A distributed approach to cluster multi-view relational data

Renê Pereira de Gusmão
Departamento de Computação
Universidade Federal de Sergipe
São Cristovão, Sergipe
rene@dcomp.ufs.br

Allan Juan Silva Araújo
Departamento de Computação
Universidade Federal de Sergipe
São Cristovão, Sergipe
ajsaraujo@dcomp.ufs.br

Francisco de Assis Tenório de Carvalho
Centro de Informática
Universidade Federal de Pernambuco
Recife, Pernambuco
fatc@cin.ufpe.br

***Abstract*—Clustering of multi-view data has become an important research field. The efficient clustering of multi-view data is a challenging problem. This work aimed to investigate a distributed approach to cluster multi-view relational data. A PSO-based hybrid method was used to generate clustering from all views independently. Five different objective functions were explored to induce diversity to the clusterings since each function looks for different cluster structures. Five different consensus functions were compared to produce the final partition from the ensembles. Three multi-view real-world data sets were considered in this study. The Adjusted Rand Index, the F-measure and Silhouette clustering validity indexes were used to assess obtained clusterings. The distributed approach found better clusterings for all data sets considering at least one consensus function.**

***Index Terms*—cluster analysis, relational multi-view data, distributed approach**

I. Introduction

Clustering methods divide data sets into clusters (groups) such that similar objects are placed in the same cluster and dissimilar objects are placed in different groups. Clustering algorithms have applications in several areas: Statistics, Biology, Pattern Recognition, Data mining [1].

Data can be described in two ways: vectorial or relational. In vectorial form, each object is described by a vector of quantitative or qualitative values. Usually, data sets are represented by a $n \times p$ matrix (n objects with p attributes). In relational form, each pair of objects is represented by a dissimilarity measure. Therefore, a $n \times n$ dissimilarity matrix describes the data set. Relational data is more practical when data has a high dimension. Another advantage of relational data is due to information privacy [2].

In several practical applications, data has multiple representations or sources that are complementary. This complementary information can be useful in the clustering process [3]. As stated by [4], the process of combining multiple views to achieve better performance is a significant research challenge.

Three main strategies are used to handle multi-view data clustering depending on whether the view combination is performed before, during, or after the clustering process: centralized approach, distributed approach, and concatenation approach. In a centralized approach, all views are considered simultaneously in clustering process. Relevance weights are estimated for each view such that some views may be more important than others for the clustering task. In concatenation approach, all views are concatenated into a single view and traditional clustering algorithms can be applied. In a distributed approach, the views are considered independently providing different clustering of the data set. Finally, a solution that represents a consensus among the set of clusterings is searched [5]. Most approaches focus on centralized strategy, while there is a lack of studies investigating the distributed approach.

Metaheuristic algorithms have proven to produce good solutions to the data clustering problem. Nature inspired metaheuristics gained attention due to its capability achieving a good balance in exploration-exploitation during the solution search. Particle Swarm Optimization (PSO) is one well-known nature inspired method and PSO-based algorithms were proposed to solve clustering problem for single view and multi-view relational data [6]–[8].

This work introduces a distributed approach to cluster multi-view relational data given by multiple dissimilarity matrices. The proposed approach uses a hybrid method based on Particle Swarm Optimization (PSO) to cluster all views independently first and then combine the generated clustering using a consensus function.

The remainder of this work is structured as follows. Section II presents a review of some related works. Section III presents an overview of basic related concepts. Section IV introduces the proposed distributed approach. Three multi-view data sets were used in experiments to assess the proposed approach according to two external clustering validity indexes, and their results are shown in Section V. Section VI presents the final remarks.

II. Related work

Most existing works in the literature studied the multi-view clustering problem in a centralized approach dealing with data sets described in vectorial form. Some works dealing with relational data will be discussed in this section.

The authors in [6] proposed a hard clustering algorithm based on PSO applied to single view relational data described by a single dissimilarity matrix. In this study, the proposed method was compared to the other three single-view algorithms for relational data. Five clustering validity indexes were considered to assess the quality of the generated clustering. In most cases, the proposed method found better clusterings.

The work [9] proposed a multi-view fuzzy clustering approach based on multiple medoids and minimax optimization called M4-FC for relational data. In M4-FC method, every object is considered as a cluster representative candidate with a weight. The weight represents the probability of the object to be chosen as a medoid. Minimax optimization is also applied to find consensus clustering results. M4-FC was tested on several multi-view data sets, in which M4-FC outperformed the compared approaches.

The authors in [8] proposed fuzzy clustering algorithms based on PSO for the clustering of multi-view relational data. The proposed methods use a centralized approach in which all views are used simultaneously in the clustering process. In this study, several clustering validity indexes were adapted to consider all views and weights, these adapted indexes were used as fitness functions. The proposed methods outperformed other algorithms suitable for multi-view relational data from the literature.

The study performed in [7] proposed hard clustering algorithms based on PSO for the clustering of multi-view relational data. The proposed methods use a centralized approach in which all views are used simultaneously in the clustering process. Eleven fitness functions were considered. It was observed that the top three fitness functions were Silhouette index, Xu index and Intra-cluster homogeneity. The proposed methods outperformed the compared methods for multi-view relational data from the literature.

III. Theoretical basis

Let $E = \{e_1, ..., e_n\}$ be a set of n objects and let T dissimilarity matrices $D_j = [d_j(e_i, e_l)]$, where $d_j(e_i, e_l)$ measures the dissimilarity between objects e_i and e_l in $j-$th view.

$$D_j = \begin{bmatrix} d_j(e_1, e_1) & \cdots & d_j(e_1, e_n) \\ \vdots & \ddots & \vdots \\ d_j(e_n, e_1) & \cdots & d_j(e_n, e_n) \end{bmatrix}$$

The PSO-hybrid method proposed in [6] will be used as cluster all views independently, in which each data set view is described by a dissimilarity matrix. In that method, each cluster is represented by a single medoid (object). The use of a single medoid may be insufficient to characterize the clusters. In this work, multiple medoids will be used to represent each cluster. The use of multiple medoids may represent each cluster more accurately [10]. In this work, we modify the method proposed in [6] to consider multiple medoids to represent each cluster.

Each particle p_i of the swarm is defined as a vector of representatives $(G_{i,1}, ..., G_{i,K})$, in which $G_{i,1}$ is a subset of E with fixed cardinality q ($|G_{i,1}| = q$). In a simplified way, each cluster is represented by q medoids. Each particle looks for a partition $\mathcal{P} = (C_1, ..., C_K)$ of E into K clusters and the corresponding representatives $G_1, \ldots, G_K$ which will represent the clusters in P such that the fitness function is optimized.

The best position of each particle, denoted as $pbest_i$, represents the best set of representatives $p_i^* = (G_1^*, .., G_K^*)$ found based on the value of fitness. The best position of the whole swarm denoted as $gbest$ corresponds to the best position $G_{best} = (G_1^*, .., G_K^*)$ found by the entire swarm. Each particle p_i have a velocity $V_i^{(t)} = (v_1^{(t)}, .., v_n^{(t)})$ at iteration t, which will be used to update the current position represented by the set of representatives. The PSO-based hybrid method is described in Alg 1.

Algorithm 1 PSO

INPUT:
- the dissimilarity matrix;
- the number K of clusters;
- the number of particles n_p in the swarm.
- the maximum number of iterations;
- suitable parameters c_1 and c_2;
OUTPUT:
- the vector of medoids $(G_1^*, .., G_K^*)$
- the partition $\mathcal{P} = (C_1, ..., C_K)$ of E into K clusters.
ALGORITHM
  For $i = 1$ to n_p
    Initialize p_i
    Set $pbest_i$ as current position
  Set $gbest$ as $(G_1, \ldots, G_K)$ of $p_i \in S$ with best fitness
REPEAT
  For each particle p_i in swarm S
    Update the velocity
    Update the position
    If $(particle\,improvement)$
      Update $pbest_i$
  If $(global\,improvement)$
    Update $gbest$
UNTIL stopping criterion is not satisfied

More details regarding position update step, velocity update step, and initialization of particles can be found in Ref. [6].

IV. Methodology

This section introduces the proposed distributed approach to cluster multi-view relational data. The distributed approach is divided into two steps. In the first step, each view of the p views is considered independently by the PSO-based hybrid method. Besides that, the clustering algorithm is executed independently for a given objective function. In this work, five objective functions were considered, and they are described in Subsection IV-A. These different objective functions were used to induce diversity since each function looks for different cluster structures that can complement each other. Therefore, the idea is to combine these different clustering results and potentially assure better data clusterings.

In the second step, all clustering results provided by the optimization of each view and each objective function are used as input for one consensus method. The consensus method will aggregate all clustering to form a final partition. In this work, we consider five different consensus functions (Subsection IV-B). Figure IV illustrates the distributed approach.

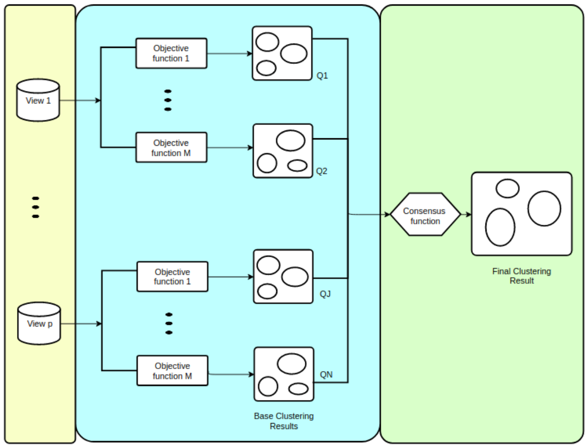

Fig. 1. Distributed approach for multi-view data

A. Ensemble generation

The Silhouette [11], Intra-cluster homogeneity, Davies-Bouldin [12], Dunn [13] and CS [14] clustering validity indexes were used as fitness functions. These different objective functions were used to induce diversity to the clusterings since each function looks for different cluster structures. It's important to emphasize that these clustering validity indexes were adapted to consider the multiple medoids as cluster representatives. The intra-cluster homogeneity is defined in Eq. 1. Lower values of this index indicate better clusterings.

$$HM = \sum_{k=1}^{K} \sum_{e_l \in C_k} \sum_{e \in G_k} d(e_l, e) \tag{1}$$

The silhouette is defined in Eq. 2. This index considers the homogeneity in each cluster and the separation between the clusters. The index is defined in [-1,1], higher values indicate better clusterings.

$$SIL = \frac{1}{n} \sum_{e_l \in E} s(e_l) \tag{2}$$

$$s(e_l) = \begin{cases} b(e_l) - a(e_l)/max(b(e_l), a(e_l)) & if \quad a(e_l) \neq b(e_l) \\ 0 & if \quad a(e_i) = b(e_i) \end{cases} \tag{3}$$

$$b(e_l) = \min_{C_j \neq C_k} d(e_l, C_j) \tag{4}$$

$$a(e_l) = \frac{\sum_{e \in C_k} d(e_l, e)}{|C_k|} \tag{5}$$

$$d(e_l, C_t) = \frac{\sum_{e \in C_t} d(e_l, e)}{|C_t|} \tag{6}$$

The Davies-Bouldin index also aims to minimize the intra-cluster distance and maximize the inter-cluster distances. To compute this index, the similarity and dispersion of clusters need to be computed. The index is defined in Eq. 7.

$$DB = \frac{1}{K} \sum_{k=1}^{K} max_{m \in \{1,..,K\}} \left(\frac{o_k + o_m}{M_{km}} \right) \tag{7}$$

$$o_k = \left(\frac{1}{|C_k|} \sum_{e_l \in C_k} \sum_{e \in G_k} d(e_l, e)) \right)^{\frac{1}{q}} \tag{8}$$

$$M_{km} = \max_{e_l \in G_k, e_t \in G_m} d(e_l, e_t) \tag{9}$$

The Dunn index defines the ratio between the minimum distance inter-cluster and the maximum diameter of clusters. The index is defined in Eq. 10 and should be maximized.

$$DN = \frac{min_{k \in \{1,..,K\}, k \neq i} \delta(C_i, C_k)}{max_{k \in \{1,..,K\}} \Delta(C_k)} \tag{10}$$

$$\delta(C_i, C_k) = \min_{e_m \in G_i, e_t \in G_k} d(e_m, e_t) \tag{11}$$

$$\Delta(C_k) = \max_{e_m, e_t \in C_k} d(e_m, e_t) \tag{12}$$

The CS index is a combination of the diameter of clusters and distances to the representatives of each cluster. This index was designed to deal with clusters of different sizes and shapes. The index is defined in Eq. 13 and should be minimized.

$$CS = \frac{\sum_{k=1}^{K} (\frac{1}{|C_k|} \sum_{e_j \in C_k} max_{e_l \in C_k} d(e_j, e_l))}{\sum_{i=1}^{K} (min_{j \in \{1,..,K\}, j \neq i} D(G_i, G_j))} \tag{13}$$

$$D(G_i, G_k) = \min_{e_m \in G_i, e_t \in G_k} d(e_m, e_t) \tag{14}$$

B. Consensus function

The second step of the distributed approach aims to obtain a final partition after obtaining the clustering results exploring all views under different objective functions. The consensus functions explored in this work are: Clustering Agglomerative [15], Clustering-based Similarity Partitioning Algorithm [16], Locally Weighted Evidence Accumulation [17], Iterative Voting Consensus [18]. These methods are briefly described below.

1) Agglomerative clustering: The agglomerative method is a bottom-up algorithm for the correlation clustering problem [15]. This method starts placing every object into a singleton cluster. The algorithm tries to merge clusters based on a calculated distance. The initial ensemble is used to compute the distance between each pair of objects based on the number of times they were placed in the same cluster on different clusterings.

2) Clustering-based Similarity Partitioning Algorithm : The Cluster-based Similarity Partitioning Algorithm (CSPA) creates a similarity graph from the co-ocurrence matrix [19]. After that, METIS [20] method is used to partition the graph into K clusters of roughly equal size.

3) Locally Weighted Ensemble Accumulation: This method is based on hierarchical agglomerative clustering. Ensemble-driven cluster uncertainty estimation and local weighting strategy are used in LWEA. The uncertainty of each cluster is estimated by considering the cluster labels in the entire ensemble via an entropic criterion. The local weighting strategy refines the co-association matrix using an ensemble-driven cluster validity [17].

4) Locally Weighted Graph Partitioning: This method is based on bipartite graph formulating and partitioning. First the algorithm constructs the locally weighted bipartite graph. Then, the Tcut algorithm is used to partition the graph. Therefore, a final partition can be obtained [17].

5) Iterative Voting Consensus: The Iterative Voting Consensus (IVC) method aims to obtain the final partition from the label-assignment data matrix of the ensemble.

V. Empirical results

This section discusses the performance of the distributed approach to cluster multi-view relational data. Three real-world data sets were considered in this study: Image Segmentation, Multiple features and Corel images. Table I summarizes the information of data sets. The number of objects, number of attributes, number of clusters, and the number of views are presented in Table I. Each view is represented by a subset of the variables.

TABLE I
Summary of the data sets

Data sets	# views	# variables	# clusters	# objects
Image	2	16	7	2310
Multiple features	6	649	10	2000
Subcorel	7	338	7	400

A. Experimental settings

The PSO-based hybrid method was run with 100 iterations for all views and all objective functions. The parameters used for the PSO algorithm were the default values: $w_{min} = 0.4$, $w_{max} = 0.9$, $c_1 = c_2 = 2$. The size of swarm used was 20 particles.

For each view, one dissimilarity matrix was computed considering all attributes in the view. In this work, the dissimilarity between each pair of objects was calculated according to the Euclidean (L_2) distance. All matrices were normalized as performed in [7]. After the execution of the clustering algorithm for all views independently and for all objective function, the obtained clusterings were assessed using three evaluation indexes: the Adjusted Rand Index (ARI) [21], F-measure [22], and Silhouette index [11]. Higher values of these two indexes indicate better clustering results.

For all data sets, each objective function was considered by the PSO-based method to generate a clustering for each view independently, i.e, the optimization of each function generates p clusterings (one per view). All consensus functions were executed to compare the performance of these methods and to verify if the optimization of each objective function could generate better final partitions, i.e, the ensembles with low diversity. Finally, the consensus functions were applied again considering all clustering results.

B. Results

Tables II-IV show the summary of the results of the clustering generated by the optimization of each view independently. The average performance and the standard variation of the indexes are also presented. The best value for each data set and each external index is shown in bold. The mean values and standard variation are also shown for each ensemble. The results of the final clustering found by each consensus function is also presented.

It is possible to observe in Table II, for Image data set, that View 2 presented better results independently of the optimized objective function. The consensus functions were applied to the sets containing only two clusterings for each objective function. Consequently, in general, the final clusterings obtained had lower quality in terms of the considered indexes.

From Table III, for Multiple features data set, it is possible to observe that, except for CS function, View 3 produced better clustering in all cases for F-measure and ARI indexes. It is noteworthy that View 4 provided poor clustering for all objective functions. Therefore, these results reinforce that some views may have noisy information and, ideally, should not be considered. The consensus functions were applied to the sets containing six clusterings for each objective function. CSPA and LWGP methods produced better results for all objective functions and all views, except for silhouette index.

From Table IV, some observations can be done. First, view 1 produced better clusterings in terms of F-measure and ARI indexes in most objective functions. IVC consensus function found the best final clusterings, considering F-measure and ARI, for the ensembles in three objective functions. The consensus functions were applied to the sets containing seven clusterings for each objective function. For all objective functions, one consensus function found a clustering better evaluated, for F-measure and ARI, than the best clustering found by optimizing all views separately.

Table V shows the summary of the results considering all clusterings found for all objective functions, i.e, the complete ensemble. The mean and standard variation is presented for the three indexes. The results of the final clustering found by each consensus function are also presented. In general, as can be seen in Table V, the proposed distributed approach achieved better results compared to the results from the optimization of each view separately considering the external indexes and the data sets used for at least one consensus function.

TABLE II
RESULTS FOR IMAGE SEGMENTATION DATA SET

	Homogeneity objective function			Consensus				
Index	View 1	View 2	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.318	**0.556**	0.437(.17)	0.087	0.115	0.011	0.112	**0.165**
F-Measure	0.405	**0.615**	0.510(.15)	0.429	0.535	0.452	0.534	**0.595**
ARI	0.219	**0.457**	0.338(.17)	0.275	0.326	0.188	0.375	**0.431**
	Silhouette objective function			Consensus				
Index	View 1	View 2	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.355	**0.580**	0.468(.16)	0.019	0.129	0.075	**0.141**	0.105
F-Measure	0.410	**0.653**	0.532(.17)	0.497	**0.565**	0.549	0.535	0.53
ARI	0.204	**0.467**	0.336(.19)	0.334	**0.376**	0.286	0.368	0.343
	Davies-Bouldin objective function			Consensus				
Index	View 1	View 2	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.155	**0.484**	0.320(.23)	-0.037	0.044	0.125	**0.229**	0.053
F-Measure	0.435	**0.666**	0.551(.16)	0.568	0.465	0.58	**0.665**	0.479
ARI	0.159	**0.529**	0.344(.26)	0.355	0.221	0.312	**0.526**	0.237
	Dunn objective function			Consensus				
Index	View 1	View 2	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.304	**0.552**	0.428(.18)	0.089	0.195	0.166	0.101	**0.201**
F-Measure	0.400	**0.580**	0.490(.13)	0.424	**0.618**	0.564	0.504	0.538
ARI	0.203	**0.396**	0.300(.14)	0.248	0.436	0.353	0.339	**0.447**
	CS objective function			Consensus				
Index	View 1	View 2	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.350	**0.554**	0.452(.14)	**0.112**	0.041	0.043	0.104	0.031
F-Measure	0.398	**0.579**	0.489(.13)	0.422	0.431	0.438	**0.503**	0.433
ARI	0.185	**0.402**	0.294(.15)	0.259	0.233	0.219	**0.343**	0.247

TABLE III
RESULTS FOR MULTIPLE FEATURES DATA SET

	Homogeneity objective function							Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.220	0.157	0.143	**0.572**	0.120	0.169	0.230(.17)	-0.002	0.123	0.050	**0.128**	0.120
F-Measure	0.723	0.707	**0.835**	0.455	0.823	0.667	0.702(.13)	0.834	**0.873**	0.633	0.798	**0.873**
ARI	0.541	0.541	**0.679**	0.317	0.661	0.447	0.531(.13)	0.716	**0.749**	0.470	0.655	0.747
	Silhouette objective function							Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.248	0.160	0.147	**0.578**	0.131	0.178	0.240 (0.17)	-0.016	**0.114**	0.056	0.076	0.111
F-Measure	0.655	0.731	**0.762**	0.457	0.714	0.629	0.658(.11)	0.758	**0.871**	0.551	0.773	0.858
ARI	0.476	0.548	**0.606**	0.323	0.559	0.435	0.491(.10)	0.619	**0.743**	0.38	0.632	0.723
	Davies-Bouldin objective function							Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.19	0.093	0.118	**0.495**	0.08	0.145	0.187 (.16)	-0.0166	0.096	0.097	**0.101**	0.095
F-Measure	0.579	0.522	**0.663**	0.479	0.582	0.563	0.565(.06)	0.665	**0.776**	0.711	0.743	0.769
ARI	0.396	0.368	**0.489**	0.326	0.395	0.346	0.387(.06)	0.512	**0.602**	0.537	0.571	0.595
	Dunn objective function							Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.209	0.127	0.134	**0.574**	0.12	0.169	0.222(.18)	-0.015	**0.119**	0.08	0.108	0.118
F-Measure	0.68	0.555	**0.785**	0.461	0.675	0.625	0.630(.11)	0.752	**0.807**	0.684	0.761	0.801
ARI	0.512	0.408	**0.628**	0.322	0.516	0.412	0.466(.11)	0.603	**0.686**	0.534	0.631	0.675
	CS objective function							Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.223	0.128	0.132	**0.562**	0.128	0.169	0.224(.17)	0.013	0.1	0.056	**0.131**	0.098
F-Measure	0.59	0.683	0.656	0.468	**0.686**	0.579	0.610(.08)	0.692	**0.808**	0.592	0.706	0.794
ARI	0.424	0.493	0.501	0.313	**0.542**	0.36	0.439(.09)	0.561	**0.675**	0.434	0.569	0.66

VI. CONCLUSION

This study proposed a distributed approach for the clustering of multi-view relational data. Most approaches in literature deal with multi-view data through a centralized approach. In the proposed approach, each view is used independently to generate clusterings of the data sets considering five different objective functions. A hybrid method based on the PSO was used to produce the clusterings by optimizing the different objective functions. Five different consensus functions were compared to generate a final clustering of the data sets.

Three real-world multi-view data sets and three clustering validity indexes were considered in this study. The empirical results showed that the proposed approach was able to improve the clustering results for at least one of the considered consensus functions.

As future work, we intend to investigate a strategy to perform view selection. View selection may be important due to complexity reduction and also to improve the clustering

TABLE IV
RESULTS FOR SUBCOREL-1 DATA SET

	Homogeneity objective function								Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	View 7	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.198	0.240	0.189	0.226	**0.345**	0.23	0.292	0.246(.05)	-0.063	**0.094**	0.08	0.083	0.092
F-Measure	**0.699**	0.664	0.585	0.457	0.567	0.55	0.427	0.564(.09)	0.614	0.639	**0.743**	0.631	0.637
ARI	**0.415**	0.311	0.3	0.114	0.336	0.184	0.079	0.248(.09)	0.388	0.396	**0.465**	0.367	0.393
	Silhouette objective function								Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	View 7	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.224	0.284	0.193	0.238	**0.378**	0.24	0.325	0.269(.06)	-0.036	0.089	0.073	**0.095**	0.088
F-Measure	0.611	**0.642**	0.611	0.479	0.577	0.517	0.443	0.554(.07)	0.601	0.675	**0.698**	0.61	0.67
ARI	0.32	0.306	0.331	0.123	**0.335**	0.143	0.091	0.236(.11)	0.339	0.411	**0.418**	0.385	0.404
	Davies-Bouldin objective function								Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	View 7	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.182	0.279	0.173	0.208	**0.330**	0.163	0.242	0.225(.06)	-0.069	**0.083**	0.074	0.060	0.082
F-Measure	0.648	**0.668**	0.604	0.443	0.541	0.496	0.461	0.552(.09)	0.648	0.647	0.647	**0.683**	0.65
ARI	0.349	**0.350**	0.285	0.106	0.323	0.137	0.117	0.238(.11)	0.38	**0.411**	0.397	0.410	**0.411**
	Dunn objective function								Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	View 7	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.201	0.279	0.150	0.233	**0.334**	0.209	0.292	0.243(.06)	-0.066	0.101	0.082	**0.129**	0.085
F-Measure	**0.709**	0.663	0.648	0.455	0.618	0.524	0.395	0.573(.11)	0.617	0.677	**0.741**	0.712	0.735
ARI	**0.413**	0.343	0.343	0.116	0.363	0.167	0.073	0.260(.13)	0.352	0.427	**0.487**	0.417	0.459
	CS objective function								Consensus				
Index	View 1	View 2	View 3	View 4	View 5	View 6	View 7	$\mu(\sigma)$	Agglomerative	CSPA	IVC	LWEA	LWGP
Silhouette	0.206	0.279	0.187	0.204	**0.288**	0.239	0.131	0.219(.04)	-0.047	0.084	0.056	0.02	**0.098**
F-Measure	**0.685**	0.645	0.563	0.466	0.526	0.52	0.415	0.546(.09)	0.595	0.662	0.56	0.691	**0.694**
ARI	**0.375**	0.292	0.246	0.119	0.235	0.148	0.076	0.213(.10)	0.355	0.396	0.249	0.405	**0.415**

TABLE V
SUMMARY OF RESULTS FOR ALL VIEWS

Data set	Index	Ensemble		Consensus				
		Mean	Std Dev	Agglomerative	CSPA	IVC	LWEA	LWGP
Image	Silhouette	0.421	0.14	0.178	0.174	0.154	**0.285**	0.259
	F-Measure	0.514	0.11	0.588	0.629	0.562	0.658	**0.757**
	ARI	0.322	0.14	0.42	0.433	0.349	0.488	**0.601**
Mfeat	Silhouette	0.221	0.157	-0.006	0.123	0.100	**0.130**	0.126
	F-Measure	0.633	0.107	0.783	0.879	0.715	0.782	**0.895**
	ARI	0.463	0.117	0.646	0.646	0.590	0.652	**0.789**
SubCorel	Silhouette	0.245	0.056	-0.016	0.093	0.089	**0.103**	0.091
	F-measure	0.558	0.091	0.614	0.685	0.675	**0.688**	0.682
	ARI	0.239	0.112	0.374	0.432	0.394	**0.421**	0.418

accuracy since data sets may have views containing noisy information.

REFERENCES

[1] A. K. Jain, "Data clustering: 50 years beyond k-means," Pattern Recognition Letters, vol. 31, no. 8, pp. 651 – 666, 2010, award winning papers from the 19th International Conference on Pattern Recognition (ICPR).

[2] H. Frigui, C. Hwang, and F. Chung-Hoon Rhee, "Clustering and aggregation of relational data with applications to image database categorization," Pattern Recogn., vol. 40, no. 11, pp. 3053–3068, Nov. 2007.

[3] Y. Yang and H. Wang, "Multi-view clustering: A survey," Big Data Mining and Analytics, vol. 1, no. 2, pp. 83–107, June 2018.

[4] Y.-M. Xu, C.-D. Wang, and J.-H. Lai, "Weighted multi-view clustering with feature selection," Pattern Recognition, vol. 53, pp. 25 – 35, 2016.

[5] G. Cleuziou, M. Exbrayat, L. Martin, and J. Sublemontier, "Cofkm: A centralized method for multiple-view clustering," in 2009 Ninth IEEE International Conference on Data Mining, 2009, pp. 752–757.

[6] R. P. Gusmão and F. A. T. de Carvalho, "Particle swarm optimization applied to relational data clustering," in 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Oct 2016, pp. 001 690–001 695.

[7] R. P. [de Gusmão] and F. de A.T. [de Carvalho], "Clustering of multi-view relational data based on particle swarm optimization," Expert Systems with Applications, vol. 123, pp. 34 – 53, 2019.

[8] R. P. de Gusmão and F. d. A. T. de Carvalho, "Pso for fuzzy clustering of multi-view relational data," International Journal of Pattern Recognition and Artificial Intelligence, vol. 0, no. 0, p. 2050022, 0. [Online]. Available: https://doi.org/10.1142/S0218001420500226

[9] Y. Wang, L. Chen, and X. Li, "Multiple medoids based multi-view relational fuzzy clustering with minimax optimization," in Proceedings of the 26th International Joint Conference on Artificial Intelligence, ser. IJCAI'17. AAAI Press, 2017, p. 2971–2977.

[10] J.-P. Mei and L. Chen, "Fuzzy clustering with weighted medoids for relational data," Pattern Recognition, vol. 43, no. 5, pp. 1964 – 1974, 2010.

[11] P. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," J. Comput. Appl. Math., vol. 20, no. 1, pp. 53–65, Nov. 1987.

[12] D. L. Davies and D. W. Bouldin, "A cluster separation measure," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-1, no. 2, pp. 224–227, April 1979.

[13] J. C. Dunn, "A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters," Journal of Cybernetics, vol. 3, no. 3, pp. 32–57, 1973.

[14] C.-H. Chou, M.-C. Su, and E. Lai, "A new cluster validity measure and its application to image compression," Pattern Analysis and Applications, vol. 7, no. 2, pp. 205–220, Jul 2004.

[15] A. Gionis, H. Mannila, and P. Tsaparas, "Clustering aggregation," ACM Trans. Knowl. Discov. Data, vol. 1, no. 1, p. 4–es, Mar. 2007.

[16] A. Strehl and J. Ghosh, "Cluster ensembles — a knowledge reuse

framework for combining multiple partitions," J. Mach. Learn. Res., vol. 3, p. 583–617, Mar. 2003.

[17] D. Huang, C. Wang, and J. Lai, "Locally weighted ensemble clustering," IEEE Transactions on Cybernetics, vol. 48, no. 5, pp. 1460–1473, 2018.

[18] N. Nguyen and R. Caruana, "Consensus clusterings," in Seventh IEEE International Conference on Data Mining (ICDM 2007), 2007, pp. 607–612.

[19] T. Boongoen and N. Iam-On, "Cluster ensembles: A survey of approaches with recent extensions and applications," Computer Science Review, vol. 28, pp. 1 – 25, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574013717300692

[20] G. Karypis and V. Kumar, "A fast and high quality multilevel scheme for partitioning irregular graphs," SIAM J. Sci. Comput., vol. 20, no. 1, p. 359–392, Dec. 1998.

[21] L. Hubert and P. Arabie, "Comparing partitions," Journal of Classification, vol. 2, no. 1, pp. 193–218, 1985.

[22] C. J. V. Rijsbergen, Information Retrieval, 2nd ed. Newton, MA, USA: Butterworth-Heinemann, 1979.