IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2016

An Efficient Locally Weighted Spectral Cluster for Automatic Image Segmentation

Vishnu Priya M [1], J Santhosh[2]

M.Phil Scholar, Department of Computer Science, Sree Narayana Guru College, Coimbatore, India [1]

Assistant Professor, Department of Computer Science, Sree Narayana Guru College, Coimbatore, India [2]

Abstract: Spectral clustering is one of the most important algorithms in data mining and machine intelligence; however, its computational complexity limits its application to truly large scale data analysis. However, ensample clustering suffers from a scalability problem in both memory use and computational time when the size of a data set is large. The proposed system introduced are three: (a) The local scale, rather than a global one, (b) estimating the scale value of the data, and (c) weighted based eigenvectors value rotating to create the maximally sparse representation. The proposed an automated spectral clustering algorithm based on these ideas: it computes automatically the large data and the number of groups and it can handle multi-scale data which are difficult for previous ensample approaches. Experimental results on multiple real-world image datasets demonstrate the effectiveness and efficiency of our approach. In particular, given a cluster, investigate its uncertainty by considering how the objects inside this cluster are grouped in the multiple base clustering's. Based on cluster uncertainty estimation, a spectral cluster index (SCI) is then presented to measure the reliability of clusters. The proposed algorithm the crowd of diverse clusters in the spectral can provide an effective indication for evaluating each individual cluster in the subspace. By evaluating and weighting the clusters in the spectral via the SCI measure, the present the concept of locally weighted co-association matrix, which incorporates local adaptively into the conventional co-association matrix and serves as a summary for the spectral of diverse clusters. Finally, to achieve the final clustering result, propose novel locally weighted Ng-Jordan-Weiss (WNJW) Algorithm, respectively, with the diversity of clusters exploited and the local weighting strategy incorporated.

Keywords: Locally Weighted Spectral Cluster, matrix, Local Scaling, Estimating Weight based Clusters.

I. INTRODUCTION

The spectral methods recently emerge as effective methods for data clustering, image segmentation, Web ranking analysis and dimension reduction. They start with well-motivated objective functions; optimization eventually leads to eigenvectors, with many clear and interesting algebraic properties. At the core of spectral clustering is the Laplacian of the graph adjacency (pairwise similarity) matrix, evolved from spectral graph partitioning. This tutorial provides a survey of recent advances after brief historical developments.

Mathematical proofs will be outlined and examples in gene expressions and internet newsgroups will give to illustrate the ideas and results. Traditional clustering algorithms, such as k-means, GM EM, etc, while simple, most of them are based on convex spherical sample space, and their ability for dealing with complex cluster structure is poor. When the sample space is not convex, these algorithms may be trapped in a local optimum. The spectral clustering algorithm has been proposed to solve this issue. Spectral clustering algorithm is based on spectra graph theory that partition data using eigenvectors of an affinity matrix derived from the data. It can cluster arbitrarily shaped data. In recent years, spectral clustering has been successfully applied to a large number of challenging clustering applications. It is simple to implement, can be solved efficiently by standard linear algebra software, and often outperforms traditional clustering algorithms such as the k-means algorithm.

Clustering is one of the building blocks of modern data analysis. Two commonly used methods are K-means and learning a mixture-model using EM. These methods, which are based on estimating explicit models of the data, provide high quality results when the data is organized according to the assumed models. However, when it is arranged in more complex and unknown shapes, these methods tend to fail. An alternative clustering approach, which was shown to handle such structured data is spectral clustering. It does not require estimating an explicit model of data distribution, rather a spectral analysis of the matrix of point-to-point similarities.

Although spectral clustering algorithms have shown good results in various applications, it relies on the dataset where each cluster is approximately well separated to a certain extent. The spectral clustering algorithm will fail to recognize one cluster as different clusters when the cluster has an obvious inflection point within a non-convex space. The reason is that the constructed affinity matrix, which the spectral clustering heavily relies on, will be corrupted with poor pairwise affinity values from the area of inflection points. Especially for most of the recent spectral clustering algorithms that use the traditional central grouping techniques to cluster the affinity matrix, e.g., k-

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2016

means, it will amplify the misguidance of clustering because these centralized algorithms that are based on a radius distance between two data points cannot separate clusters that are very long or nonlinearly separable.

Weighted Spectral Cluster approach based on cluster uncertainty estimation and local weighting strategy. In particular, the uncertainty of each cluster is estimated by considering the cluster labels in the entire ensemble via an entropic criterion. A novel spectral-driven cluster validity measure is introduced, and a locally weighted co-association matrix is presented to serve as a summary for the ensemble of diverse clusters. With the local diversity in spectral exploited, two novel consensus functions are further proposed. Extensive experiments on a variety of real-world datasets demonstrate the superiority of the proposed approach over the state-of-the-art. Here use a number of higher-dimensional synthetic datasets to show that the weight based spectral clustering procedure can eliminate the misleading information from different kinds of datasets so the obtained spectral method could cluster the dataset more accurately.

II. RELATED WORK

The pair-wise similarity based approaches represent the ensemble information by some pair-wise similarity measure. The evidence accumulation clustering (EAC) proposed is probably the best known pair-wise similarity based approach. In EAC, a co-association matrix is constructed by counting how many times two objects occur in the same cluster in the ensemble of multiple base clustering's. By treating the CA matrix as a new similarity matrix, clustering algorithms, such as the agglomerative clustering methods, can be further utilized to obtain the consensus clustering. Proposed a hierarchical clustering algorithm to construct the consensus clustering using the CA matrix. The concept of normalized edges is introduced to measure the similarity between clusters.

The median partition based approaches aim to find a clustering (or partition) that maximizes the similarity between this clustering and all of the base clustering's, which can be viewed as finding the median point of the base clustering's. Due to the huge space of all possible clustering's, it is generally infeasible to find the optimal solution for the median partition problem. In fact, the median partition problem is NP-complete. The existing algorithm to find an approximatively solution for the ensemble clustering problem by exploiting the genetic algorithm, in which clustering's are represented as chromosomes. To formulate the median partition problem into a maximum likelihood problem and solved it by the EM algorithm.

The graph partitioning based approaches are another main category of ensemble clustering. The exiting method formulated the ensemble clustering problem into a graph partitioning problem and proposed three ensemble clustering approaches: cluster-based similarity partitioning algorithm (CSPA), hypergraph partitioning algorithm (HGPA), and meta-clustering algorithm (MCLA). The clustering ensemble into a bipartite graph by treating both clusters and objects as graph nodes and obtained the consensus clustering by partitioning the bipartite graph. Ren et al. [8] proposed to assign weights to data objects with regard to how difficult it is to cluster them and presented three graph partitioning algorithms based on the weighted object scheme, that is, weighted-object meta clustering (WOMC), weighted-object similarity partition (WOSP) clustering, and weighted-object hybrid bipartite (WOHB) graph partition clustering. Despite the significant success, there are still two limitations to most of the existing ensemble clustering approaches. First, the existing approaches mostly overlook the problem of uncertain links which may mislead the consensus process. Second, most of them lack the ability to incorporate global structure information to refine local links accurately and efficiently. The existing approach addresses the issue of uncertain links in an effective and efficient manner. To identify the uncertain links by the ENS strategy and build a sparse graph with a small number of probably reliable links. The empirical study shows the advantage of using only a small number of probably reliable links rather than all graph links regardless of their reliability. The incorporate global information to construct more accurate local links by exploiting the random walk trajectories. The random walkers driven by a new probability transition matrix are utilized to explore the graph structure. A dense similarity measure is further derived from the sparse graph K-ENG using probability trajectories of the random walkers. uncertain links in a locally adaptive manner and construct a sparse graph with a small number of probably reliable links. It has been shown that the use of a small number of probably reliable links can lead to significantly better clustering's than using all graph links regardless of their reliability.

Researchers have investigated how to combine clustering ensembles with subspace clustering in an effort to address both the ill-posed nature of clustering and the curse-of dimensionality in high dimensional spaces. A subspace clustering is a collection of weighted clusters, where each cluster has a weight vector representing the relevance of features for that cluster. In a subspace clustering ensemble, the consensus function makes use of both the clusters and the weight vectors provided by the base subspace clustering. Work has also been done to evaluate the relevance of each base clustering (and assign weights accordingly), in an effort to improve the final consensus clustering. To the best of our knowledge, no previous work has investigated how to use weights associated to objects within the clustering ensemble framework.

Temporal data clustering provides underpinning techniques for discovering the intrinsic structure and condensing information over temporal data. In this paper, we present a temporal data clustering framework via a weighted clustering ensemble of multiple partitions produced by initial clustering analysis on different

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2016

temporal data representations. In our approach, we propose a weighted consensus function guided by clustering validation criteria to reconcile initial partitions to candidate consensus partitions from different perspectives and then introduce an agreement function to further reconcile those candidate consensus partitions to a final partition. As a result, the exiting weighted clustering ensemble algorithm provides an effective enabling technique for the joint use of different representations, which cuts the information loss in a single representation and exploits various information sources underlying temporal data. In addition, our approach tends to capture the intrinsic structure of a data set, e.g., the number of clusters.

Laplacian matrix and calls a sparse Eigen solver. Several methods are available for scarifying the similarity matrix. A sparse representation effectively handles the memory bottleneck, but some scarification schemes still require calculating all elements of the similarity matrix. Another popular approach to speed up spectral clustering is by using a dense sub-matrix of the similarity matrix. In particular, The Nyström approximation to avoid calculating the whole similarity matrix; this approach trades accurate similarity values for shortened computational time. In another work, propose a method that does not use eigenvectors, but they assume the availability of the similarity matrix. The existing work, the developing a parallel spectral clustering package on distributed environments. We begin by analysing 1) the traditional method of scarifying the similarity matrix and 2) the Nystrom approximation. While the scarification approach may be more computationally expensive, our experimental results indicate that it may yield a slightly better solution.

III.PROPOSED APPROACH

Given a dataset which contains n data points x={X = $\{x_1, x_2 \ldots . x_n\}$,$x_i \in R^l$. Construct an undirected weighted graph G = (V,E,W). Treat each data point as a vertex V in graph G. Each edge E between vertices (xi ,x j) has a similarity value W. Then the clustering problem can be transformed into a graph partitioning problem on graph G. The optimal partitioning criteria based on graph theory is maximizing the internal similarity of the two divided sub-graphs, and minimizing the similarity between sub-graphs. Most spectral clustering algorithms search clusters utilizing the eigenvectors of similarity matrix. The similarity matrix W ∈ Rn×n of spectral clustering algorithm is composed of wij, which is usually represented by Gaussian kernel function:

$$w_{ij} = \exp\left(-\frac{d^2(x_i, x_j)}{2\sigma^2}\right) \quad (1)$$

Where $d(x_i, x_j)$the Euclidean distance between point x_i and x_j ; σ is the scaling parameter which controls how rapidly the similarity w_{ij} falls off with $d(x_i, x_j)$.

In graph G , the sum of the weights of the edges connected to vertex i , is defined as the degree of vertex i , which can be represented by d_i:

$$d_i = \sum_{j=1}^{n} w_{ij} \quad (2)$$

In graph cut methods, seeking the optimal solution of the objective function is often NP-hard. With the help of spectral method, the original problem can be solved in polynomial time by relaxing the original discrete optimization problem to the real domain. For graph partitioning, a point can be considered part belonging to subset A and part belonging to subset B, rather than strictly belongs to one cluster. It can be proved that the classification information of vertices is contained in the eigenvalues and eigenvectors of graph Laplacian matrix. and it can get good clustering results.

Step 1 Create similarity matrix and Laplacian matrix to describe the samples;
Step 2 Calculate the eigenvalues of the Laplacian matrix, and choose appropriate eigenvalues and their corresponding eigenvectors to create the space Rk ;
1) 2-way: the initial samples are distributed into one-dimensional space (k=1).
2) k-way: the initial samples are distributed into space Rk and we can obtain matrix Y which consists of k orthogonal vectors.
Step 3 Treat Y as a new representation space of data samples, and then cluster the data samples according to the new representation space.
1) 2-way: optimize the objective function in the one-dimensional space, and then repeatedly partition the obtained sub-graphs.
2) k-way: use a typical clustering algorithm such as kmeans to deal with the eigenvectors in Rk .

A. Local Scaling

The scaling parameter is some measure of when two points are considered similar. This provides an intuitive way for selecting possible values for σ. The selection of σ is commonly done manually. The selecting σ automatically by running their clustering algorithm repeatedly for a number of values of σ and selecting the one which provides least distorted clusters of the rows of Y. This increases significantly the computation time. Additionally, the range of values to be tested still has to be set manually. Moreover, when the input data includes clusters with different local statistics there may not be a single value of σ that works well for all the data. When the data contains multiple scales, even using the optimal σ fails to provide good clustering.

Instead of selecting a single scaling parameter σ propose to calculate a local scaling parameter σi for each data point si. The distance from si to sj as 'seen' by si is d(si, sj)/σi while the converse is d(sj, si)/σj . Therefore the square

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2016

distance d2 of the earlier papers may be generalized as d(si, sj)d(sj, si)/σiσj = d2(si, sj)/σiσj The affinity between a pair of points can thus be written as:

$$\widehat{A_{ij}} = \exp\left(\frac{-d^2(x_i, y_i)}{\sigma_i \sigma_j}\right) \quad (3)$$

Using a specific scaling parameter for each point allows self-tuning of the point-to-point distances according to the local statistics of the neighbourhoods surrounding points i and j. where sK is the K'th neighbor of point si. The selection of K is independent of scale and is a function of the data dimension of the embedding space. Nevertheless, in all the experiments (both on synthetic data and on images) we used a single value of K = 7, which gave good results even for high-dimensional data (the experiments with high-dimensional data were left out due to lack of space).

B. Estimating Weight based Clusters

Having defined a scheme to set the scale parameter automatically it is left with one more free parameter: the number of clusters. This parameter is usually set manually and not much research has been done as to how might one set it automatically. It can approach to discovering the number of clusters. The suggested scheme turns out to lead to a new spatial clustering algorithm. Examining the eigenvalues of our locally scaled matrix, corresponding to clean data-sets, indeed shows that the multiplicity of eigenvalue 1 equals the number of groups. However, if the groups are not clearly separated, once noise is introduced, the values start to deviate from 1, thus the criterion of choice becomes tricky.

Following this objective, propose a weight assignment to objects using the results of all base clustering's, and then embed the weights into the successive consensus clustering process. Similar to boosting, points that are hard to cluster receive larger weights, while easy-to-cluster points are given smaller weights. The difference is that boosting is an iterative process, while our weight assignment scheme is performed in one-shot. The details are given below. the level of uncertainty in clustering two points **x***i* and **x***j* as

$$\text{confusion}(x_i, x_j) = \widehat{A_{ij}}\left(1 - \widehat{A_{ij}}\right) \quad (4)$$

The confusion index reaches its maximum of 0.25 when $\widehat{A_{ij}}$ = 0.5, and its minimum of 0 when $\widehat{A_{ij}}$ = 0 or Aij = 1. Here uses this confusion measure to define the weight associated with each object as follows:

$$w_i' = \frac{4}{n}\sum_{j=1}^{n} \text{confusion}(x_i, x_j) \quad (5)$$

The normalization factor 4 n is used to guarantee that w' i ∈ [0, 1].

To avoid a value of 0 for a weight, which can lead to instability, add a smoothing term

$$w_i = \frac{w_i' + e}{1 + e} \quad (6)$$

where e is a small positive number (e = 0.01 in the experiments). As a result, wi ∈ (0, 1]. A large wi value means that confusion (**x***i*, **x***j*) is large for different **x***j* values. As such it

C. Weighted Spectral Cluster Algorithm

The proposed method for estimating the number of groups automatically has two desirable by-products: (i) After aligning with the canonical coordinate system, one can use non-maximum suppression on the rows of Z, thus eliminating the final iterative k-means process, which often requires around 100 iterations and depends highly on its initialization. (ii) Since the final clustering can be conducted by non-maximum suppression, here obtain clustering results for all the inspected group numbers at a tiny additional cost. When the data is highly noisy, one can still employ k-means, or better, EM, to cluster the rows of Z.

Algorithm:

Given a set of points S = {s1, . . . , sn} in Rl that we want to cluster:

1. Compute the local scale σi for each point si ∈ S using Eq. (1).
2. Form the locally scaled affinity matrix Aˆ ∈ Rn×n where Aˆij is defined according to Eq. (1) for i = j and Aˆii = 0.
3. Define D to be a diagonal matrix and construct the normalized affinity matrix L = D−1/2AˆD−1/2. using Eq. (3)
4. Find x1, . . ., xC the C largest eigenvectors of L and form the matrix X =[x1, . . . , xC] ∈ Rn×C, where C is the largest possible group number.
5. Recover the rotation R which best aligns X's columns with the canonical coordinate system using the incremental gradient descent scheme using Eq. (4)
6. Grade the cost of the alignment for each group number, up to C, according to Eq. (5).
7. Set the final group number Cbest to be the largest group number with minimal alignment cost. using Eq. (6)
8. Take the alignment result Z of the top Cbest eigenvectors and assign the original point si to cluster c if and only if maxj(Z2 ij) = Z2 ic.
9. If highly noisy data, use the previous step result to initialize weight, clustering on the rows of Z.

The tested the quality of this algorithm on real data. The number of groups and the corresponding segmentation were obtained automatically. In this case same quality of results were obtained using non-scaled affinities, however, this required manual setting of both σ (different values for different images) and the number of groups, whereas our result required no parameter settings.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2016

IV. EXPERIMENT RESULTS

Weight based spectral Cluster proposed an automated spectral clustering algorithm based on these ideas: it computes automatically the scale and the number of groups and it can handle multi-scale data which are problematic for previous approaches. the normalized mutual information (NMI) to evaluate the quality of the consensus clustering's, which provides a sound indication of the shared information between two clustering's. Note that a higher NMI indicates a better test clustering.

TABLE I AVERAGE RATIOPL VARYING PARAMETERS

K	1	2	3
Ensemble Clustering	3.2	4.0	5.9
Weighted Spectral Cluster	6.8	7.1	8.0

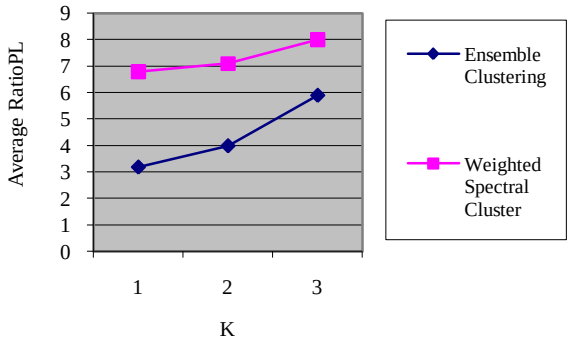

Fig 1 Compare the Avg RatioPL existing with proposed algorithm.

TABLE III AVERAGE PERFORMANCE IN TERMS OF NMI

	True-k		
K	1	2	3
Ensemble Clustering	0.513	0.411	0.376
Weighted Spectral Cluster	0.629	0.609	0.577

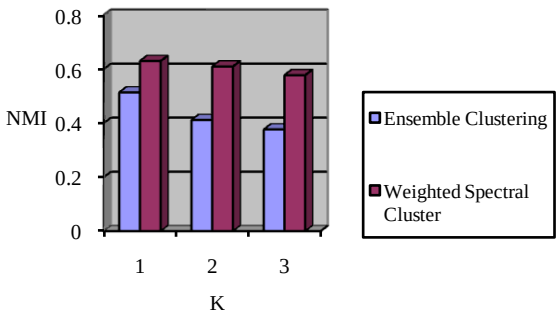

Fig 2 Compare the NMI score existing with proposed algorithm.

The existing PTA and PTGP metods and the baseline methods 100 times on each dataset. But in proposed system take minimum time to run the image dataset.

TABLE IIIII EXECUTION TIME OF DIFFERENT CLUSTERING APPROACHES

	Datasize		
K	200	400	600
Ensemble Clustering	8.687	12.856	16.983
Weighted Spectral Cluster	5.432	7.654	9.473

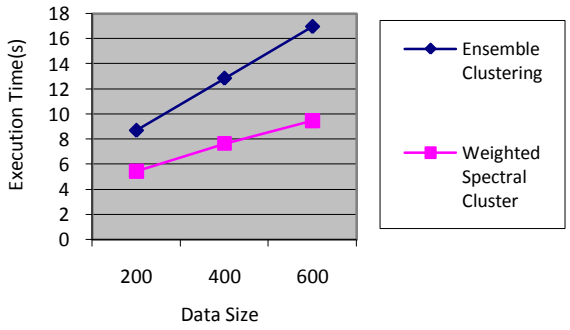

Fig. 3. Execution time of different clustering approaches as the data size varies.

Extensive experiments have been conducted on ten real-world datasets. The experimental results show that our approach significantly outperforms the state-of-the-art approaches in both clustering accuracy and efficiency.

V. CONCLUSION

The proposed an automated spectral clustering algorithm based on these ideas: it computes automatically the large data and the number of groups and it can handle multi-scale data which are difficult for previous ensample approaches. Experimental results on multiple real-world image datasets demonstrate the effectiveness and efficiency of our approach. In particular, given a cluster, investigate its uncertainty by considering how the objects inside this cluster are grouped in the multiple base clustering's. Based on cluster uncertainty estimation, a spectral cluster index (SCI) is then presented to measure the reliability of clusters. The proposed algorithm the crowd of diverse clusters in the spectral can provide an effective indication for evaluating each individual cluster in the subspace. By evaluating and weighting the clusters in the spectral via the SCI measure, the present the concept of locally weighted co-association matrix, which incorporates local adaptively into the conventional co-association matrix and serves as a summary for the spectral of diverse clusters. Finally, to achieve the final clustering result, propose novel locally weighted Ng-Jordan-Weiss (WNJW) Algorithm, respectively, with the

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2016

diversity of clusters exploited and the local weighting strategy incorporated.

REFERENCES

[1] A. L. N. Fred and A. K. Jain, "Combining multiple clusterings using evidence accumulation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 6, pp. 835–850, Jun. 2005.
[2] T. Li and C. Ding, "Weighted consensus clustering," in Proc. SDM Int. Conf. Data Mining, 2008, pp. 798–809.
[3] Y. Yang and K. Chen, "Temporal data clustering via weighted clustering ensemble with different representations," IEEE Trans. Knowl. Data Eng., vol. 23, no. 2, pp. 307–320, Feb. 2011.
[4] N. Iam-On, T. Boongoen, S. Garrett, and C. Price, "A link-based approach to the cluster ensemble problem," IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 12, pp. 2396–2409, Dec. 2011.
[5] Y. Zhang and T. Li, "Extending consensus clustering to explore multiple clustering views," in Proc. SIAM Int. Conf. Data Mining, 2011, pp. 920–931.
[6] N. Iam-On, T. Boongeon, S. Garrett, and C. Price, "A link-based cluster ensemble approach for categorical data clustering," IEEE Trans. Knowl. Data Eng., vol. 24, no. 3, pp. 413–425, Mar. 2012.
[7] J. Yi, T. Yang, R. Jin, and A. K. Jain, "Robust ensemble clustering by matrix completion," in Proc. Int. Conf. Data Mining, 2012, pp. 1176–1181.
[8] Y. Ren, C. Domeniconi, G. Zhang, and G. Yu, "Weighted-object ensemble clustering," in Proc. Int. Conf. Data Mining, 2013, pp. 627–636.
[9] A. Strehl and J. Ghosh, "Cluster ensembles: A knowledge reuse framework for combining multiple partitions," J. Mach. Learn. Res., vol. 3, pp. 583–617, 2003.
[10] X. Z. Fern and C. E. Brodley, "Solving cluster ensemble problems by bipartite graph partitioning," in Proc. 21st Int. Conf. Mach. Learn., 2004, pp. 36–43.
[11] S. Mimaroglu and E. Erdil, "Combining multiple clusterings using similarity graph," Pattern Recog., vol. 44, no. 3, pp. 694–703, 2011.
[12] D. Cristofor and D. Simovici, "Finding median partitions using information-theoretical-based genetic algorithms," J. Universal Comput. Sci., vol. 8, no. 2, pp. 153–172, 2002.
[13] A. Topchy, A. K. Jain, and W. Punch, "Clustering ensembles: Models of consensus and weak partitions," IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 12, pp. 1866–1881, Dec. 2005.
[14] X.Wang, C. Yang, and J. Zhou, "Clustering aggregation by probability accumulation," Pattern Recog., vol. 42, no. 5, pp. 668–675, 2009.
[15] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks," Phys. Rev. E, vol. 69, no. 2, 2004.
[16] P. Pons and M. Latapy, "Computing communities in large networks using random walks," in Proc. 20th Int. Symp. Comput. Inf. Sci., 2005, pp. 284–293.
[17] D. Lai, H. Lu, and C. Nardini, "Enhanced modularity-based community detection by random walk network preprocessing," Phys. Rev. E, vol. 81, 2010.
[18] C.-D. Wang, J.-H. Lai, and P. S. Yu, "NEIWalk: Community discovery in dynamic content-based networks," IEEE Trans. Knowl. Data Eng., vol. 26, no. 7, pp. 1734–1748, Jul. 2014.
[19] D. Huang, J.-H. Lai, and C.-D. Wang, "Combining multiple clusterings via crowd agreement estimation and multi-granularity link analysis," Neurocomputing, vol. 170, pp. 240–250, 2015.
[20] S. Vega-Pons and J. Ruiz-Shulcloper, "A survey of clustering ensemble algorithms," Int. J. Pattern Recog. Artif. Intell., vol. 25, no. 3, pp. 337–372, 2011.