changes its cluster assignment during two consecutive iterations. Usually, a fast convergence can be observed (less than 50 iterations on our experimental data), but there are some rare cases in which IKM does not converge. Analogously to standard K-means, it can be straightforwardly proven that the assignment and the update step strictly decrease the objective function.

However, due to the greedy stepwise algorithm applied for model finding, cf. the strictly monotonic property is lost. In particular, in different iterations of IKM, BIC may select different numbers of explanatory variables to be included in the cluster models. We therefore propose to terminate after max iterations. Our experiments demonstrate that this has no negative effects on the quality of the clustering result. Recently, following the approach of smoothed analysis Arthur et al. [2] have proven that K-means converges in polynomial time on an arbitrary input data set subjected to random perturbations.

A major advantage of IKM is possibility to interpret the detected interaction patterns. To facilitate interpretation, we focus on a subset of the models which best differentiates among the clusters. For each pair of clusters, the best discriminating models are selected by leave-one-out validation using objects of the corresponding clusters. This result theoretically supports our approach, since the fact that slightly different explanatory variable selection can happen in consecutive iterations is similar to minor perturbations of the data.

```
algorithm IKM (data set DS, integer K):
Clustering C
Clustering best Clustering;
//initialization
for init := 1 . . . maxInit do
    C := randomInit(DS, K);
    for each C ∈Cdo
        MC:= findModel(C);
    while not converged or iter < maxIter do
        //assignment
        for each O ∈ DS do
            O.cid = $\min_{C \in C} EO,C$
        //update
        for each C ∈Cdo
            Mc:= findModel(C);
        if improvement of objective function
            bestClustering := C;
    end while
end for
return bestClustering;
```

Figure2. Algorithm Interaction K-Means

As for ordinary K-means, the runtime of IKM[2] scales linearly with the number of objects n, since the complexity of assignment step is linear in n and usually only a few iterations are performed. Clearly, the update step is the most computationally expensive step, since model finding involves matrix inversion with complexity of $O(d^2)$ combined with the greedy stepwise algorithm with complexity of $O(d^3)$.

V. Results And Discussion

A major advantage of IKM is possibility to interpret the detected interaction patterns. To facilitate interpretation, we focus on a subset of the models which best differentiates among the clusters. For each pair of clusters, the best discriminating models are selected by leave-one-out validation using objects of the corresponding clusters. The clustering result together with the information about which models best discriminate among clusters is a good basis for user interaction.

Our results show reduced FA values in all sub-regions of the corpus callosum and confirms the altered interhemispheric connectivity in TS (Tourette syndrome). The alterations in the corpus callosum in our data are associated with TS itself, not with the comorbidities, as the results ofthe subsample without comorbidities indicate.

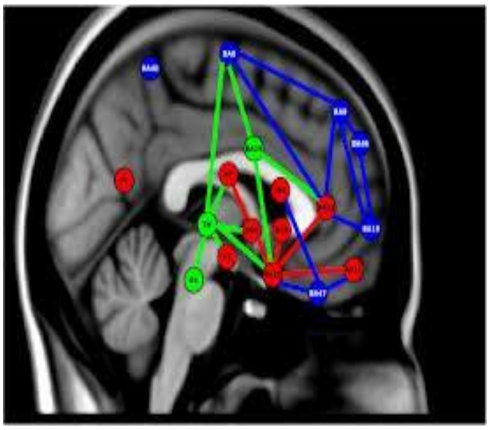

Figure 3. Clustering Interaction Patterns

- FA reductions in TS patients are displayed in the upper row in red.
- Alterations overlap to a large degree with the increases in radial diffusivity which are depicted in blue in the lower row.
- Background: mean tract skeleton (green)

In summary, reduced FA and increased radial diffusivity in adult Tourette patients affects different systems and brain regions. The pathologic pattern reaches beyond the corticospinal tract and affects interhemispheric connections such as the corpus callosum as well as intrahemishperic long associations fibre tracts. The affection of limbic structures such as the uncinate fascicle and its close correlations with the motor and The alteration pattern of decreased FA and increased radial diffusivity might indicate a deficit in myelination as the underlying pathophysiologic principle.

VI. Conclusion

We define a cluster as a set of objects sharing a specific interaction pattern among the dimensions. In addition, we propose interaction K-means (IKM) [2], an efficient algorithm for interaction-based clustering. Our experimental evaluation demonstrates that the interaction-based cluster notion is a valuable complement to existing methods for clustering multivariate time series. IKM achieves good results on synthetic data and on real world data from various domains, but especially excellent results on EEG and FMRI data. Our algorithm is scalable and robust against noise. Moreover, the interaction patterns detected by IKM are easy to interpret and can be visualized.

REFERENCES

[1] M. D. Fox and M. E. Raichle, "*Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging," Nat Rev Neurosci, vol. 8, no. 9, pp. 700–711, 2007.*

[2] M. Vlachos, J. Lin, E. Keogh, and D. Gunopulos, "*A waveletbased anytime algorithm for k-means clustering of time series*," in *In Proc. Workshop on Clustering High Dimensionality Data and Its Applications*, 2003, pp. 23–30.

[3] Tetsuya Yuasa and Tohoru Takeda, "*Fluorescent X-Ray Computed Tomography Using Synchrotron Radiation Towards Molecular Imaging*" Prof. Bernhard Schaller (Ed.), ISBN: 978-953-51-0359-2.

[4] Boris V. Baryshnikov*, BarryD.van Veen, Fellow, *IEEE*, and Ronald T. Wakai "*Maximum-Likelihood Estimation of Low-Rank Signals for Multiepoch MEG/EEG Analysis*" in IEEE TRANSACTIONS ON BIOMEDICALENGINEERING, VOL. 51, NO. 11, NOVEMBER 2004

[5] Ryusuke Kakigi,Koji Inui,Diep Tuan Tran Yunhai Qiu,XiahongWang, ShokoWatanabe Minoruhoshiyama," *Human Brain Processing and Central Mechanisms of Pain as Observed by ElectroandMagnetoencephalography*" in Ryusuke Kakigi *et al.* Journal of the Chinese Medical Association Vol. 67, No. 8

[6] L. Xiao, H. Begleiter, B. Porjesz, W. Wenyu, and A. Litke, "*Event related potentials during object recognition tasks," Brain Research Bulletin*, vol. 38, pp. 531–538(8), 1995.

[7] J. Hamilton, G. Chen, M. Thomason, M.Schwartz,andI.Gotlib, "*Investigating neural primacy in major depressive disorder: multivariate granger causality analysis of restingstate fmri time-series data*," *Mol Psychiatry*, vol. 16, no. 7, pp.763–72, 2011.

[8] Ahmadreza Rezaei1, Michel Defrise2, Girish Bal3, Christian Michel3, Maurizio Conti3, Charles Watson3, Johan Nuyts1," *Simultaneous Reconstruction of Activity and Attenuation in Time-of-Flight PET*".

[9] T.Oates, "*Identifying distinctive subsequences in multivariate time series by clustering*," in *Proc. ACM SIGKDD*, 1999, pp.322–326.

[10] T. Oates, L. Firoiu, and P. R. Cohen, "*Clustering time series with hidden markov models and dynamic time warping*," in *In Proceedings of the IJCAI-99 Workshop on Neural, Symbolic and Reinforcement Learning Methods for Sequence Learning*, 1999, pp.17–21.

[11] Ji He*y*, Man Lan*z*, Chew-Lim Tan*z*, Sam-Yuan Sung*z*, and Hwee-Boon Low*]*, *"Initialization of Cluster Refinement Algorithms:A Review and Comparative Study" yz*School of Computing, (National University of Singapore,3 Science Drive 2, Singapore 117543).

[12] E. H. C.Wu and P. L. H. Yu, "*Independent component analysis for clustering multivariate time series data,*" in *ADMA*, 2005,pp. 474–482.

[13] J. Lin, M. Vlachos, E. J. Keogh, D. Gunopulos, J.-W. Liu, S.-J.Yu, and J.-J. Le, "*A mpaa-based iterative clustering algorithm augmented by nearest neighbors search for time-series data streams*," in *PAKDD*, 2005, pp. 333–342.

[14] A. Trouve and Y. Yu, "*Unsupervised clustering trees by nonlinear principal component analysis." Pattern recognition and Image Analysis*, vol. 2, pp. 108–112, 2001.

[15] Michel Menard," *Fuzzy clustering and switching reg ression models usingambiguity and distance rejects*" in (Laboratoire d'Informatique et d'Imagerie Industrielle, Universit_e de La Rochelle, Avenue Marillac,17042 La Rochelle Cedex 1, France Received 12 August 1998); received in revised form 7 March 2000; accepted 22 March 2000.

[16] J. B. MacQueen, "*Some methods for classification and analysis of multivariate observations,"* in *Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, L. M. L.Cam and J. Neyman, Eds., vol. 1. (University of California Press, 1967, pp. 281–297.)

[17] Leckman 2003, Brain Dev. 2. Albin, Mink 2006, Trends Neurosc. 3. Smith et al. 2006, NeuroImage. 4. Smith et al. 2004, NeuroImage. 5. Smith, Nichols 2009, NeuroImage. 6. Wakana et al. 2004, Radiology. 7. Hua "*White Matter Abnormalities in Tourette Syndrome Extend Beyond Motor Pathways*"et al. 2008, NeuroImage.