Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 12, December 2014, pg.146 – 150

RESEARCH ARTICLE

Mining Interaction Patterns among Brain Regions by Clustering Based Interaction K-means

Amruta S. Chougule
Department of Computer Engineering, Pune University, India
Email: chouguleamruta24@gmail.com

S.A.Bajpai
Department of Computer Engineering, Pune University, India
Email: sanchi.scriet@gmail.com

Abstract—Functional magnetic resonance imaging (fMRI) provides the potential to study brain function in a non-invasive way. Massive in volume and complex in terms of the information content, fMRI data requires effective, and efficient data mining techniques. Recent results from neuroscience suggest a modular organization of the brain. 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. In contrast to previous approaches, we base our cluster notion on the interactions between the univariate time series within a data object. Our objective is to assign objects exhibiting a similar intrinsic interaction pattern to a common cluster. To formalize this idea, we define a cluster by a set of mathematical models describing the cluster-specific interaction patterns. Based on this novel cluster notion, we propose interaction K-means (IKM), an efficient algorithm for partitioning clustering. An extensive experimental evaluation on benchmark data demonstrates the effectiveness and efficiency of our approach. The results on two real fMRI studies demonstrate the potential of IKM to contribute to a better understanding of normal brain function and the alternations characteristic for psychiatric disorders.
Keywords— Clustering, multivariate time series, interaction patterns

INTRODUCTION

Human brain activity is very complex and far from being fully understood. Many psychiatric disorders like Schizophrenia and Somatoform Pain Disorder can so far neither be identified by biomarkers, nor by physiological or histological abnormalities of the brain. Aberrant brain activity often is the only resource to understand psychiatric disorders. Functional magnetic resonance imaging (fMRI) opens up the opportunity to study human brain function in a non-invasive 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. In atypical fMRI experiment, the subject performs some cognitive task while in the scanner. Recently, resting-state fMRI has attracted considerable attention in the neuroscience community [1]. Surprisingly, only about 5% of the energy consumption of the human brain can be explained by the task-related activity. Many essential brain functions, e.g. long-term memory are largely happening during rest, most of them without consciousness of the subject and many of them are still not well understood. Therefore recent findings support the potential of resting-state fMRI to explore the brain function in healthy subjects and reveal alternations characteristic for psychiatric disorders (e.g. [2]). 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. The data is traditionally analyzed within a mass-univariate framework essentially relying on classical inferential statistics, e.g. contained in the software package SPM [3]. A typical statistical analysis involves comparing groups of subjects or different experimental conditions based on univariate statistical tests on the level of the single 3-d pixels called voxels. 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. Only a small part of this information is accessible by univariate statistics. To make more of the potentially available information accessible, we need effective and efficient multivariate data mining methods.

EXISTING SYSTEM:

- In existing system for determining the really relevant dimensions there is no greedy stepwise algorithm for model finding and Bayesian Information Criterion (BIC) as evaluation criterion.

- There is clustering of data set is available but re-clustering is not used.

- A compression-based similarity measure is also proposed to compare long time series structure using co-compressibility as a dissimilarity measure. The authors report impressive results in many applications, but this technique requires certain statistical conditions from data

PROPOSED SYSTEM:

- We introduce a novel cluster notion for clustering multivariate time series based on attribute interactions.
- We proposed Interaction K-means(IKM), a partitioning clustering algorithm suitable to detect clusters of objects with similar interaction patterns.

- We demonstrate that the information on interaction patterns provide valuable insights for interpretation.

- Motivated by a real challenge from a neuroscience application IKM outperforms state-of-the-art techniques for clustering multivariate time series on synthetic data as well as on benchmark data sets from different applications.

- On FMRI data from studies on Somatoform Pain Disorder and Schizophrenia, our algorithm detects very interesting and meaningful interaction patterns.

RELATED WORK:

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective.

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

Modules:

a) Browse Dataset
b) Finding Model
c) IKM Clustering

Descriptions:

a) Browse Dataset:-

In this module user first logged into the System then browse dataset to view all images from dataset. Then finding model and IKM clustering is apply on this dataset .

b) Finding Model :-

Greedy stepwise algorithm is use for model finding in combination with the Bayesian Information Criterion (BIC). BIC is determines a balance between goodness-of-fit and complexity of the model and is defined by:

$BIC(Ma) = -2 \cdot LL(a,Ma) + \log(m^*)(|V| + 1).$

c) Iterative K-means:-

```
algorithm IKM (data set DS, integer K): Clustering 𝒞
  Clustering bestClustering;
//initialization
  for init := 1...maxInit do
    𝒞 := randomInit(DS, K);
    for each C ∈ 𝒞 do
      ℳ_C := findModel(C);

    while not converged or iter < maxIter do
//assignment
      for each O ∈ DS do
        O.cid = min_{C∈𝒞} ℰ_{O,C}
//update
      for each C ∈ 𝒞 do
        ℳ_C := findModel(C);
       if improvement of objective function
        bestClustering := 𝒞;
    end while

  end for
return bestClustering;
```

CONCLUSION

In this paper, we propose a novel cluster notion for multivariate time series. 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), 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. Nonlinear models show their superiority in the corresponding real world data. In ongoing and future work, we plan to extend our ideas to differential equations. We want to consider different models for different regions of the time series. We intend to work on methods for suitable initialization of IKM, since existing strategies for K-means can not be straightforwardly transferred to IKM because of the special cluster notion. We are also investigating in feature selection for interaction-based clustering.motivated by our interest in text clustering, FRECCA is a generic fuzzy clustering algorithm that can in principle be applied to any relational clustering problem, and application to several nonsentence data sets has shown its performance to be comparable to Spectral. Graph-based methods are an exciting area of research within the pattern recognition community. We have already mentioned some of the new work we are conducting in this area; however, what we are most excited about is extending the technique to perform hierarchical clustering. The concepts present in natural language documents usually display some type of hierarchical structure, whereas the algorithm we have presented in this paper identifies only flat clusters. Our main future objective is to extend these ideas to the development of a hierarchical fuzzy relational clustering algorithm.

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] C. Sorg *et al*., "Selective changes of resting-state networks in individuals at risk for alzheimer's disease," *PNAS*, vol. 104, no. 47, pp. 18760–18765, 2007.

[3] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Boston, MA, USA: Elsevier, 2007.

[4] S. Smith *et al*., "Correspondence of the brain's functional architecture during activation and rest," *PNAS*, vol. 106, no. 31, p. 13040, 2009.

[5] C. Li, L. Khan, and B. Prabhakaran, "Feature selection for classification of variable length multiattribute motions," in *Multimedia Data Mining and Knowledge Discovery*, V. A. Petrushin and L. Khan, Eds. London, U.K.: Springer, 2007.

[6] C. Pamminger, *Bayesian Clustering of Categorical Time Series: An Approach Using Finite Mixtures of Markov Chain Models*. Saarbrücken, Germany: VDM Publishing Group, 2008.

[7] T. W. Liao, "Clustering of time series data – A survey," *Pattern Recognit.*, vol. 38, no. 11, pp. 1857–1874, 2005.

[8] E. Keogh and S. Kasetty, "On the need for time series data mining benchmarks: A survey and empirical demonstration," in *Proc. SIGKDD*, Edmonton, AB, Canada, 2002, pp. 102–111.

[9] J. Lin, M. Vlachos, E. Keogh, and D. Gunopulos, "Iterative incremental clustering of time series," in *Proc. EDBT*, Heraklion, Greece, 2004, pp. 106–122.

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

Acknowledgement

This is a great pleasure & immense satisfaction to express my deepest sense of gratitude & thanks to everyone who has directly or indirectly helped me in my project work successfully.

I express my gratitude towards project guide **Prof. S.A.Bajpai**, and **Prof. G.M.Bhandari**, Head of Department of Computer Engineering , Bhivarabai Sawant Institute Of Technology and Research College Of Engineering, Pune who guided & encouraged me in my project work in scheduled time. I would like to thanks our **Principal Dr. D.M.Yadav**, for allowing us to pursue my project in this institute.

