Integrating Mutiview clusters with Tensor Methods

Dr.A.Bharathi, [1]
Professor,
Dept of Information Technology,
Bannari Amman Institute of Technology,
Erode, India,
E-Mail id: bharathia@bitsathy.ac.in.

S.Anitha, [2]
PG Scholar,
Dept of Information Technology,
Bannari Amman Institute of Technology,
Erode, India,
E-Mail id:anithas.se12@bitsathy.ac.in.

Abstract-Integrating multiview cluster is an crucial issue in heterogeneous environment. Spectral clustering is used for integrating cluster in heterogeneous environment. In this paper, we used Principal Modularity Maximization based tensor decomposition for identifying hidden pattern in the context of spectral clustering. This gives the good result when compared to other methods. Here synthetic datasets are used for evaluating the results by comparing with Average Modularity Maximization and Total Modularity Maximization methods.

Key words Component- Multiview clustering, Tensor decomposition, Principal Modularity Maximization, Spectral Clustering.

INTRODUCTION

Datasets consists of multiple similarities. For a group of people, we might know their height, weight, age, education, location, designation and their family related. For a set of authors have their own papers, citations and methods. For a set of drives, they have own files, documents, locations and their contents. Our approach is to cluster people, researchers, methods, or files, to treat all the similarities concurrently.

In a set of multiple networks, they share same set of nodes but possess different types of connection between nodes. Multiple relationship can be formed through individual activity is called as multiview learning [2]. The recent development in clustering is the spectral clustering. Spectral clustering is based on the Ncut algorithm [1]. This can work well in the single view data as it is based on matrix decompositions. Many clustering algorithms have been proposed in comparison with the single view data. Therefore, these algorithms have some limitation.

Tensors are the higher order generalization of matrices .They can be applied to several domains such as web searching, image processing, data mining, and image recognition. Here tensor based methods are used to model multiview data. This is also used to detect the hidden pattern in multiview data subspace by tensor analysis. It works based on the tensor decomposition [1] which captures the multilinear structures in higher order data, where data has more than two modes. From the above example, in tensor, similarity of researchers is one slice, and then similarity citations are one slice. Likewise all slices will be combined to form tensor.

Tensor decomposition is used to cluster all the similarity matrices into set of compilation feature vector. Many clustering algorithms like k Means, SVD, HOSVD [15] are used for many tensor methods. Spectral clustering [1] is used for clustering the similarity matrices based on tensor methods.

MATERIALS AND METHODS

Tensor

Scalars, vectors and matrices are facilitated by higher order tensors. Scalars denote lower case letters (A, B…). Vectors are written in italic capitals (*, A, B….)* matrices correspond to boldface capitals (A, B**…)** and tensors are written as casteller lettering (**A, B**…). This notation is consistently used for ordering lower parts of a given quantity such as a vector a, matrix **A**, and tensor *A*, correspondingly. The Kronecker product is denoted by ⊗. Vector is first order, matrix is a second order and tensors are the third order or the higher order tensor.

Multiview Clustering

A multiview clustering method that extends k-means and hierarchical clustering to deal with data as two conditionally independent views [13]. Canonical correlation analysis in multiview clustering assumes that the views are uncorrelated in the given cluster label. These algorithms can concentrate only on two view data. Long et al formulated a multiview spectral clustering method while investigating multiple spectral dimension reduction.

Zhou and Burges developed a multiview clustering strategy through generalizing the Ncut from a single view to multiple views and subsequently they build a multiview transductive inference. In tensor-based strategy, the multilinear relationship among multiview data is taken into account. The strategy focuses on the clustering of multitype interrelated data objects, rather than clustering of the same objects using multiple representations as in our research.

Spectral Clustering

Spectral clustering was derived based on relaxation of the Ncut formulation for clustering. Spectral clustering involves a matrix trace optimization problem [14]. In this paper, we proposed that the spectral clustering formalism can be extended to deal with multiview problems based on tensor computations.

Given a set of N data points $\{x_i\}$ where $x_i \varepsilon IR^d$ is the *i* th data point, a similarity s_{ij} can be defined for each pair of data points x_i and x_j based on some similarity measure. An intuitive way for representing the data set by using a graph G= (V, E) in which the vertices *V* represents the data points and the edges characterize the similarity between data points which are quantified by s_{ij} the similarity measure of the graph is symmetric and undirected. The matrix of the graph G is the matrix S with entry in row i and column j equal to s_{ij}.The degree of the vertex can be written as

$$d_i = \sum_{j=1}^{N} s_{ij} \qquad (1)$$

Where v_i is connected to the sum of all weight of the edges. The degree of the matrix D is a diagonal matrix containing the vertex degrees from $d_1 \ldots . d_N$ As the diagonal, It follows from the spectral formalism of embedding the Laplacian matrix can be defined as L=D-S and Ncut is defined by corresponding to the normalized Laplacian matrix

$$L_{Ncut} = D^{-1/2} L D^{-1/2} = \text{I-}S_N \qquad (2)$$

Where s_N and L_{Ncut} are the normalized similarity eigenvector and their eigenvalues.

Multiview Spectral Clustering

In integration of multiview data in spectral clustering, there are two different strategies

Multiview Clustering by Trace Maximization (MC-TR-I)

Different views can be added to the objective function. The function can be as follows,

$$\max_{U} \sum_{K=1}^{K} trace\left(U^T S_N^{(K)} U\right). = \left(U^T \left(\sum_{k=1}^{K} S_N^{(K)}\right) U\right), \qquad (3)$$

$$U^T \text{U=I},$$

Where $S_N^{(K)}$ a normalized matrix for kth value and U is the common factor shared by the views. This corresponds to the MKF with linear kernel. In alternative, weighted combination of objective functions, where the weights are learned from the data[6],[7]

$$= \max_{U,W} \sum_{K=1}^{K} W_K trace\left(U^T S_N^{(K)} U\right)$$

$$= \max_{U,W} \text{trace}\left(U^T \left(\sum_{K=1}^{K} W_K S_N^{(K)}\right) U\right), \qquad (4)$$

$$U^T \text{U = I, W} \geq 0 \text{ and } \|W\|_F = 1.$$

Modularity for Multidimensional network

The modules are used to identify the communities among the multiple network analysis. The d dimensional analysis can be represented as

$$A = \{A_1, A_2, \ldots \ldots . A_d\}$$

In multidimensional analysis, the symmetric networks can be concatenated by representing the interaction of actors in various forms. This is implemented to concatenate the communities among them.

Average Modularity Maximization (AMM)

This is used to handle single dimensional network [2] .The average interaction network can be calculated as

$$\bar{A} = \frac{1}{d}\sum_{i=1}^{d} A_i \qquad (5)$$

Correspondingly,

$$\bar{m} = \frac{1}{d}\sum_{i=1}^{d} m_i \;,\; \bar{d} = \frac{1}{d}\sum_{i=1}^{d} d_i \qquad (6)$$

Where, $\bar{A}$ is used to find the community detection in single dimensional network. For multiview network, it can be representated as

$$\text{Max Q} = \frac{1}{2m}\text{Tr}\left(S^T\left[\bar{A} - \frac{\bar{d}\,\bar{d}^T}{2\bar{m}}\right] S\right) \qquad (7)$$

Total Modularity Maximization (TMM)

This is used to maximize the modularity among the all the dimensions. It can be represented as,

$$\text{Max } \bar{Q} = \frac{1}{d}\sum_{i=1}^{d} Q_i = \frac{1}{d}\sum_{i=1}^{d} Tr\left(S^T \frac{B_i}{2m_i} S\right)$$

$$= \text{Tr}\left(S^T\left[\frac{1}{d}\sum_{i=1}^{d}\left\{\frac{A_i}{2m_i} - \frac{d\,d^T}{(2mi)^2}\right\}\right] S\right) \qquad (8)$$

Where Q_i is the modularity in i-th dimension. To compute the top Eigen vectors, matrix vector multiplication can be calculated as:

$$\frac{1}{d}\left[\sum_{i=1}^{d}\frac{A_i}{2m_i}x - \sum_{i=1}^{d}\frac{d_i^T x}{(2m_i)^2} d_i\right] \qquad (9)$$

Here degree dimension is expressed whereas in average modularity, degree distributions are not mentioned.

Principal Modularity Maximization by using Spectral Clustering

In AMM and TMM, single views are extracted and they are integrated. In PMM, it integrates the multiple dimensions of the network. It consists of two steps, they are

i) Feature Extraction of Multiview dimension.

ii)Integration of Cross dimension

Feature Extraction of multivew dimension

In Feature Extraction. Spectral clustering is used for calculating similarity matrix from mutiview objects. After calculating similarity matrix, then we compute top Eigen vectors of the modularity matrix. The Eigen vectors represent the possible community partitions. When it is compared with AMM and TMM method, noise is reduced. Eigen vectors can be either positive or negative. The Eigen vectors with positive values alone selected.

Integration of Cross Dimension

After extracting similar community partitions, the features of the multiview dimensions are extracted. They are extracted based on the modularity maximization. Let *S* be the top *l* Eigen vectors that maximize the *Q* and *V* an orthogonal matrix such that,

$$V \varepsilon R^{l \times l}, VV^T = I_l, V^T V = I_l$$

SV which also maximize Q:

$$= \frac{1}{2m}\text{tr}\left(S^T BS\right) = Q_{max} \qquad (10)$$

Where *SV* and *S are* equivalent in an orthogonal transformation. Correlation between the multiple sets of variables can be founded using Canonical Correlation Analysis. The correlation between two dimensions can be expressed as

$$\sum_{i=1}^{d} w_i^T C_{ii} w_i = 1 \qquad (11)$$

By using the Lagrange multiplier, the derivatives, we obtain the equation as

Algorithm: Principal Modularity Maximization by Spectral Clustering

Input: Net = $\{A_1, A_2, \ldots, A_d.\}$

Number of communities of communities *K*

Number of structure features to extract *l*

Output: Community is assigned to *idx*

Step 1: Compute top *l* Eigen vectors of modularity matrix
For each ai via Lancoz method;
Step 2: Select the vectors with positive Eigen values as *Si;*
Step 3: Compute slim SVD of X= [$S_1, S_2 ...... S_n$] = UDV^T;
Step 4: Obtain the lower dimensional embedding U=U (;, k-1);
Step 5: Normalize the row of $\overline{U}$ to unit length;
Step 6: Calculate the cluster *idx* to unit length;

The top Eigen vectors of the modularity are extracted from the structural features. The diaog $(C_{11,} C_{22 .....} C_{dd}\ )$ matrix becomes identity matrix. Hence W ($w_{1,} w_2 \ldots\ldots w_3$) corresponds to the top Eigen vector of full covariance matrix.

X= ($S_{1,} S_2 \ldots\ldots S_d$)

EXPERIMENTAL EVALUATION

We evaluated and compared with different strategies which applied to multi dimensional networks. The synthetic dataset which consists of three clusters having 50,100 and 200 members in them. The interaction probability is calculated for each group member. Based on the dimensions, interaction probability differs. When we add one member of the group to the other group with low probability. Normalized Mutual Information (NMI) is used to measure the performance. NMI value is 1 when two clusters are exactly same. Normally NMI values exist between 0 and 1.

Our principle Modularity Maximization method is used to cover the hidden patterns and it also performs well, when compared with other two methods. For example if we include some noises in the second dimension, the performance is reduced from 0.5 to 0.1. For this dimension, the pattern cannot be analyzed. By using PMM based method, it achieves good performance. This method is more robust in noisy dimensions.

Table 1

	Sample cluster	Performance
Single Dimensional Network	A_1 A_2 A_3 A_4	0.7337 ± 0.1934 0.6898 ± 0.1898 0.6772 ± 0.1858 0.6907 ± 0.1986
Multi Dimensional Network	AMM TMM PMM	0.7956 ± 0.1633 0.9167 ± 0.1147 0.9361 ± 0.1069

Performance on Synthetic dataset with PMM

Here A_1, A_2, A_3, A_4 are sample single dimensional community detection when compared with the multidimensional community. In multidimensional clustering, the performance of PMM is higher when compared with the other maximization techniques. Single dimensions have the higher variance when compared with all multidimensional community.

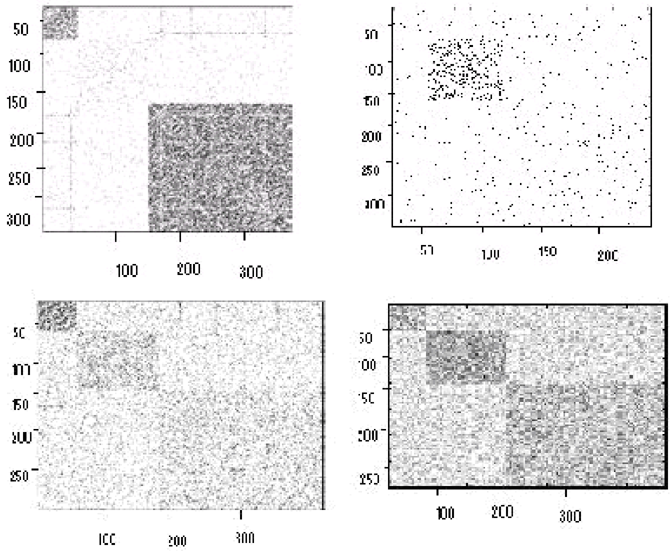

Fig 1. Example of Multi dimensional Network.

In fig 1. First dimension consists of two clusters and second dimension consists of one cluster. The first two dimensions are based on single view dimensional with low dimensionality. Therefore, the hidden patterns cannot be viewed properly. They are indicated with the low NMI. When comparing with the single view clustering, hidden patterns can be shown clearly. The NMI value exists between 0 and 1.

We first evaluate and compare the different clustering strategies and applied to the multiview clustering. Here clustering can be formed by the extension of the spectral clustering. Here dataset consist of three communities which consist of 50,100 and 200 members.

We can generate different view of interaction that is in each view; network shares the same vertices but has a different interaction pattern. The group members within the group can interact with the others random manner. The probability of interaction differs with respect to distinct views. Two vertices are connected randomly in low probability by adding some noise. The different views demonstrate the different interaction pattern in them.

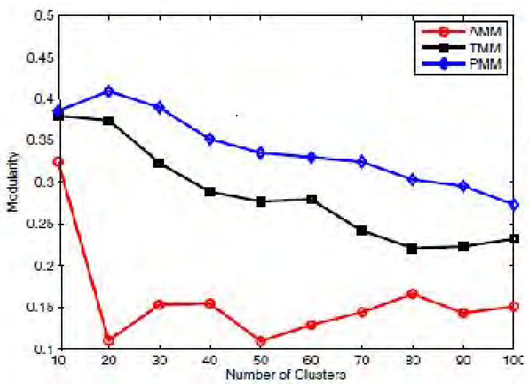

Fig 2. Performance of cluster in dimensions

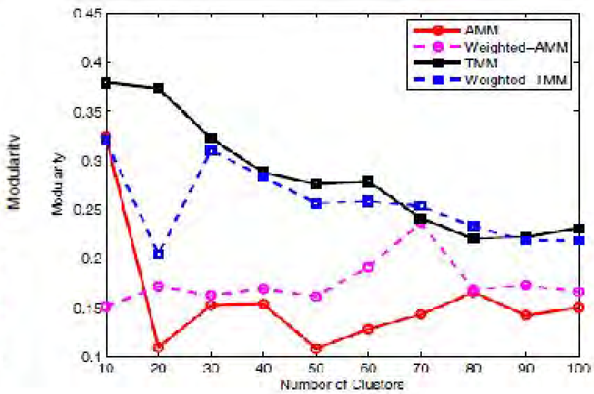

Fig 3.Performance of weighted AMM &TMM

In fig 2.The clustering evaluation for single view and multiview clustering is analyzed. In multiview clustering hidden patterns can be clearly viewed when compared to the single view clustering. Therefore, most of the multiview clustering results are better than the single view clustering. Multiview clustering helps to reduce the noise and shows the shared cluster. In fig 3. The performance of TMM and AMM are compared with each other. The weight of AMM and TMM are calculated in which TMM performance is high.

RESULT AND DISCUSSION

In multidimensional networks, hidden pattern can be viewed clearly when compared to the single dimensional networks. In Average Modularity Maximization (AMM) and Total Modularity Maximization (TMM) methods are not robust to handle. In Principal Modularity Maximization (PMM), the hidden patterns are clearly shown. Here we have proposed the extension of spectral clustering by implementing the maximization algorithm. By using this algorithm, multidimensional networks can be included in heterogeneous environment. Tensor based solutions are proposed for including multiple networks in the form of tensor.

Multiview clustering performance can be analyzed using many tensor based strategies. Multiple similarities can be founded through many methods. In cluster ensemble method, single view partition clusters can be integrated. So, this is not efficient.

In LMF method, clustering performance get lower at initialization and the partition cluster at the end are unstable. The optimization results also consumes much time.

In tensor based multiview clustering, spectral clustering can be extended by Principal Modularity Maximization where multiple networks can be linked together. This method is more efficient when compared with other methods. NMI is used to measure the performance of the clustering.

REFERENCES

[1] Xinhai Liu, Shuiwang Ji, Wolfgang Glanzel, and Bart De moor, Fellow, “Multiview clustering via Tensor Methods” IEEE Trans. Knowledge and Data Engineering, vol.25, no. 5, May 2013.

[2] Lei Tang, Xufei Wang, Huan Liu, “Uncovering Groups via Heterogeneous Interaction Analysis”,

[3] H.G. Ayad and M.S. Kamel, “Cumulative Voting Consensus Method for Partitions with Variable Number of Clusters,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 30, no. 1, pp. 160-173, Jan. 2008.

[4] Teresa M. Selee,, Tamara G. Kolda, W. Philip Kegelmeyer, And Joshuda. Griffin,” Extracting Clusters from Large Datasets with Multiple Similarity Measures Using IMSCAND”, summer proceedings, 2007

[5] S. Bickel and T. Scheffer,” Multi-View Clustering”, Proc. IEEE Fourth Int’l Conf. Data Mining (ICDM ’04), pp. 19-26, 2004.

[6] J.D. Carroll and J.J. Chang,” Analysis of Individual Differences in Multidimensional Scaling via an n-Way Generalization of Echart-Young, Decomposition”, Psychometricka, vol. 35, pp. 283-319, 1970.

[7] K. Chaudhuri, S.M. Kakade, K. Livescu, and K. Sridharan,” Multi-View Clustering Via Canonical Correlation Analysis”, Proc. 26th Ann. Int’l Conf. Machine Learning (ICML ’09), pp. 129-136, 2009.

[8] A. Cichocki, R. Zdunek, A.-H. Phan, and S. Amari,” Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation”. John Wiley, 2009.

[9] L. De Lathauwer, B.D. Moor, and J. Vandewalle, “A Multilinear Singular Value Decomposition, SIAM J. Matrix Analysis and Applications”, vol. 21, no. 4, pp. 1253-1278, 2000.

[10] L. De Lathauwer, B.D. Moor, and J. Vandewalle, “On the Best Rank-1 and Rank Approximation of Higher-Order Tensors, SIAM J. Matrix Analysis and Applications”, vol. 21, no. 4, pp. 1324-1342, 2000.

[11] D.M. Dunlavy, T.G. Kolda, and E. Acar, Poblano v1.0:” A MATLAB Toolbox for Gradient-Based Optimization”, Technical Report SAND2010-1422, Sandia Nat’l Laboratories, Mar. 2010.

[12] D.M. Dunlavy, T.G. Kolda, and W.P. Kegelmeyer,” Multilinear Algebra for Analyzing Data with Multiple Linkages,” Technical Report SAND2006-2079, Sandia Nat’l Laboratories, 2006.

[13] H. Huang, C. Ding, D. Luo , and T. Li, “Simultaneous Tensor Subspace Selection and Clustering: The Equivalence of High Order SVD and k-Means Clustering,” Proc. 14th ACM SIGKDDInt’l Conf. Knowledge Discovery and Data Mining, pp. 327-335, 2008.

[14] B. Long, Z.M. Zhang, X. Wu´ , and P.S. Yu, “Spectral Clustering for Multi-Type Relational Data,” Proc. 23rd Int’l Conf. Machine Learning, pp. 585-592, 2006.

[15] U. Luxburg, “A Tutorial on Spectral Clustering,” Statistics and Computing, vol. 17, no. 4, pp. 395-416, 2007