ISSN 2395-695X (Print)
ISSN 2395-695X (Online)

International Journal of Advanced Research in Biology Engineering Science and Technology (IJARBEST)
Vol. 2, Issue 4, April 2016

MINING TIME VARYING DATA BASED ON K- MEANS AND POSSIBILISTICS FUZZY C- MEANS ALGORITHM

S.PORTHAN PAUL[1] ,T.KUMARAGURUBARAN[2]

Department of Computer Science and Engineering[1], Department of Computer Science and Engineering[2]
Mohamed Sathak Engineering College, kilakarai, Ramanathapuram.
[1]porthanpaul796@gmail.com

Abstract**— The aim of co-clustering is identifying the block structures of the data matrices by clustering the rows and columns simultaneously into co-clusters. Here the data matrices evolve smoothly over the time in all the applications. Many approaches are simple to learn from these time evolving data matrices and analyze them separately but they neglect the time–dependent nature of those data. In this paper, we learn the time varying data by proposing the two formulations such as evolutionary co-clustering and feature selection process which is based on the fused Lasso regularization. The evolutionary co-clustering formulation is done by one based on sparse SVD and spectral co-clustering method which is able to identify smoothly varying hidden block structures embedded into the matrices along the temporal dimension. This formulation is flexible and allows for imposing smoothness constraints over only one dimension of the data matrices. Then the feature selection process is used for uncovering the shared features from the clustering such as K-means clustering and Possibilistic Fuzzy C-means clustering from the time evolving data matrices. The optimization process involved here are non-convex, non-smooth, non-separable. For the efficient computation problem, we develop a two step procedure for the iterative objective function process. Our evaluation process is done using the Allen Developing Mouse Brain Atlas data which shows the consistent formulation.**

Keywords**— Sparsity learning, Time-varying data,Co-Clustering, Feature Selection, Optimization, Bioinformatics, Neuroinformatics.**

I.INTRODUCTION

Sparse machine learning refers to a collection of methods to learning that seek a trade-off between some goodness-of-fit measure and sparsity of the result,the latter property allowing better interpretability.In a sparse learning classification task for example,the prediction accuracy or some other classical measure of performance is not the sole concern: we also wish to be able to explain what the classifier means to a non-expert.Thus,if the classification task involves say gene data,one wish to provide not only a high-performance classifier,but one that only involves a few genes,allowing biologists to focus their research efforts on those specific genes.There is an extensive literature on the topic of sparse machine learning,with terms such as compressed sensing,l1-norm penalties and convex optimization,often associated with the topic.Successful applications of sparse methods have been reported,mostly in image and signal processing, see for example.Due to the intensity of research in this area,and despite an initial agreement that sparse learning problems are more computationally difficult than their nonsparse counterparts,many very efficient algorithms have been developed for sparse machine learning in the recent past.

A new consensus might soon emerge that sparsity constraints or penalties actually help reduce the computational burden involved in learning.Sparse machine learning has recently emerged as powerful tool to obtain models of high-dimensional data with high degree of interpretability, at low computational cost.Clustering is the grouping together of similar objects. Given a collection of unlabeled documents, document clustering can help in organizing the collection thereby facilitating future navigation and search. A starting point for applying clustering algorithms to document collections is to create a vector space model. The basic idea is (a) to extract unique content-bearing words from the documents treating these words as features and (b) to then represent each document as a vector in this feature space.Thus the entire document collection may be represented by a word by-document matrix A whose rows correspond to words and columns to documents.A non-zero entry in A, say Aij ,indicates the presence of word i in document j,while a zero entry indicates an absence.Typically, a large number of words exist in even a moderately sized set of documents; for example, in one test case we use 4303 words in 3893 documents. However, each document generally contains only a small number of words and hence, A is typically very sparse with almost 99% of the matrix entries being zero.

Evolutionary network analysis has found an increasing interest in the literature because of the importance of different kinds of dynamic social networks, email networks, biological networks, and social streams. When a network evolves, the results of data mining algorithms such as community detection need to be correspondingly updated. Furthermore, the specific kinds of changes to the structure of the network, such as the impact on community structure or the impact on network structural parameters, such as node degrees, also need to be analyzed. Some dynamic networks have a much faster rate of edge arrival and are referred to as network streams or graph streams.

The analysis of such networks is especially challenging, because it needs to be performed with an online approach, under the one-pass constraint of data streams. The incorporation of content can add further complexity to the evolution analysis process. This survey provides an overview of the vast literature on graph evolution analysis and the numerous applications that arise in different contexts.

One major factor that is contributing to the growth of data size is the increasingly widespread ability to perform very large scale time-varying simulations. Although intensive research has been undertaken to optimize the performance of visualizing very large data sets, there remain many challenges for us to perform thorough analysis for large scale time-varying data sets. For the past ten years, the GRAVITY Research Group has strived to perform a comprehensive study of end-to-end solutions to facilitate efficient and effective analysis of large-scale time-varying data. The focus of our research includes:

- Temporal Trend Analysis and Visualization
- Time-Varying Feature Tracking
- Transfer Function Design for Time-Varying Volume Rendering
- Efficient Rendering of Time-Varying Data
- Data Representations and Indexing for Time-Varying Data

Bioinformatics has become an important part of many areas of biology.In experimental molecular biology,bioinformatics techniques such as image and signal processing allow extraction of useful results from large amounts of raw data.In the field of genetics and genomics,it aids in sequencing and annotating genomes and their observed mutations.It plays a role in the text mining of biological literature and the development of biological and gene ontologies to organize and query biological data. It also plays a role in the analysis of gene and protein expression and regulation. Bioinformatics tools aid in the comparison of genetic and genomic data and more generally in the understanding of evolutionary aspects of molecular biology. At a more integrative level it helps analyze and catalogue the biological pathways and networks that are an important part of systems biology.

Bioinformatics is the application of computer technology to the management of biological information. Computers are used to gather, store, analyze and integrate biological and genetic information which can then be applied to gene-based drug discovery and development. The need for Bioinformatics capabilities has been precipitated by the explosion of publicly available genomic information resulting from the Human Genome Project.

In this project, we focus on mining the time varying data by proposing the two formulations such as evolutionary co-clustering and feature selection process which is based on the fused Lasso regularization. The evolutionary co-clustering formulation is done by one based on sparse SVD and spectral co-clustering method which is able to identify smoothly varying hidden block structures embedded into the matrices along the temporal dimension. This formulation is flexible and allows for imposing smoothness constraints over only one dimension of the data matrices. Then the feature selection process is used for uncovering the shared features from the clustering such as K-means clustering and Possibilistic Fuzzy C-means clustering from the time evolving data matrices. The optimization process involved here are non-convex, non-smooth, non-separable.Evolutionary network analysis has found an increasing interest in the literature because of the importance of different kinds of dynamic social networks, email networks, biological networks, and social streams. When a network evolves, the results of data mining algorithms such as community detection need to be correspondingly updated. Furthermore, the specific kinds of changes to the structure of the network, such as the impact on community structure or the impact on network structural parameters, such as node degrees, also need to be analyzed. An evolutionary clustering should simultaneously optimize two potentially conflicting criteria: first, the clustering at any point in time should remain faithful to the current data as much as possible; and second, the clustering should not shift dramatically from one timestep to the next. Evolutionary co-clustering formulation for uncovering patterns from time-evolving data matrices.The proposed formulation encourages smooth changesin the row and/or column patterns over time, thereby capturing the time-evolving nature of the underlying process faithfully. The proposed framework is very flexible and can be applied to applications in which only one dimension of the data matrices evolves. In many application domains, each data matrix is usually associated with a particular time point, and it evolves smoothly over time. For example, in the developing mouse brain gene expression analysis, the spatial gene expression patterns at a particular developing time point is captured by a data matrix in which one dimension corresponds to the genes and the other dimension corresponds to the spatial locations. Since gene regulation acts sequentially, the expression patterns usually evolves smoothly over time, thereby resulting a series of time-stamped data matrices, one for each sampled developing time point. A simple approach for mining these time-evolving data matrices is to treat the data matrices at different time points separately.This approach, however, ignores the time-dependent nature of the underlying process, thereby yielding results that are not amenable to domain interpretation. In this project, we propose an evolutionary co-clustering formulation for uncovering patterns from time-evolving data matrices.The proposed formulation encourages smooth changes in the row and/or column patterns over time, thereby capturing the time-evolving nature of the underlying process faithfully. The proposed framework is very flexible and can be applied to applications in which only one dimension of the data matrices evolves.Our solutions to the evolutionary spectral clustering problems provide more stable and consistent clustering results that are less sensitive to short-term noises while at the same time are adaptive to longterm cluster drifts.

Furthermore, we demonstrate that our methods provide the optimal solutions to the relaxed versions of the corresponding evolutionary k-means clustering problems. Performance experiments over a number of real and synthetic data sets illustrate our evolutionary spectral clustering methods provide more robust clustering results that are not sensitive to noise and can adapt to data drifts.

Consider the minimization of a smooth convex function regularized by the composite prior models. This problem is generally difficult to solve even if each subproblem regularized by one prior model is convex and simple. In this paper, we present two algorithms to effectively solve it. First, the original problem is decomposed into multiple simpler subproblems. Then, these subproblems are efficiently solved by existing techniques in parallel. Finally, the result of the original problem is obtained by averaging solutions of subproblems in an iterative framework. The proposed composite splitting algorithms are applied to the compressed MR image reconstruction and low-rank tensor completion. Numerous experiments demonstrate the superior performance of the proposed algorithms in terms of both accuracy and computation complexity.

Spatial gene expression profiles provide a novel means of exploring the structural organization of the brain. Computational analysis of these patterns is made possible by genome-scale mapping of the C57BL/6J mouse brain in the Allen Brain Atlas.Biclustering consists in simultaneous partitioning of the set of samples and the set of their attributes (features) into subsets (classes). Samples and features classified together are supposed to have a high relevance to each other. In this paper we review the most widely used and successful biclustering techniques and their related applications.

II. RELATED WORKS

The system proposes an evolutionary co-clustering formulation for identifying the co-clusters from the time variant data. This formulation employs the sparsity-inducing regularization for identifying the block structures from the time variant data matrices. This process applies the fused lasso regularization encourage the temporal smoothness over the block structures identified from the contiguous time points. This formulation is very flexible and can be applied to encourage temporal smoothness over either one or both dimensions of the data matrices. It also uses an iterative two-step procedure to compute the solution of t general optimization problem. The experimental result shows the formulation is best compared with the existing system. Simultaneous row and column clustering for identifying block structures from matrix data has been initially studied in . Recent surge of interests in co-clustering is motivated by biological applications, which aim at identifying subset of genes co-expressed in a subset of samples from microarray gene expression data . Co-clustering has also been applied in many other applications, including simultaneous clustering of words and documents authors and conference , etc. Early work on co-clustering focuses on defining an error measure and then identifying blocks that minimize this measure using heuristic search algorithms. These early work has recently been reformulated using matrix and optimization techniques. Following the spectral clustering formalism, it has been shown recently that co-clustering is closely related to the SVD of the data matrix . In , co-clustering is formulated as a bipartite graph cut problem, and the data are projected onto the left and right singular vector spaces before they are concatenated and clustered to identify row and column co-clusters. It is shown in that sparsityinducing regularization can be employed to compute sparse singular vectors, which in turn can be used to form co-clusters. In a framework for simultaneous co-clustering and predictive learning is proposed. This work is also related to recent studies on mining from time-evolving data. In the spectral clustering formalism is systematically extended to the evolutionary setting by incorporating a temporal cost into the objective function, leading to a suite of formulations for evolutionary spectral clustering. In the nonnegative matrix factorization is employed for soft clustering, and a temporal cost is included for mining from timeevolving data. The broad area of evolutionary network analysis is reviewed in The fused Lasso penalty was originally proposed in for encouraging smoothness over related coefficients in regression problems. This type of penalty is very attractive and has been applied for encouraging smoothness over spatial and temporal smoothness in many applications, including biological data analysis and social studies. A critical challenge in employing the fused Lasso formalism is that this class of penalty is non-smooth and non-separable and thus is very challenging to optimize. In a modified coordinate descent algorithm is developed to solve the fused Lasso formulation. However, this algorithm is not guaranteed to give the exact solution. In a path algorithm is proposed to solve the fused Lasso signal approximator. Instead of solving the original primal problem, Liu et al. developed a dual formulation for the fused Lasso signal approximator and devised a gradient descent algorithm for computing the dual solution .Similar formulations and algorithms have been studied in the compressive sensing literature,The evolutionary clustering paradigm is related, but different from, the currently studied evolutionary feature selection formalism. Specifically, the smoothness constraints are imposed on the sample dimension in evolutionary clustering, while similar constraints are imposed on the feature dimension in evolutionary feature selection. Consequently, the clustering results are expected to evolve smoothly in evolutionary clustering, while the selected features are shared across time points in evolutionary feature selection.

III. PROPOSED SCHEME

In this proposed system,we learn the time varying data by proposing the two formulations such as evolutionary co-clustering and feature selection process which is based on the fused Lasso regularization.The evolutionary co-clustering formulation is done by one based on sparse SVD and spectral co-clustering method which is able to identify smoothly varying hidden block structures embedded into the matrices along the temporal dimension.This formulation is flexible and it allows for imposing smoothness constraints over only one dimension of the data matrices.Then the feature selection process is used for uncovering the shared features from the clustering such as K-means clustering and Possibilistic Fuzzy C-means clustering from the time evolving data matrices.The optimization process involved here are non-convex,non-smooth,non-separable.For the efficient computation problem,we develop a two step procedure for the iterative objective function process.Our evaluation process is done using the Allen Developing Mouse Brain Atlas data which shows the consistent formulation. Results show that our formulations consistently outperform prior methods. Experimental results on the Allen Developing Mouse Brain Atlas data show that the proposed methods yield consistently higher performance in comparison to other methods.Performance results will compare with the other clustering methods results with the help of fuzzed lasso reguglaization.It will reduce the sparsity between the data intervals through two step procedure.We evaluate the proposed evolutionary co-clustering formulation and evolutionary feature selection formulation using the Allen Developing Mouse Brain Atlas data.This dataset contains in situ hybridization gene expression pattern images in the developing mouse brain seven developmental ages E11.5, E13.5, E15.5, E18.5, P4, P14, and P28. The 3D images are registered to a reference atlas separately for each age, and a regular grid is applied to partition the 3D brain space into voxels. The expression energy within each voxel is given as a numerical value. There is one data matrix associated with each of the seven developing ages. The rows of the matrices correspond to brain voxels while the columns correspond to genes. The reference atlas ontology is organized into a hierarchy, and we up-propagate the annotations to Level 3 and Level 5 in the experiments. It is well-known that the developing mouse brain is divided into grid-like patterns along the longitudinal and transversal dimensions and identification of genes co-expressed in these domains might elucidate the genetic mechanisms governing the mouse brain development. The transversal and longitudinal dimensions correspond to the Level 3 and Level 5 ontology, respectively. Table 1 shows the statistics on the number of genes, voxels and brain regions for each dataset on Level 3 and Level 5 annotations respectively. To measure the performance of our proposed methods, we consider the annotated brain region of each voxel as its class and compare the clustering results with the region labels of voxels, since it has been shown that the results of gene expression data clustering are largely consistent with classical neuroanatomy Followingthe normalized mutual information (NMI) and Rand index are used to quantitatively measure the correspondence of the clustering results with the classical neuroanatomy reflected in the region annotations. We use the duality gap as the stopping criterion for the gradient descent algorithm and the error tolerance is set to 108 in the experiments. Overall, the proposed formulations are efficient to solve on a regular desktop PC, but we do not provide detailed timing results due to space constraints. The performance of the three methods on the seven datasets is reported in Fig. 1. We observe that the best performance is achieved when ¼ 0:05 max where max is defined in Eq. (4.3) and report the results under this parameter setting. Detailed studies on parameter sensitivity are reported in the following. It can be observed from Fig. 1 that incorporation of the smoothness constraints between contiguous age data yield improved performance.

	E11.5	E13.5	E15.5	E18.5	P4	P14	P28
# of genes	1948	1948	1930	1946	1918	1906	1944
# of voxels	7122	13194	12148	12045	21845	24325	28023
# of Level 3 structures	20	20	20	20	20	19	20
# of Level 5 structures	82	77	76	65	64	71	74

TABLE 1
Statistics about the Mouse Brain Data at Annotation Level 3 and Level 5

IV. ALGORITHMS

a)K Means Clustering

k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed Apriori. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because of different location causes different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as barycenter of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated.

As a result of this loop we may notice that the k centers change their location step by step until no more changes are done or in other words centers do not move any n[illegible] nimizing an objective f[illegible] given by:

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\|x_i - v_j\|\right)^2$$

where,

'$||x_i - v_j||$' is the Euclidean distance between x_i and v_j.

'c_i' is the number of data points in i^{th} cluster.

'c' is the number of cluster centers.

Algorithmic steps for k-means clustering:

Let X = $\{x_1,x_2,x_3,\ldots,x_n\}$ be the set of data points and V = $\{v_1,v_2,\ldots,v_c\}$ be the set of centers.

1) Randomly select 'c' cluster centers.

2) Calculate the distance between each data point and cluster centers.

3) Assign the data point to the cluster center whose distance from the cluster center is minimum of all the cluster centers..

4) Recalculate the new cluster center using:

$$v_i = (1/c_i) \sum_{j=1}^{c_i} x_i$$

Where, 'c_i' represents the number of data points in i^{th} cluster.

5) Recalculate the distance between each data point and new obtained cluster centers.

6) If no data point was reassigned then stop, otherwise repeat from step 3).

b)Possibilistic Fuzzy C-means

This Algorithm establishes a connection between the possibilistic and probabilistic approaches. The FPCM creates both memberships (in the sense of relative belonging) and typicalities, along with the usual centroids through a standard alternating optimization process. The current approach solves the noise susceptibility characteristic of the FCM, and also overcomes the coincident clusters problem of the PCM. A modified version of the objective function is obtained by adding a possibilistic term, within the typicality, based on all the n data points rather on all the c centroids

$$J(U,V) = \sum_{i=1}^{c} \sum_{k=1}^{n} M_{ik} d^2(x_k, v_i)$$

Where $M_{ik} = u_{ik}^m + t_{ik}^\eta$, given positive reals m and ŋ. The membership formula constraint remains the same as in (2) and (3) respectively. Through distributing the typicalities w.r.t. all n samples, we get:

$$\sum_{k=1}^{n} t_{ik} = 1 , \quad \forall i$$

$$t_{ik} = \left(\sum_{j=1}^{c} \left(\frac{d^2(x_k, v_i)}{d^2(x_k, v_j)} \right)^{2/(\eta-1)} \right)^{-1} , \quad \forall i,k$$

$$v_i = \frac{\sum_{k=1}^{n} M_{ik} x_k}{\sum_{k=1}^{n} M_{ik}}, \quad \forall i$$

V. IMPLEMENTATION TECHNIQUES

a)Time Variant Data

It is the first module in our process. Here the Time-varying data are dynamic in nature and can be categorized by different temporal behaviors they exhibit. The first category of time-varying data is regular, which usually involves a certain phenomenon that grows, persists, and declines in several (distinct) stages. The rate of change at each stage could vary dramatically in space and time. Many natural phenomena and their simulations, such as the earthquake, fall into this category. The second category of time-varying data is periodic. For this type of data with recurring patterns, special attentions are paid to space-time abnormal events.

In our project we use the Allen Developing Mouse Brain Atlas data that contains high-resolution, three dimensional gene expression patterns in the mouse brain at multiple development stages.

b) Evolutionary Clustering

It is the second module in our project; here we use the co-clustering formulations for the uncovering patterns from time-evolving data matrices. For the co-clustering formulation, we use two methods they are, one based on sparse SVD and the spectral co-clustering method. Here the formulation encourages smooth changes in the row or the column patterns over time, based on capturing the time-evolving nature of the underlying process. This formulation method is very flexible and can be applied to the one dimension of the data matrices evolving applications.We propose an evolutionary co-clustering formulation for uncovering patterns from time-evolving data matrices.The proposed formulation encourages smooth changes in the row and/or column patterns over time, thereby capturing the time-evolving nature of the underlying process faithfully. The proposed framework is very flexible and can be applied to applications in which only one dimension of the data matrices evolves.

Sparse singular value decomposition (SSVD) is proposed as a new exploratory analysis tool for bi-clustering or identifying interpretable row–column associations within high-dimensional data matrices. SSVD seeks a low-rank, checkerboard structured matrix approximation to data matrices. The desired checkerboard structure is achieved by forcing both the left- and right-singular vectors to be sparse, that is, having many zero entries. By interpreting singular vectors as regression coefficient vectors for certain linear regressions.

c) Feature selection

This is the third process called Feature selection process in our project. Here we use the two step procedure for solving the optimization problem. In this process we need to analyze the data set at the different points separately. Here we use K-means clustering and Possibilistic Fuzzy C-means clustering for this process. To encourage the selection of shared features among the time varying data, we introduce the fused lasso term for the process the data. A simple approach for mining these time-evolving data matrices is to analyze them at different time points separately. In this section, we propose an evolutionary feature selection formulation for uncovering shared features from time evolving
data matrices. The proposed formulation encourages smooth changes of the features over time, thereby capturing the time-evolving nature of the underlying process faithfully.Consequently, the clustering results are expected to evolve smoothly in evolutionary clustering, while the selected features are shared
across time points in evolutionary feature selection.

k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed Apriori. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because of different location causes different result.

The FPCM creates both memberships (in the sense of relative belonging) and typicalities, along with the usual centroids through a standard alternating optimization process. The current approach solves the noise susceptibility characteristic of the FCM, and also overcomes the coincident clusters problem of the PCM. A modified version of the objective function is obtained by adding a possibilistic term, within the typicality, based on all the n data points rather on all the c centroids. We propose an iterative two-step procedure to compute the solution of the general optimization problem. Each of the iterative step involves a convex, but non-smooth and non-separable problem. To enable efficient optimization, we derive the dual form of this problem and employ a gradient descent algorithm to solve the smooth dual problem.

d) Performance Results

This is the final process in our project; we display the performance evaluation result and resultant graph. Here the resultant graph shows the comparison of the existing clustering methods and the proposed clustering methods such as K-means and the Possibilistic fuzzy c-means clustering process. To evaluate the proposed evolutionary feature selection formulation,we compare it with two other clustering methods; namely the K-means and the sparse K-means methods on the Allen Developing Mouse Brain Atlas data. Specifically, the smoothness constraints are imposed on the sample dimension in evolutionary clustering, while similar constraints are imposed on the featuredimension in evolutionary feature selection. Consequently, the clustering results are expected to evolve smoothly in evolutionary clustering, while the selected features are shared across time points in evolutionary feature selection. More importantly, the evolutionary feature selection method yields a set of shared features across developmental ages. These features correspond to genes in our datasets. Hence, our method identifies a set of genes that act continuously in multiple developmental ages. These genes might play important roles in the mouse brain development. We will analyze their functional and developmental roles in the future.

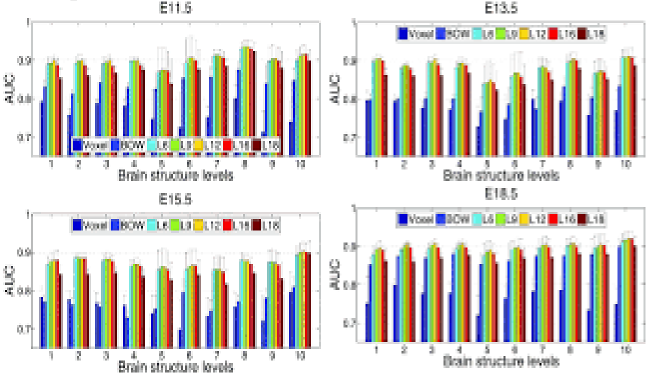

Fig. 1. Performance of the proposed method

VI CONCLUSION AND FUTUREWORKS

This system proposes an evolutionary co-clustering formulation for identifying the co-clusters from the time variant data. This formulation employs the sparsity-inducing regularization for identifying the block structures from the time variant data matrices. This process applies the fused lasso regularization encourage the temporal smoothness over the block structures identified from the contiguous time points. This formulation is very flexible and can be applied to encourage temporal smoothness over either one or both dimensions of the data matrices. It also uses an iterative two-step procedure to compute the solution of the general optimization problem. The experimental result shows the formulation is best compared with the existing system.In our future work, we learn the time varying data by proposing the two formulations such as evolutionary co-clustering and feature selection process which is based on the fused Lasso regularization. The evolutionary co-clustering formulation is done by one based on sparse SVD and spectral co-clustering method which is able to identify smoothly varying hidden block structures embedded into the matrices along the temporal dimension. This formulation is flexible and it allows for imposing smoothness constraints over only one dimension of the data matrices. Then the feature selection process is used for uncovering the shared features from the clustering such as K-means clustering and K-means++ clustering from the time evolving data matrices.

REFERENCES

[1] C. Aggarwal and K. Subbian, “Evolutionary network analysis: A survey,” ACM Comput. Surveys, vol. 47, no. 1, p. 10, 2014.

[2] S. Alelyani, J. Tang, and H. Liu, “Feature selection for clustering: A review,” Data Clustering: Algorithms and Applications, C. Aggarwal, and C. Reddy, Eds., Boca Raton, FL, USA: CRC Press.

[3] Allen Institute for Brain Science. (2012). Allen Developing Mouse Brain Atlas [Internet]. [Online].Available: http://developingmouse. brain-map.org

[4] O. Banerjee, L. E. Ghaoui, and A. d’Aspremont, “Model selection through sparse maximum likelihood estimation for multivariate Gaussian or binary data,” J. Mach. Learn. Res., vol. 9, pp. 485–516, 2008.

[5] J. W. Bohland, H. Bokil, S. D. Pathak, C. Lee, L. Ng, C. Lau, C. Kuan, M. Hawrylycz, and P. P. Mitra, “Clustering of spatial gene expression patterns in the mouse brain and comparison with classical neuroanatomy,” Methods, vol. 50, no. 2, pp. 105–112, 2010.

[6] S. Busygin, O. Prokopyev, and P. M. Pardalos, “Biclustering in data mining,” Comput. Oper. Res., vol. 35, pp. 2964–2987, Sep. 2008.

[7] D. Chakrabarti, R. Kumar, and A. Tomkins, “Evolutionary clustering,” in Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2006, pp. 554–560.

[8] C. Chen, J. Huang, L. He, and H. Li, “Preconditioning for accelerated iteratively reweighted least squares in structured sparsity reconstruction,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2013, pp. 2713–2720.

[9] Y. Cheng and G. M. Church, “Biclustering of expression data,” in Proc. Eighth Int. Conf. Intell. Syst. Mol. Biol., 2000, pp. 93–103.

[10] Y. Chi, X. Song, D. Zhou, K. Hino, and B. L. Tseng, “On evolutionary spectral clustering,” ACM Trans. Knowl. Discov. Data, vol. 3, pp. 17:1–17:30, Dec. 2009.

[11] H. Cho, I. S. Dhillon, Y. Guan, and S. Sra, “Minimum sum-squared residue co-clustering of gene expression data,” in Proc. Fourth SIAM Int. Conf. Data Min., 2004, pp. 114–125.

[12] M. Deodhar and J. Ghosh, “SCOAL: A framework for simultaneous co-clustering and learning from complex data,” ACM Trans. Knowl. Discov. Data, vol. 4, no. 3, pp. 11:1–11:31, 2010.

[13] I. S. Dhillon, “Co-clustering documents and words using bipartite spectral graph partitioning,” in Proc. Seventh ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2001, pp. 269–274.

[14] I. S. Dhillon, S. Mallela, and D. S. Modha, “Information-theoretic co-clustering,” in Proc. Ninth ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2003, pp. 89–98.

[15] C. Eckart and G. Young, “The approximation of one matrix by another of lower rank,” Psychometrika, vol. 1, no. 3, pp. 211–218, 1936.

[16] J. Friedman, T. Hastie, H. H€ofling, and R. Tibshirani, “Pathwise coordinate optimization,” Ann. Appl. Statist., vol. 1, no. 2, pp. 302–332, 2007.

[17] E. Giannakidou, V. Koutsonikola, A. Vakali, and Y. Kompatsiaris, “Co-clustering tags and social data sources,” in Proc. Ninth Int. Conf. Web-Age Inf. Manag., 2008, pp. 317–324.

[18] J. A. Hartigan, “Direct clustering of a data matrix,” J. Am. Stat. Associat., vol. 67, no. 337, pp. 123–129, 1972.

[19] H. H€ofling, “A path algorithm for the fused Lasso signal approximator,” J. Comput. Graph. Stat., vol. 19, no. 4, pp. 984–1006, 2010.

[20] J. Huang, S. Zhang, H. Li, and D. Metaxas, “Composite splitting algorithms for convex optimization,” Comput. Vis. Image Understanding, vol. 115, no. 12, pp. 1610–1622, 2011.