A Novel Multi-View Clustering Method for Unknown Mapping Relationships Between Cross-View Samples

Hong Yu

yuhong@cqupt.edu.cn

Chongqing Key Laboratory of Computational Intelligence
Chongqing University of Posts and Telecommunications
Chongqing, China

Guoyin Wang

wanggy@cqupt.edu.cn

Chongqing Key Laboratory of Computational Intelligence
Chongqing University of Posts and Telecommunications
Chongqing, China

Jia Tang

tangjia0726@163.com

Chongqing Key Laboratory of Computational Intelligence
Chongqing University of Posts and Telecommunications
Chongqing, China

Xinbo Gao

gaorb@cqupt.edu.cn

Chongqing Key Laboratory of Computational Intelligence
Chongqing University of Posts and Telecommunications
Chongqing, China

ABSTRACT

The existing multi-view clustering algorithms require that a sample in a view is completely or partially mapped onto one or more samples in a different corresponding view. However, this requirement could not be satisfied in many practical applications. Fortunately, there is a common cognition that the graph structure formed from each view should be as consistent as possible. Thus, this paper proposes a novel multi-view clustering method for unknown mapping relationships between cross-view samples based on the framework of non-negative matrix factorization, as an attempt to solve this problem. The objective function is designed by effectively building reconstruction error terms, local structural constraint terms, and cross-view mapping loss terms by exploring cross-view relationships. The experimental results show that the proposed method not only performs well to reveal the real mapping relationships between cross-view samples but also outperforms the comparison algorithms on the obtained clustering results.

CCS CONCEPTS

• **Computing methodologies** → **Machine learning; Cluster analysis.**

KEYWORDS

multi-view clustering, unknown mapping relationships, non-negative matrix factorization

ACM Reference Format:

Hong Yu, Jia Tang, Guoyin Wang, and Xinbo Gao. 2021. A Novel Multi-View Clustering Method for Unknown Mapping Relationships Between Cross-View Samples. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '21), August 14–18, 2021*.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

KDD '21, August 14–18, 2021, Virtual Event, Singapore

© 2021 Association for Computing Machinery.

ACM ISBN 978-1-4503-8332-5/21/08...\$15.00

<https://doi.org/10.1145/3447548.3467294>

Virtual Event, Singapore. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3447548.3467294>

1 INTRODUCTION

Multi-view clustering is of great ability to explore both correlated and complementary information carried by multi-view data. Based on the given information, the mapping relationships between cross-view samples, the existing researches on this topic have provided significant insights into both completely and partially mapped multi-view data. However, such a requirement could not be satisfied in many practical applications. The data from different views are usually collected, processed and stored separately, and it is expensive and time-consuming to establish the mapping relationships between cross-view samples. For example, it is difficult to determine whether two accounts from different platforms belong to the same user, due to privacy protection. In this case, the different platforms are the different views and the accounts of platforms are the samples in the corresponding views, so the collected data are the typical multi-view data with unknown mapping relationships between cross-view samples.

The challenges of this question are specified as follows. First, as multi-view data are heterogeneous from multi-sources, it is beyond the capabilities to directly explore mappings between cross-view samples utilizing vector norm on the completely unmapped multi-view data. Second, it is difficult to capture a consistent representation for all views, due to the lack of the association between a sample of a view and the underlying corresponding one in another view. Third, the mapping relationships are complex. A sample may not have corresponding samples in some views, or it has more than one corresponding samples in a different view. So it is still a challenging problem to determine automatically the complex mapping relationship between cross-view samples.

To address the problem, this work proposes a novel method, aiming at clustering on multi-view data with unknown mapping relationships between cross-view samples while finding the real cross-view mapping relationships. The basic idea of solving this problem comes from the common cognition that the graph structure formed from each view should be as consistent as possible. Thus, a novel multi-view clustering method for unknown mapping

relationships (MVC-UM) is proposed, and it is based on the framework of non-negative matrix factorization (NMF). In this work, the original multi-view data are viewed as various projections in the same non-negative space and restored by constructing reconstruction error terms. Then, the cross-view mapping loss terms are constructed via aligning the graph structures from views by vertex permutation, which consist of the structure-matching loss terms and the non-negative consistent representation loss terms. Finally, the local structural constraint terms are introduced to preserve the intrinsic geometrical structural information.

We summarize our main contributions as follows:

- To the best of our knowledge, it is the first attempt to present and discuss the problem of clustering on unmapped multi-view data. Moreover, we take the lead in sorting out the existing multi-view clustering work from the perspective of mapping relationships.
- A novel multi-view data clustering for unknown mapping relationships between cross-view samples is proposed based on the framework of NMF. The proposed method can explore the true cross-view mapping relationships while obtaining the final clustering results.

The remainder of this paper is organized as follows. Section 2 briefly reviews the related work from the perspective of mapping relationships. Section 3 describes the proposed method. Section 4 reports the results of comparison experiments, and conclusions are given in Section 5. Codes are available at <https://github.com/yuhongcqupt/MVC-UM>.

2 RELATED WORK

The multi-view clustering algorithms are divided into three categories from the perspective of mapping relationships between cross-view samples. Without loss of generality, the following descriptions are based on two views.

2.1 Clustering on Completely Mapped Multi-View Data

Completely mapped multi-view data means that all mapping relationships between any two samples are given. The mappings between cross-view samples and the mappings between samples in the same view are all known, as shown in Fig. 1. There are two classes, namely, the complete simple mapping relationships and the complete complex mapping relationships.

Clustering on multi-view data with complete simple mapping relationships has been widely studied. As shown in Fig. 1(a), for any sample x_i^a in the view X^a , there is one and only one sample x_i^b in the view X^b that associates with it. That is, $f_b(x_i^a) = x_i^b$, where f_b is the mapping function. Obviously, the number of samples in X^a is equal to the number of samples in X^b , i.e. $n_a = n_b$. Bickel and Scheffer [1] took the lead in extending the traditional K -means clustering algorithm to two views, which can be regarded as one of the earliest multi-view clustering researches. To improve the robustness of multi-view K -means, the RMKMC [4] assigns different weights to different views and adopts $\mathcal{L}_{2,1}$ -norm to guide multi-view clustering. Besides, NMF is also widely used in multi-view clustering [9, 22]. The main idea is to utilize matrix decomposition

to find the matrix of potential consistency indicators shared by all views, so as to get the final clustering result. Other common approaches [12, 15, 16] try to obtain the consistent graph structure representation by exploring the geometric consistency of different views. Furthermore, Gao et al. [8] extended the subspace clustering to multi-view clustering. Each view uses $\mathcal{F}$ -norm to find a potentially consistent self-representation matrix to cluster multi-view data. To further utilize complementary information and obtain robust clustering results, Zhang et al. [18] proposed Latent Multi-view Subspace Clustering (LMSC). On the basis of obtained potential representation, LMSC generates a common subspace for all views and reconstructs raw multi-view data.

The complete complex mapping relationship, as shown in Fig. 1(b), means that any object can have $m(\geq 0)$ samples in any view. In the example, there are three samples in the view X^a to reflect a certain object, and they form a set $\{1, 2, 3\}$. So, its mapped samples in the view X^b form the corresponding set $\{a, b\}$, i.e., $f_b(\{1, 2, 3\}) = \{a, b\}$. Similarly, the set of samples in the view X^b , $\{a, b\}$, its mapped samples in the view X^a form the set $\{1, 2, 3\}$, i.e., $f_a(\{a, b\}) = \{1, 2, 3\}$. In this situation, only few methods have been proposed [6, 17]. The CGC [6] is the first algorithm to solve multi-view data with complete complex mapping relationship, which is based on matrix decomposition and cross-view weight allocation. Yu et al. [17] argued that the requirement for CGC that all mapping relationships should be given, but the requirement cannot be satisfied in many practical applications. They proposed the CMMVC algorithm to automatically explore the complex mapping relationships between samples in the same view by constructing a complex matching relationship matrix in each view.

It is obvious that the multi-view data with complete simple mapping relationships are a special case of multi-view data with the complete complex mapping relationships.

Figure 1: Completely mapped multi-view data. Solid black lines indicate the real and known mapping relationships between cross-view samples. The connecting line represents that the adjacent samples describe the same object in different views.

2.2 Clustering on Partially Mapped Multi-View Data

Partially mapped multi-view data means that only partial mapping relationships between samples are given. As shown in Fig. 2, there are also two classes.

In Fig. 2(a), the connection between the sample 2 in the view X^a and the sample b in the view X^b is not given, though they describe the same object; and it is also unknown that the sample 2, the sample 3 and the sample b are the samples of the same object due to the lack of mapping relationships in Fig. 2(b). In earlier studies, Eaton et al. [7] proposed a multi-view clustering algorithm to handle partially mapped multi-view data. The main idea is to construct constraints by the existing mapping relationships and expand the constraints through cross-view propagation. Based on the clustering framework of NMF, Zhang et al. [19] proposed a novel method named CMVNMf, which constructs existing mapping relationships as different types of constraint matrixes. Zong et al. [21] further extended the CMVNMf via adding a constraint propagation mechanism.

(a) Multi-view data with partial simple mapping relationships (b) Multi-view data with partial complex mapping relationships

Figure 2: Partially mapped multi-view data. Solid black lines indicate the real and known mappings, dashed red lines indicate the real but unknown mappings.

2.3 Clustering on Unknown Mapped Multi-View Data

Unknown mapped multi-view data means that no mapping relationships between any two samples are given. So, all connections are dashed in Fig. 3. No mapping is given though they exist in reality. There are also two classes.

To the best of our knowledge, there is still a gap in the research on unknown mapped multi-view data. This work is the first attempt to cluster on multi-view data with unknown mappings.

3 THE PROPOSED METHOD

In this section, we introduce a multi-view clustering method for unknown mapping relationships based on the framework of NMF. An overview of our framework is given by the graphical model depicted in Fig. 4.

(a) Multi-view data with unknown simple mapping relationships (b) Multi-view data with unknown complex mapping relationships

Figure 3: Unknown mapped multi-view data. Dashed red lines indicate the real but unknown mappings.

3.1 Construct Reconstruction Error Terms

Non-negative Matrix Factorization [13] aims to decompose a matrix into two or more non-negative matrices. In such a way, the product of these matrices reduces the original matrix as much as possible. The focus of NMF is on the non-negativity of the matrix elements. Given the original data matrix $X \in \mathbb{R}_+^{d \times n}$, where n denotes the number of samples in the original data and d denotes the sample dimension, the ultimate goal of NMF is to optimize the following problem:

$$\min_{W, H} \|X - WH^T\|_{\mathcal{F}}^2 \text{ s.t. } W \geq 0, H \geq 0, \quad (1)$$

where $W \in \mathbb{R}_+^{d \times k}$, $H \in \mathbb{R}_+^{n \times k}$, $\|\cdot\|_{\mathcal{F}}$ is $\mathcal{F}$ -norm and k is the number of hidden patterns. It's known that the objective function above is not convex in W and H together, and it's hard to find its global minimum. Lee et al. [13] presented “multiplicative update rules” to execute iteratively to minimize the objective function as follows:

$$W_{ij} \leftarrow W_{ij} \frac{(XH)_{ij}}{(WH^T H)_{ij}}, H_{ij} \leftarrow H_{ij} \frac{(X^T W)_{ij}}{(HW^T W)_{ij}}. \quad (2)$$

Given multi-view data $X = \{X^1, X^2, \dots, X^v\}$, where v represents the number of views, MVC-UM aims to factorize the original view $X^a \in \mathbb{R}^{d_a \times n_a}$ into two non-negative matrices $W^a \in \mathbb{R}_+^{d_a \times k}$ and $H^a \in \mathbb{R}_+^{n_a \times k}$, where d_a represents the dimension of samples in the view X^a , n_a represents the total number of samples in the view X^a , and k is consistent non-negative dimension. The matrix H^a can be regarded as a non-negative representation of the original data in the non-negative space. In traditional multi-view clustering with NMF, a shared non-negative representation matrix H^* can be found straight for multi-view data since the mapping relationships are complete and simple. The corresponding objective function is as follows:

$$\min_{W^a, H^a, H^*} \sum_{a=1}^v \|X^a - W^a (H^a)^T\|_{\mathcal{F}}^2 + \sum_{a=1}^v \lambda_a \|H^a - H^*\|_{\mathcal{F}}^2 \quad (3)$$

$$\text{s.t. } \forall 1 \leq a \leq v, W^a \geq 0, H^a \geq 0, H^* \geq 0,$$

where λ_a represents the balance parameter, which can adjust the relative weight between different views, as well as between standard

Figure 4: Overview of the framework of MVC-UM. The purple connecting lines denote the mapping relationships found between the cross-view samples and the blue connecting lines represent the mappings found within a view.

NMF reconstruction error terms and uniform non-negative representation terms. The number of views is denoted by v . However, due to the lack of cross-view mapping relationships, it is difficult to learn a unified non-negative representation matrix for all views directly. Therefore, to solve the different dimensions of multi-view data, this work finds a non-negative representation in the same non-negative space for each view respectively. The corresponding objective function is as follows:

$$\min_{\mathbf{W}^a, \mathbf{H}^a} \sum_{a=1}^v \|\mathbf{X}^a - \mathbf{W}^a (\mathbf{H}^a)^T\|_{\mathcal{F}}^2 \quad (4)$$

$$s.t. \forall 1 \leq a \leq v, \mathbf{W}^a \geq 0, \mathbf{H}^a \geq 0.$$

3.2 Construct Cross-view Mapping Loss Terms

To perform multi-view clustering, we need a mechanism to obtain the real cross-view mapping relationships. The cross-view mapping matrix $\mathbf{P}^{ab} \in \mathbb{R}_+^{n_a \times n_b}$ is used to represent the matching degree of cross-view samples between view $\mathbf{X}^a$ and view $\mathbf{X}^b$. Based on the common cognition that the graph structure formed from multi-view data should be as consistent as possible, structure-matching loss terms are constructed to explore the real cross-view mapping relationships. First, we build the respective adjacency matrix $\mathbf{S}^a$ for each view. And then, the connection structures of the two views are made as similar as possible by transposing the adjacency matrix $\mathbf{S}^a$ or $\mathbf{S}^b$ through the cross-view mapping matrix $\mathbf{P}^{ab}$. Finally, structure-matching loss term is formulated as follows:

$$\sum_{1 \leq a < b \leq v} \left\| (\mathbf{P}^{ab})^T \mathbf{S}^a (\mathbf{P}^{ab}) - \mathbf{S}^b \right\|_{\mathcal{F}}^2 \quad (5)$$

$$s.t. \mathbf{P}^{ab} \geq 0$$

where $s_{ij}^a = \|\mathbf{x}_i^a - \mathbf{x}_j^a\|^2$.

Besides, we require that the non-negative representations of two well-matched cross-view samples should be as similar as possible. Accordingly, non-negative consistent representation loss term can

be constructed as follows:

$$\sum_{1 \leq a < b \leq v} \left\| (\mathbf{P}^{ab})^T \mathbf{H}^a - \mathbf{H}^b \right\|_{\mathcal{F}}^2 \quad (6)$$

$$s.t. \mathbf{P}^{ab} \geq 0.$$

Finally, through the above analysis, the formula of the entire cross-view mapping loss term is as follows:

$$\sum_{1 \leq a < b \leq v} \lambda_a \left\| (\mathbf{P}^{ab})^T \mathbf{S}^a (\mathbf{P}^{ab}) - \mathbf{S}^b \right\|_{\mathcal{F}}^2 + \mu_a \left\| (\mathbf{P}^{ab})^T \mathbf{H}^a - \mathbf{H}^b \right\|_{\mathcal{F}}^2 \quad (7)$$

$$s.t. \mathbf{P}^{ab} \geq 0.$$

3.3 Construct Local Structural Constraint Terms

Considering that the traditional multi-view clustering with NMF ignores the internal structure of data, we use the GNMF method proposed by Cai et al. [3] to construct the local structural constraint terms, which can maintain the local structure of data. That is, if the $\mathbf{x}_i^a$ and $\mathbf{x}_j^a$ are similar, $\mathbf{h}_i^a$ and $\mathbf{h}_j^a$ should be similar in reconstructed representation. Let's use $\mathcal{D}(\mathbf{h}_i^a, \mathbf{h}_j^a) = \|\mathbf{h}_i^a - \mathbf{h}_j^a\|^2$ to represent the similarity between any two samples from the same view in the reconstructed representation matrix $\mathbf{H}^a$. Each adjacency matrix $\mathbf{S}^a$ is used to construct the local structural constraint terms to keep the original structure. Therefore, the local structural constraint term is as follows:

$$\mathcal{R}^a = \frac{1}{2} \sum_{i,j} \left\| \mathbf{h}_i^a - \mathbf{h}_j^a \right\|_{\mathcal{F}}^2 s_{ij}^a = \text{tr}((\mathbf{H}^a)^T \mathbf{L}^a \mathbf{H}^a), \quad (8)$$

where $\text{tr}(\cdot)$ is the trace of the matrix, $\mathbf{L}^a = \mathbf{D}^a - \mathbf{S}^a$ denotes the Laplace matrix, and $\mathbf{D}^a$ is the degree matrix, a diagonal matrix meeting with $\mathbf{D}_{ii}^a = \sum_j s_{ij}^a, i = 1, 2, \dots, n_a$.

3.4 The Objective Function of MVC-UM

Based on the above analysis, the final objective function of MVC-UM is as follows:

$$\begin{aligned}
\min_{\mathbf{W}^a, \mathbf{H}^a, \mathbf{P}^{ab}} & \sum_{a=1}^v \left\| \mathbf{X}^a - \mathbf{W}^a (\mathbf{H}^a)^T \right\|_{\mathcal{F}}^2 + 2\alpha_a \mathcal{R}^a \\
& + \sum_{1 \leq a < b \leq v} \lambda_a \left\| (\mathbf{P}^{ab})^T \mathbf{S}^a (\mathbf{P}^{ab}) - \mathbf{S}^b \right\|_{\mathcal{F}}^2 \\
& + \sum_{1 \leq a < b \leq v} \mu_a \left\| \mathbf{P}^{ab} \mathbf{H}^b - \mathbf{H}^a \right\|_{\mathcal{F}}^2 \\
\text{s.t. } & \forall 1 \leq a \leq v, \mathbf{W}^a \geq 0, \mathbf{H}^a \geq 0, \\
& \forall 1 \leq a < b \leq v, \mathbf{P}^{ab} \geq 0.
\end{aligned} \tag{9}$$

In the formula (9), α_a denotes the local structure constraint balance parameter, λ_a is the structure-matching loss balance parameter, and μ_a represents the non-negative consistent loss balance parameter.

Algorithm 1 Multi-view Clustering for Unknown Mappings(MVC-UM)

Input: Non-negative multi-view data matrices $\mathbf{X} = \{\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^v\}$ with unknown mapping relationships between cross-view samples, parameters $\{\lambda_1, \lambda_2, \dots, \lambda_v, \mu_1, \mu_2, \dots, \mu_v, \alpha_1, \alpha_2, \dots, \alpha_v\}$, the number of clusters c .

Output: Basis matrices $\{\mathbf{W}^1, \mathbf{W}^2, \dots, \mathbf{W}^v\}$, reconstructed representation matrices $\{\mathbf{H}^1, \mathbf{H}^2, \dots, \mathbf{H}^v\}$, the cross-view mapping matrices $\mathbf{P}^{ab}$, $1 \leq a < b \leq v$ and the clustering results.

- 1: Normalize each view $\mathbf{X}^a$ for $1 \leq a \leq v$.
 - 2: Construct respective adjacency matrix $\mathbf{S}^a$.
 - 3: Construct Graph Laplacians $\mathbf{L}^a$ for each view.
 - 4: Initialize $\mathbf{W}^a$ and $\mathbf{H}^a$ for $1 \leq a \leq v$ by adopting the random initialization method.
 - 5: Initialize the cross-view mapping matrices $\mathbf{P}^{ab}$ for $1 \leq a < b \leq v$ by setting 1 for all elements.
 - 6: **repeat**
 - 7: **for** $a = 1$ to v **do**
 - 8: Fixing $\mathbf{H}^a$ and each $\mathbf{P}^{ab}$, update $\mathbf{W}^a$ by the formula (12).
 - 9: Fixing $\mathbf{W}^a$ and each $\mathbf{P}^{ab}$, update $\mathbf{H}^a$ by the formula (14).
 - 10: **end for**
 - 11: **for** $a = 1$ to v **do**
 - 12: **for** $b = a + 1$ to v **do**
 - 13: Fixing $\mathbf{H}^a$ and $\mathbf{W}^a$, update $\mathbf{P}^{ab}$ by the formula (16).
 - 14: $\mathbf{P}^{ba} := (\mathbf{P}^{ab})^T$.
 - 15: **end for**
 - 16: **end for**
 - 17: **until** objective function in the formula (3) converges.
 - 18: Apply K -means on each $\mathbf{H}^a$ to get the clustering result.
 - 19: According to the corresponding cross-view mapping matrix $\mathbf{P}^{ab}$, construct mapping relationships of samples between view $\mathbf{X}^a$ and view $\mathbf{X}^b$.
-

3.5 Optimization

In this section, we design an algorithm to optimize the objective function. Since the whole objective function is a non-convex optimization function, it is difficult to simultaneously solve $\{\mathbf{W}^a\}_{a=1}^v$, $\{\mathbf{H}^a\}_{a=1}^v$ and the cross-view mapping matrix $\mathbf{P}^{ab}$ for $1 \leq a, b \leq v$, $a \neq b$. But for each variable, the objective function is a convex optimization. So we find a local minimum by alternately solving for each variable. Taking view $\mathbf{X}^a$ as an example, the objective function (9) is changed into the following form by introducing Lagrange multiplier Ψ, Φ, Θ :

$$\begin{aligned}
\mathcal{L}_a = & \left\| \mathbf{X}^a - \mathbf{W}^a (\mathbf{H}^a)^T \right\|_{\mathcal{F}}^2 \\
& + \lambda_a \sum_{\substack{1 \leq b \leq v \\ a \neq b}} \left\| (\mathbf{P}^{ab})^T \mathbf{S}^a \mathbf{P}^{ab} - \mathbf{S}^b \right\|_{\mathcal{F}}^2 \\
& + \mu_a \sum_{\substack{1 \leq b \leq v \\ a \neq b}} \left\| \mathbf{P}^{ab} \mathbf{H}^b - \mathbf{H}^a \right\|_{\mathcal{F}}^2 \\
& + \alpha_a \text{tr}((\mathbf{H}^a)^T \mathbf{L}^a \mathbf{H}^a) \\
& + \text{tr}(\Psi \mathbf{W}^a) + \text{tr}(\Phi \mathbf{H}^a) + \text{tr}(\Theta \mathbf{P}^{ab}).
\end{aligned} \tag{10}$$

Fixing $\mathbf{H}^a, \mathbf{P}^{ab}$, minimize $\mathcal{L}_a$ over $\mathbf{W}^a$: the derivative of $\mathcal{L}_a$ in the formula (10) with respect to $\mathbf{W}^a$ is

$$\frac{\partial \mathcal{L}_a}{\partial \mathbf{W}^a} = 2\mathbf{W}^a (\mathbf{H}^a)^T \mathbf{H}^a - 2\mathbf{X}^a \mathbf{H}^a + \Psi. \tag{11}$$

Using KKT conditions[2], i.e., $\Psi_{ij} \mathbf{W}_{ij}^a = 0$, we can get the following iterative updating rule about $\mathbf{W}^a$:

$$\mathbf{W}_{ij}^a \leftarrow \mathbf{W}_{i,j}^a \frac{(\mathbf{X}^a \mathbf{H}^a)_{ij}}{(\mathbf{W}^a (\mathbf{H}^a)^T \mathbf{H}^a)_{ij}}. \tag{12}$$

Fixing $\mathbf{W}^a, \mathbf{P}^{ab}$, minimize $\mathcal{L}_a$ over $\mathbf{H}^a$: the derivative of $\mathcal{L}_a$ in the formula (10) with respect to $\mathbf{H}^a$ is

$$\begin{aligned}
\frac{\partial \mathcal{L}_a}{\partial \mathbf{H}^a} = & 2\mathbf{H}^a (\mathbf{W}^a)^T \mathbf{W}^a - 2(\mathbf{X}^a)^T \mathbf{W}^a \\
& + 2\alpha_a (\mathbf{D}^a + (\mathbf{D}^a)^T) \mathbf{H}^a - 2\alpha_a (\mathbf{S}^a + (\mathbf{S}^a)^T) \mathbf{H}^a \\
& + 2 \sum_{\substack{1 \leq b \leq v, \\ b \neq a}} \mu_a (\mathbf{H}^a - \mathbf{P}^{ab} \mathbf{H}^b) + \Phi.
\end{aligned} \tag{13}$$

Using KKT conditions, i.e., $\Phi_{ij} \mathbf{H}_{ij}^a = 0$, we can get the following iterative updating rule about $\mathbf{H}^a$:

$$\mathbf{H}_{ij}^a \leftarrow \mathbf{H}_{i,j}^a \frac{\left((\mathbf{X}^a)^T \mathbf{W}^a + \alpha_a (\mathbf{S}^a + (\mathbf{S}^a)^T) \mathbf{H}^a + \mu_a \sum_{\substack{1 \leq b \leq v, \\ b \neq a}} \mathbf{P}^{ab} \mathbf{H}^b \right)_{ij}}{(\mathbf{H}^a (\mathbf{W}^a)^T \mathbf{W}^a + (\alpha_a (\mathbf{D}^a + (\mathbf{D}^a)^T) + (v-1)\mu_a) \mathbf{H}^a)_{ij}}. \tag{14}$$

Fixing W^a , H^a , minimize $\mathcal{L}_a$ over H^a : the derivative of $\mathcal{L}_a$ in the formula (10) with respect to P^{ab} is

$$\begin{aligned} \frac{\partial \mathcal{L}_a}{\partial P^{ab}} = & 2\mu_a(P^{ab}H^b(H^b)^T - H^a(H^b)^T) \\ & - \lambda_a(S^a P^{ab} S^b + (S^a)^T P^{ab} S^b) \\ & - \lambda_a(S^a P^{ab} (S^b)^T + (S^a)^T P^{ab} (S^b)^T) \\ & + 2\lambda_a(S^a P^{ab} (P^{ab})^T S^a + S^a P^{ab} (P^{ab})^T S^a) P^{ab} + \Theta. \end{aligned} \quad (15)$$

Using KKT conditions, i.e., $\Theta_{ij} P_{ij}^{ab} = 0$, we can get the following iterative updating rule about P^{ab} :

$$P_{ij}^{ab} \leftarrow P_{ij}^{ab} \frac{(\mu_a H^a (H^b)^T + 2\lambda_a S^a P^{ab} S^b)_{ij}}{(2\lambda_a S^a P^{ab} (P^{ab})^T S^a P^{ab} + \mu_a P^{ab} H^b (H^b)^T)_{ij}}. \quad (16)$$

Summarizing the former analysis, we give the MVC-UM algorithm framework in Algorithm 1.

3.6 Computational Complexity Analysis

In this subsection, we discuss the computational complexity of the proposed algorithm. The main computation cost of MVC-UM is the iterative updating of W^a , H^a , and the cross-view mapping matrix P^{ab} . The computational cost of update W^a is $O(d_a n_a K + K^2(d_a + n_a))$. We need $O(K^2(n_a + d_a) + n_a K(d_a + n_a + \sum_{b=1, b \neq a}^v n_b))$ to compute H^a . The d_a , n_a , and K are the dimension of the view, the number of samples, and the consistent dimension respectively. Updating each cross-view mapping matrix P^{ab} requires $O(K(n_a)^2 + 4(n_a)^2 n_b + 2(n_b)^2 n_a + (n_b)^2 K + (n_a)^3)$. If the dimension of each view is d and the number of samples is N , the total time required to optimize the objective function is $O(tC_v^2(KN^2 + N^3) + tK(2dN + Kd + 2KN + vN^2))$, where t is the number of iterations.

4 EXPERIMENTS

We conduct experiments on 7 datasets with 5 algorithms. The comparison algorithms can be divided into three categories: 1) the clustering algorithm on a single-view, including SymNMF [11], 2) the clustering algorithms on completely mapped multi-view data, including MultiNMF [9], MVSC-CEV [10] and CMMVC [17], and 3) the clustering algorithm on partially mapped multi-view data, including CMVNMF[19]. To avoid randomness, we construct 20 sub-datasets for each dataset by randomly changing the relative position of each view and ignoring real mapping relationships. All algorithms are run 20 times, and the average results are reported.

4.1 Datasets and Evaluation Criteria

The 7 real-world multi-view datasets are BBCSport¹, SensIT², ORL³, Prokaryotic⁴, Moives617⁵, Digits⁶ and Wine. The dataset Wine is sliced into two views as described in Reference [5], containing 6

¹<https://github.com/GPMVCDummy/GPMVC/tree/master/partialMV/PVC/recreateResults/data>

²<https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/multiclass.html>

³<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>

⁴<https://github.com/mbrbic/Multi-view-LRSSC/blob/master/datasets/prokaryotic.mat>

⁵<http://membres-lig.imag.fr/grimal/data.html>

⁶<https://www.dropbox.com/s/b00cdxc2s96o1to/handwritten.mat>

Figure 5: Parameters studies on the proposed MVC-UM using the SensIT dataset

Table 1: Information about the datasets

Datasets	Size	Dimensions	Views	Clusters
Wine	178	{6,7}	2	3
BBCSport ¹	282	{2582,2544,2465}	3	5
SensIT ²	300	{50,50}	2	3
ORL ³	400	{228,1024}	2	40
Prokaryotic ⁴	551	{438,3}	2	4
Moives617 ⁵	617	{1878,1398}	2	17
Digits ⁶	2000	{240,76}	2	10

and 7 attributes, respectively. The statistical information of the 7 real-world multi-view data is summarized in Table 1.

The commonly used evaluation criteria such as accuracy (ACC) [14] and normalized mutual information (NMI) [20] are used in experiments. Besides, we define a new evaluation criterion, $MACC@q$, to evaluate the correctness of the cross-view mapping relationships found. The $MACC@q$ denotes the proportion of the correct mapping relationships in the top q mapping relationships found.

Given $\forall x_i^a \in X^a$, let $A_{(a,i)}^b$ be the set of samples which includes the first q mapped samples found in the view X^b for x_i^a , according to the cross-view mapping matrix P^{ab} . The $MACC@q$ is defined as

follows:

$$\text{MACC}@q(\mathbf{p}^{ab}) = \frac{\sum_{i=1}^{n_a} \text{count}(\mathbf{x}_i^a)}{n_a}, \quad (17)$$

$$\text{count}(\mathbf{x}_i^a) = \begin{cases} 1, & \hat{\mathbf{x}}_{(a,i)}^b \in A_{(a,i)}^b, \\ 0, & \text{others} \end{cases}, \quad (18)$$

where $\hat{\mathbf{x}}_{(a,i)}^b$ represents the real sample in the view $\mathbf{X}^b$ corresponding to the sample $\mathbf{x}_i^a$.

4.2 Parameter Study

The proposed MVC-UM method contains three sets of hyperparameters λ_a , α_a and μ_a , $a = 1, \dots, v$. Fixing the parameters α , we test λ and μ within the set of $\{0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000\}$. Fig. 5 reports the MACC@1 performance of MVC-UM on the SensIT dataset. From Fig. 5, we can find that MVC-UM performs stably when the parameter μ is around 10 to 1000, the parameter λ is around 0.0001 to 0.1, and the parameter α is around 1 to 1000. Consequently, we set the values of the parameters within the corresponding range for the subsequent experiments.

4.3 Experiments on Multi-View Data With Unknown Simple Mapping Relationships

Since there are no similar multi-view clustering algorithms to deal with unknown mapping relationships, we first verify the clustering performance of the proposed MVC-UM.

Figure 6: Variation of MACC@q with increasing number of iterations.

The SymNMF [11] is a single-view clustering algorithm. The CMVNMF [19] is a clustering algorithm for partially mapped multi-view data. The input and output of the CMVNMF and the proposed MVC-UM are similar. The input is the multi-view data and the output is the respective result of each view. To simulate the scene

that there is unknown mapping relationships between cross-view samples, we set the ratio of mapping relationships as 0% in the CMVNMF algorithm.

The results in Table 2 and Fig. 6 show that the proposed MVC-UM can find a large number of real mapping relationships while obtaining stable clustering results. The NMI and the ACC of MVC-UM are best on SensIT, Digits, ORL, Wine, Prokaryotic and Moives617. For these datasets, it is easy to know the averages of MACC@1, MACC@3, and MACC@10 are 84%, 89%, 97% respectively. The results show that the mapping accuracy found by MVC-UM is very high. Almost all the real mapping relationships are found over a short times iterations. The MVC-UM only performs not well enough on the BBCSport with the clustering performance and it performs not well enough on the Movies617 with the ability of finding the mapping relationships. On the contrary, the CMVNMF only performs well on the BBCSport but performs poorly on the others. Both the BBCSport and the Movies617 are sparse datasets. One reason maybe that the high sparsity of the raw data makes the graph structure formed from each view not obvious, thus making it relatively difficult to find the real mapping relationships.

Nonetheless, it is clear that the overall MACC@q increases steadily with the number of iterations for all multi-view data with unknown mapping relationships between cross-view samples.

4.4 Experiments on Multi-View Data With a Percentage of Mapping Relationships

To further validate the effectiveness of the proposed method in multi-view clustering, we conduct experiments on more multi-view clustering algorithms. The MultiNMF [9], MVSC-CEV [10] and CMMVC [17] are clustering algorithms for completely (i.e. 100%) mapped multi-view data. The CMVNMF [19] is a clustering algorithm for partially mapped multi-view data. In the last subsection, we set 0% of the mapping relationships, in which we see the proposed method with superior performance. In this subsection, we retain 5% of the mapping relationships to adapt to the CMVNMF. Furthermore, all parameters of the comparison algorithms are set according to the original recommendation. Table 3 and Table 4 show the NMI and ACC results of the comparison methods and the proposed MVC-UM with 0% of the mapping relationships.

It is evident that the new method MVC-UM performs well both in clustering performance and in finding the mapping. Though all of the mapping relationships are known in the MultiNMF, MVSC-CEV and CMMVC algorithms, only the CMMVC algorithm performs well in NMI on the data set Prokaryotic. The CMVNMF with 5% of the mapping relationships known performs better than MVC-UM on Dights and BBCSport. Nevertheless, the MVC-UM with 0% of the mapping relationships known performs best on SensIT, ORL, Wine and Moives617. Although the MVC-UM does not perform best on every dataset, it is the best in terms of overall performance. It is easy to analyze that the NMI of the proposed MVC-UM is averagely 49% more than that of CMVNMF, and the ACC of the proposed MVC-UM is averagely 20% more than that of CMVNMF.

It is shown that the proposed method does not need to take a large number of real mappings in advance while obtaining excellent results, which saves the resources to construct the artificial mapping relationships.

Table 2: NMI and ACC results of different algorithms on different multi-view data with unknown mapping relationships

Datasets	Algorithms	NMI			ACC		
		View 1	View 2	View 3	View 1	View 2	View 3
SensIT	SymNMF [11]	0.23±0.02	0.27±0.00	--	0.60±0.01	0.66±0.01	--
	CMVNMF [19]	0.19±0.02	0.17±0.02	--	0.33±0.05	0.28±0.07	--
	MVC-UM	0.32±0.01	0.30±0.01	--	0.68±0.01	0.68±0.00	--
ORL	SymNMF [11]	0.73±0.01	0.78±0.01	--	0.60±0.02	0.61±0.02	--
	CMVNMF [19]	0.58±0.03	0.61±0.02	--	0.38±0.02	0.37±0.02	--
	MVC-UM	0.82±0.02	0.82±0.01	--	0.67±0.05	0.65±0.04	--
Digits	SymNMF [11]	0.71±0.01	0.72±0.01	--	0.67±0.03	0.70±0.01	--
	CMVNMF [19]	0.66±0.02	0.64±0.03	--	0.64±0.04	0.70±0.04	--
	MVC-UM	0.76±0.03	0.77±0.02	--	0.80±0.02	0.81±0.06	--
BBCSport	SymNMF [11]	0.24±0.02	0.26±0.01	0.17±0.01	0.40±0.02	0.42±0.02	0.39±0.01
	CMVNMF [19]	0.28±0.02	0.31±0.03	0.30±0.03	0.42±0.05	0.45±0.06	0.44±0.05
	MVC-UM	0.18±0.04	0.18±0.04	0.17±0.05	0.44±0.07	0.40±0.08	0.41±0.08
Wine	SymNMF [11]	0.46±0.00	0.47±0.06	--	0.65±0.02	0.62±0.03	--
	CMVNMF [19]	0.06±0.03	0.04±0.03	--	0.44±0.06	0.39±0.08	--
	MVC-UM	0.86±0.05	0.86±0.03	--	0.68±0.05	0.68±0.04	--
Moives617	SymNMF [11]	0.18±0.01	0.09±0.03	--	0.18±0.01	0.10±0.01	--
	CMVNMF [19]	0.11±0.01	0.10±0.02	--	0.10±0.02	0.09±0.02	--
	MVC-UM	0.27±0.01	0.25±0.01	--	0.26±0.01	0.25±0.02	--
Prokaryotic	SymNMF [11]	0.02±0.00	0.15±0.01	--	0.18±0.01	0.54±0.01	--
	CMVNMF [19]	0.07±0.04	0.11±0.04	--	0.40±0.04	0.41±0.05	--
	MVC-UM	0.29±0.00	0.29±0.00	--	0.52±0.00	0.52±0.00	--

Table 3: NMI results of different algorithms on different datasets with (%) mapping relationships

Algorithms (%) \ Datasets	SensIT	ORL	Dights	BBCSport	Wine	Moives617	Prokaryotic
MultiNMF (100%) [9]	0.10±0.01	0.57±0.01	0.70±0.03	0.07±0.00	0.67±0.02	0.08±0.02	0.26±0.01
MVSC-CEV (100%) [10]	0.32±0.00	0.82±0.01	0.83±0.01	0.17±0.00	0.64±0.03	0.23±0.01	0.35±0.00
CMMVC (100%) [17]	0.25±0.03	0.69±0.03	0.87±0.02	0.22±0.04	0.81±0.02	0.26±0.01	0.47±0.01
CMVNMF (5%) [19]	0.29±0.03	0.79±0.01	0.89±0.02	0.58±0.04	0.22±0.06	0.19±0.01	0.16±0.06
MVC-UM (0%)	0.32±0.01	0.82±0.02	0.77±0.02	0.18±0.04	0.86±0.05	0.27±0.01	0.29±0.00

Table 4: ACC results of different algorithms on different datasets with (%) mapping relationships

Algorithms (%) \ Datasets	SensIT	ORL	Dights	BBCSport	Wine	Moives617	Prokaryotic
MultiNMF (100%) [9]	0.10±0.01	0.41±0.02	0.70±0.03	0.35±0.02	0.67±0.02	0.11±0.01	0.65±0.06
MVSC-CEV (100%) [10]	0.67±0.02	0.73±0.00	0.78±0.03	0.43±0.00	0.34±0.01	0.24±0.01	0.59±0.00
CMMVC (100%) [17]	0.67±0.01	0.54±0.04	0.93±0.01	0.53±0.06	0.39±0.03	0.30±0.01	0.71±0.02
CMVNMF 5% [19]	0.43±0.05	0.53±0.01	0.83±0.06	0.57±0.04	0.50±0.11	0.21±0.02	0.43±0.06
MVC-UM (0%)	0.68±0.01	0.67±0.05	0.81±0.06	0.44±0.07	0.68±0.05	0.26±0.01	0.52±0.00

4.5 Convergence Study

This subsection investigates the convergence analysis of MVC-UM by reporting the objective function value and the corresponding ACC and NMI with increasing iteration. Due to space restrictions, we only show the results for ORL in Fig. 7. From the experimental averaging results on the ORL dataset, our objective function values

drop rapidly in the first 16 iterations. After about 300 iterations, the ACC and NMI scores generally converge to a steady-state.

5 CONCLUSIONS

Multi-view clustering has attracted increasing attention in recent years by exploiting common clustering structure across multiple

Figure 7: Convergence analysis on the proposed MVC-UM using the ORL dataset

views. The existing multi-view clustering algorithms assume that they have some prior knowledge of the mapping relationships between cross-view samples. However, this requirement is difficult to meet in many applications. After all, to establish manually the mapping relationships is time-consuming and expensive.

To the best of our knowledge, the proposed MVC-UM method is the first attempt to challenge the problem of multi-view clustering for unknown mapping relationships between cross-view samples. First, the existing multi-view clustering methods have been regrouped into three categories from the perspective of mapping relationships. Then, the novel algorithm was proposed based on the framework of non-negative matrix factorization by exploring the graph-structure consistency for all views. The competitive experimental results show that the proposed MVC-UM not only performs well to reveal the real mapping relationships between cross-view samples but also outperforms the compared algorithms on clustering performances.

This work performs relatively weak for sparse multi-view data, because it is based on graph-structure. This suggests that we should explore different ways to solve this problem. The work is mainly for the unknown simple mapping relationships as shown in Fig. 3(a). Though it can find mapping relationships within views, the accuracy is low. The complex mapping relationship as shown in Fig. 3(b) is common in real life. Developing an efficient method for this topic is another direction for future research.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Linlin Zong for providing the code of the CMVNMf method. This work was jointly supported by the National Natural Science Foundation of China (61876027 and 61751312), and the Natural Science Foundation of Chongqing (cstc2019jcyj-cxttX0002).

REFERENCES

- [1] Steffen Bickel and Tobias Scheffer. 2004. Multi-View Clustering. In *Proceedings of the 4th IEEE International Conference on Data Mining (ICDM 2004)*, 1-4 November 2004, Brighton, UK. IEEE Computer Society, 19–26. <https://doi.org/10.1109/ICDM.2004.10095>
- [2] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. 2004. *Convex optimization*. Cambridge university press.
- [3] Deng Cai, Xiaofei He, Jiawei Han, and Thomas S. Huang. 2011. Graph Regularized Nonnegative Matrix Factorization for Data Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 33, 8 (2011), 1548–1560. <https://doi.org/10.1109/TPAMI.2010.231>
- [4] Xiao Cai, Feiping Nie, and Heng Huang. 2013. Multi-View K-Means Clustering on Big Data. In *IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013*, Francesca Rossi (Ed.). IJCAI/AAAI, 2598–2604. <http://www.aaai.org/ocs/index.php/IJCAI/IJCAI13/paper/view/6979>
- [5] Kamalika Chaudhuri, Sham M. Kakade, Karen Livescu, and Karthik Sridharan. 2009. Multi-view clustering via canonical correlation analysis. In *Proceedings of the 26th Annual International Conference on Machine Learning, ICML 2009, Montreal, Quebec, Canada, June 14-18, 2009 (ACM International Conference Proceeding Series, Vol. 382)*, Andrea Pohoreckýj Danyluk, Léon Bottou, and Michael L. Littman (Eds.). ACM, 129–136. <https://doi.org/10.1145/1553374.1553391>
- [6] Wei Cheng, Xiang Zhang, Zhishan Guo, Yubao Wu, Patrick F. Sullivan, and Wei Wang. 2013. Flexible and robust co-regularized multi-domain graph clustering. In *The 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2013, Chicago, IL, USA, August 11-14, 2013*, Inderjit S. Dhillon, Yehuda Koren, Rayid Ghani, Ted E. Senator, Paul Bradley, Rajesh Parekh, Jingrui He, Robert L. Grossman, and Ramasamy Uthrusamy (Eds.). ACM, 320–328. <https://doi.org/10.1145/2487575.2487582>
- [7] Eric Eaton, Marie desJardins, and Sara Jacob. 2014. Multi-view constrained clustering with an incomplete mapping between views. *Knowl. Inf. Syst.* 38, 1 (2014), 231–257. <https://doi.org/10.1007/s10115-012-0577-7>
- [8] Hongchang Gao, Feiping Nie, Xuelong Li, and Heng Huang. 2015. Multi-view Subspace Clustering. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*. IEEE Computer Society, 4238–4246. <https://doi.org/10.1109/ICCV.2015.482>
- [9] Jing Gao, Jiawei Han, Jialu Liu, and Chi Wang. 2013. Multi-View Clustering via Joint Nonnegative Matrix Factorization. In *Proceedings of the 13th SIAM International Conference on Data Mining, May 2-4, 2013, Austin, Texas, USA*. SIAM, 252–260. <https://doi.org/10.1137/1.9781611972832.28>
- [10] Samir Kanaan-Izquierdo, Andrey Ziyatdinov, and Alexandre Perera-Lluna. 2018. Multiview and multifeature spectral clustering using common eigenvectors. *Pattern Recognit. Lett.* 102 (2018), 30–36. <https://doi.org/10.1016/j.patrec.2017.12.011>
- [11] Da Kuang, Sangwoon Yun, and Haesun Park. 2015. SymNMF: nonnegative low-rank approximation of a similarity matrix for graph clustering. *J. Glob. Optim.* 62, 3 (2015), 545–574. <https://doi.org/10.1007/s10898-014-0247-2>
- [12] Abhishek Kumar and Hal Daumé III. 2011. A Co-training Approach for Multi-view Spectral Clustering. In *Proceedings of the 28th International Conference on Machine Learning, ICML 2011, Bellevue, Washington, USA, June 28 - July 2, 2011*, Lise Getoor and Tobias Scheffer (Eds.). Omnipress, 393–400. https://icml.cc/2011/papers/272_icmlpaper.pdf
- [13] Daniel D. Lee and H. Sebastian Seung. 2000. Algorithms for Non-negative Matrix Factorization. In *Advances in Neural Information Processing Systems 13, Papers from Neural Information Processing Systems (NIPS) 2000, Denver, CO, USA*, Todd K. Leen, Thomas G. Dietterich, and Volker Tresp (Eds.). MIT Press, 556–562. <http://papers.nips.cc/paper/1861-algorithms-for-non-negative-matrix-factorization>
- [14] Huapeng Li, Shuqing Zhang, Xiaohui Ding, Ce Zhang, and Patricia Dale. 2016. Performance Evaluation of Cluster Validity Indices (CVIs) on Multi/Hyperspectral Remote Sensing Datasets. *Remote. Sens.* 8, 4 (2016), 295. <https://doi.org/10.3390/rs8040295>
- [15] Xi Peng, Zhenyu Huang, Jiancheng Lv, Hongyuan Zhu, and Joey Tianyi Zhou. 2019. COMIC: Multi-view Clustering Without Parameter Selection. In *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, 5092–5101. <http://proceedings.mlr.press/v97/peng19a.html>
- [16] Hao Wang, Yan Yang, and Bing Liu. 2020. GMC: Graph-Based Multi-View Clustering. *IEEE Trans. Knowl. Data Eng.* 32, 6 (2020), 1116–1129. <https://doi.org/10.1109/TKDE.2019.2903810>
- [17] Hong Yu, Jing Xiong, and Xiaoxia Zhang. 2020. Multi-view clustering by exploring complex mapping relationship between views. *Pattern Recognit. Lett.* 138 (2020), 230–236. <https://doi.org/10.1016/j.patrec.2020.07.031>
- [18] Changqing Zhang, Huazhu Fu, Si Liu, Guangcan Liu, and Xiaochun Cao. 2015. Low-Rank Tensor Constrained Multiview Subspace Clustering. In *2015 IEEE International Conference on Computer Vision, ICCV 2015, Santiago, Chile, December 7-13, 2015*. IEEE Computer Society, 1582–1590. <https://doi.org/10.1109/ICCV.2015.185>
- [19] Xianchao Zhang, Linlin Zong, Xinyue Liu, and Hong Yu. 2015. Constrained NMF-Based Multi-View Clustering on Unmapped Data. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA*, Blai Bonet and Sven Koenig (Eds.). AAAI Press, 3174–3180. <http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9370>
- [20] Xingwang Zhao, Jiye Liang, and Chuangyin Dang. 2017. Clustering ensemble selection for categorical data based on internal validity indices. *Pattern Recognit.* 69 (2017), 150–168. <https://doi.org/10.1016/j.patcog.2017.04.019>
- [21] Linlin Zong, Xianchao Zhang, and Xinyue Liu. 2018. Multi-view clustering on unmapped data via constrained non-negative matrix factorization. *Neural Networks* 108 (2018), 155–171. <https://doi.org/10.1016/j.neunet.2018.08.011>
- [22] Linlin Zong, Xianchao Zhang, Long Zhao, Hong Yu, and Qianli Zhao. 2017. Multi-view clustering via multi-manifold regularized non-negative matrix factorization. *Neural Networks* 88 (2017), 74–89. <https://doi.org/10.1016/j.neunet.2017.02.003>