

Article

Community Detection in Multiplex Networks Using Orthogonal Non-Negative Matrix Tri-Factorization Based on Graph Regularization and Diversity

Yuqi Yang ¹, Shanshan Yu ², Baicheng Pan ¹, Chenglu Li ¹ and Man-Fai Leung ^{3,*} 

¹ College of Electronic and Information Engineering, Southwest University, Chongqing 400715, China; y15572188780@163.com (Y.Y.); panbaicheng@email.swu.edu.cn (B.P.); chenglulcl@email.swu.edu.cn (C.L.)

² Training and Basic Education Management Office, Southwest University, Chongqing 400715, China; yu33@swu.edu.cn

³ School of Computing and Information Science, Faculty of Science and Engineering, Anglia Ruskin University, Cambridge CB1 1PT, UK

* Correspondence: man-fai.leung@aru.ac.uk

Abstract: In recent years, community detection has received increasing interest. In network analysis, community detection refers to the identification of tightly connected subsets of nodes, which are called “communities” or “groups”, in the network. Non-negative matrix factorization models are often used to solve the problem. Orthogonal non-negative matrix tri-factorization (ONMTF) exhibits significant potential as an approach for community detection within multiplex networks. This paper explores the application of ONMTF in multiplex networks, aiming to detect both shared and exclusive communities simultaneously. The model decomposes each layer within the multiplex network into two low-rank matrices. One matrix corresponds to shared communities across all layers, and the other to unique communities within each layer. Additionally, graph regularization and the diversity of private communities are taken into account in the algorithm. The Hilbert Schmidt Independence Criterion (HSIC) is used to constrain the independence of private communities. The results prove that ONMTF effectively addresses community detection in multiplex networks. It also offers strong interpretability and feature extraction capabilities. Therefore, it is an advanced method for community detection in multiplex networks.

Keywords: community detection; orthogonal non-negative matrix tri-factorization; multiplex networks

MSC: 68R10



Citation: Yang, Y.; Yu, S.; Pan, B.; Li, C.; Leung, M.-F. Community Detection in Multiplex Networks Using Orthogonal Non-Negative Matrix Tri-Factorization Based on Graph Regularization and Diversity. *Mathematics* **2024**, *12*, 1124. <https://doi.org/10.3390/math12081124>

Academic Editors: Laura Cornejo-Bueno and Jorge Pérez-Aracil

Received: 29 February 2024

Revised: 1 April 2024

Accepted: 4 April 2024

Published: 9 April 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

There are many real systems, including social networks, collaboration networks, citation networks, and protein–protein networks, that represent complex networks. These networks express the connections between the same sets of nodes in different types of interactions through different layers. However, most traditional measures merely consider single-layer networks. But they do not consider the diversity of connections between entities. Thus, multiplex networks have been proposed recently to represent various modes of interaction. A multiplex network is a unique form of multilayer network in which every layer contains a distinct topology but the same types of nodes [1]. The most critical challenge of this network is effectively identifying and dividing the structure in the network, also called community detection.

The goal of community detection is to discover the clustering or partition of nodes in the network’s groups of nodes. There is a lot of connectivity between nodes. However, they have a shaky connection to other communities’ nodes. While there is a large body of work on community detection, most of it focuses on single-layer networks. Because they typically cannot handle multiple layers of the network simultaneously, traditional

community detection methods often face challenges when dealing with multiplex networks. Three basic types of community discovery methods are now available for multiplex networks. The first approach directly simplifies the multiplex network into a single graph. After employing a flattening algorithm to merge the layers of a multilayer network into a single graph, traditional community detection algorithms are utilized for detection [2,3]. Nevertheless, the approach can only effectively identify the common communities across each layer of the network. It might also generate some spurious communities due to the flattening process, ultimately impacting the results of community detection. In the second approach, each layer is analyzed layer by layer using standard single-layer network community detection algorithms, with the results eventually combined. The application is referred to as Principal Modularity Maximization (PMM) and Spectral Clustering on Multilayer graphs (SC-ML) [4]. The third approach involves directly operating on the multiplex network model. Based on this category of methods, several algorithms have been proposed, including Locally Adaptive Random Transitions (LART) [5], Infomap [6], and multilayer community quality measure optimization-based techniques [7–9].

The existing community detection methods based on multiplex networks typically select a partition suitable for all layers [10]. However, no method simultaneously considers the shared communities across different layers and the unique communities existing in each layer. Therefore, for real-world complex networks, existing community detection algorithms often exhibit poor performance. The reason behind this is that, in real social networks, different layers often represent distinct interaction patterns, and each layer's network is heterogeneous. For instance, in real social networks, students of a class might interact on different social platforms like WeChat, QQ, and Facebook. Each layer of the network corresponds to a specific social platform. On different platforms, each student may connect with different individuals. Hence, for a class of students, there may be common communities across each social platform. In the meantime, each layer possesses distinct communities. Based on the above, the goal is to address the challenge of simultaneously identifying both common and private communities within multiplex networks. Specific constraints are introduced for detected shared and unique communities, aiming to enhance the algorithm's performance.

Non-negative matrix factorization (NMF) is a widely utilized technique in data mining and machine learning, exhibiting particularly unique advantages when dealing with high-dimensional data. In recent years, NMF has been successfully applied in various domains, including image processing, text mining, and bioinformatics. In the field of network science, particularly in community detection within multiplex networks, NMF has shown tremendous potential. In fact, NMF has been widely employed in community identification for single-layer networks, multiplex networks, multilayer networks, and dynamic networks because of its high interpretability and effect [11–13]. In a previous paper, an orthogonal non-negative matrix tri-factorization [14] model was proposed to detect each layer of the network separately. The proposed model decomposes each layer of the multiplex network into two low-rank matrices. The first component relates to the communities that are shared by all levels. Additionally, the second relates to the exclusive groups that are found inside each stratum. Compared to standard non-negative matrix factorization, orthogonal non-negative matrix tri-factorization decomposes the matrix into three independent parts, thereby offering stronger interpretability. Additionally, it can better extract features from the data, accurately capturing structural information. It considers additional orthogonal constraints, avoiding the overfitting issues commonly found in NMF and ensuring a more stable decomposition. It demonstrates better performance and applicability in multiplex networks compared to traditional non-negative matrix factorization.

Based on the above findings, this study introduces graph regularization [15] and Hilbert–Schmidt Independence Criterion (HSIC) [16] terms into the orthogonal non-negative matrix tri-factorization model. We calculate common communities separately for each layer and then synthesize the final common communities by incorporating weights. The contributions of the paper are as follows:

1. An orthogonal non-negative matrix tri-factorization model is applied to multiplex networks. It enables the simultaneous detection of common and private communities within the network. The model has enhanced interpretability and feature extraction capabilities.
2. For the detected private communities, a graph regularization [15,17] constraint is added. Graph regularization utilizes the topological structure of networks to better capture relationships and connection patterns between nodes. It constrains the formation of community structures. And it aligns more with the actual community partitioning rules in real networks. Finally, it enhances the accuracy and robustness of community detection.
3. The Hilbert–Schmidt Independence Criterion (HSIC) [16] term is introduced. Because the private communities in each layer are independent of each other with minimal correlation, an HSIC term is added to the detected private communities [18]. The HSIC effectively imposes independence constraints on the private communities in each layer.
4. Weight constraints are applied to the identified shared communities in each layer [19]. Ultimately, the shared communities across all layers are summed, enhancing the accuracy of the detected common communities.

The remaining sections of this paper are structured as follows. The second section describes related work. Section 3 introduces some notations and definitions. The proposed algorithm is introduced in Section 4. Section 5 describes the updating process of the algorithm. In Section 6, various datasets and comparison models are introduced in detail. Then, the experimental results are analyzed. In Section 7, conclusions and further research are covered.

2. Related Works

2.1. Existing Methods

The method proposed in this paper directly operates on multiplex network models. This category of algorithms includes various types: methods based on random walks, statistical generative network models, label propagation, objective function optimization, and non-negative matrix factorization (NMF) methods.

The community detection method based on random walkers is a commonly used network analysis technique. It involves simulating the behavior of random walkers in the network to discover community structures. The basic idea is to start from a node in the network, randomly select a neighboring node to move to, and repeat this process until certain conditions are met. By simulating the behavior of random walkers multiple times, the degree of association between nodes can be obtained, finally identifying community structures. The authors of [5] proposed LART, a method based on a random walk on a multiplex network. The community detection method based on random walkers has advantages such as simple implementation, suitability for large-scale networks, and the ability to detect overlapping communities. However, it needs various parameters. The selection of parameters in the algorithm may affect the final community detection results. When processing large-scale networks, it requires a considerable amount of computational resources and time.

Statistical methods such as the weighted stochastic block model (WSBM) detect shared and unique communities in heterogeneous weighted networks. While this approach tackles the problem of network heterogeneity across layers, it overlooks communities that might be shared by specific subsets or different combinations of layers. By considering the weights of edges in the network, the weighted stochastic block model can capture more nuanced relationships between nodes compared to traditional binary network models. However, for large networks with a high number of nodes and edges, the weighted stochastic block model can be computationally intensive.

The Label Propagation Algorithm (LPA) is based on the idea of label propagation. The basic principle of the LPA is to propagate initial labels (or communities) in the network

until the nodes in the network reach a stable state. Firstly, each node is initially assigned a unique label (or community). Then, labels propagate in the network, with each node adopting the most common label among its current neighboring nodes. When the labels of all nodes no longer change, the algorithm reaches a stable state. Finally, nodes with the same label are considered to belong to the same community. In general, the LPA is suitable for community detection in large-scale networks. But it may not accurately identify overlapping communities or communities with clear hierarchical structures in some cases.

In community detection for multilayer networks, methods based on optimizing objective functions are commonly used to identify communities that span across different layers of the network. These methods aim to find the partition of nodes into communities that maximizes/minimizes a specific objective function, which captures the quality of the community structure. One popular approach for community detection in multilayer networks is to extend traditional modularity optimization to multilayer settings. Another method for optimizing objective functions in multilayer community detection is to use a variation of the stochastic block model (SBM). The SBM is a generative model that assumes nodes belong to different blocks (communities). And it can define the probability of edges between nodes based on their block assignments. In the multilayer setting, the SBM is extended to model interactions within and between layers. It succeeds in detecting communities that exist across layers.

The majority of the above-mentioned methods only consider the shared communities across all or most layers. They all overlook the possibility of shared communities existing in subsets of different layers. They also do not detect the unique communities. In methods based on NMF, the number of communities is also provided through prior information.

2.2. Multiplex Networks

A multiplex network is a special multilayer network in which each layer has the same nodes, but its network topology is different. Like single-layer networks, a finite graph sequence $\{G_l\}$ can also be used to depict multiplex networks, where l represents the number of layers, $l \in \{1, 2, \dots, L\}$, $G_l = (V_l, E_l, A_l)$ [20]. The collection of layer l nodes is denoted by V_l , and $A_l \in \mathbb{R}^{n \times n}$ is the adjacency matrix for layer l , where A_{ij} characterizes the relationship between v_i and v_j . In this paper, binary and undirected weighted adjacency matrices are employed. For both the weighted adjacency matrix and the binary adjacency matrix, $A_{ij} \in [0, 1]$. We have $A_{ij} = 1$ if $e_{ij} \in E$ and $A_{ij} = 0$ otherwise in each layer of the network.

2.3. Hilbert–Schmidt Independence Criterion (HSIC)

To determine the degree of independence between two random variables, statisticians employ the Hilbert–Schmidt Independence Criterion (HSIC). It is based on kernel methods in Hilbert–Schmidt space and was introduced by Gretton et al. [21]. The primary application domains of the HSIC include machine learning, statistical, and kernel methods. The HSIC quantifies the correlation between two variables by comparing their representations in the feature space. Each of the detected private communities is independent of the others, meaning their topological structures are entirely different from each other. Hence, the algorithm proposes the following version of the HSIC as a measure of independence between two private communities:

$$HSIC(\mathbf{H}_{p_l}, \mathbf{H}_{p_m}) = \frac{1}{(N-1)^2} \text{Tr}(\mathbf{K}_l \mathbf{H} \mathbf{K}_m \mathbf{H}) \quad (1)$$

where \mathbf{H}_{p_l} and \mathbf{H}_{p_m} represent the detected private communities in the l -th and m -th layers, respectively. And $\mathbf{H} = \mathbf{I} - 1/n$, and \mathbf{I} is the n -dimensional identity matrix, where n is the dimension of the matrix. The larger the value of the HSIC, the stronger the correlation between X and Y ; conversely, a smaller HSIC indicates a weaker correlation. The HSIC equals zero only when X and Y are independent. For the kernels \mathbf{K}_l and \mathbf{K}_m , the inner

product function is chosen, i.e., $K_l = (\mathbf{H}_{pl})^T \mathbf{H}_{pl}$, $K_m = (\mathbf{H}_{pm})^T \mathbf{H}_{pm}$. Substituting it and the expression for \mathbf{H}_{pl} and \mathbf{H}_{pm} into Equation (1), the following formulation is obtained:

$$HSIC(\mathbf{H}_{pl}, \mathbf{H}_{pm}) = \frac{1}{(N-1)^2} \text{Tr} \left(\mathbf{H}_{pl}^T \mathbf{H}_{pl} \mathbf{H} \mathbf{H}^T \mathbf{H}_{pm}^T \mathbf{H}_{pm} \mathbf{H} \right). \quad (2)$$

2.4. Community Detection with NMF in Multiplex Networks

Non-negative matrix factorization (NMF) [22] is an extensively employed algorithm in community detection. The model has been widely applied in many fields [23]. A non-negative matrix can be broken down using NMF into the product of two or more other non-negative matrices, thereby capturing the latent structures within the data. In the context of community detection, NMF has been successfully utilized to uncover potential modular structures and is particularly well suited for representing the non-negative characteristics inherent in the data. Given a non-negative matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$, NMF attempts to find two non-negative matrices, $\mathbf{X} \in \mathbb{R}^{n \times k}$ and $\mathbf{U} \in \mathbb{R}^{m \times k}$, such that $\mathbf{A} \approx \mathbf{X}\mathbf{U}^T$ and $k \ll n, m$. \mathbf{X} and \mathbf{U} are found by solving the following optimization problem:

$$\min_{\mathbf{X}, \mathbf{U}} \left\| \mathbf{A} - \mathbf{X}\mathbf{U}^T \right\|_F^2 \text{ s.t. } \mathbf{X}, \mathbf{U} \geq 0. \quad (3)$$

The community feature matrix (\mathbf{X}) and community indicator matrix (\mathbf{U}) are both community detection techniques based on NMF. The algorithm adds the orthogonality constraint to both factor matrices, $\mathbf{X}\mathbf{X}^T = \mathbf{I}$ and $\mathbf{U}\mathbf{U}^T = \mathbf{I}$, resulting in better performance. Each row in \mathbf{W} or \mathbf{U} has only one non-zero element due to the requirements of orthogonality and non-negativity, which suggests that each node is only a member of one community. The decomposed matrix \mathbf{X} is an adjacency matrix. Typically, the number of communities is expressed as k [24,25].

The algorithm research on community detection in single-layer networks has been improved, and various non-negative matrix decomposition algorithms based on single-layer networks have been widely applied and achieved good performance [26]. However, in the research of multilayer networks, the non-negative matrix factorization algorithm has not been widely applied. The main reason is that the traditional non-negative matrix factorization algorithm cannot detect the relationship between communities in a multilayer network. It only applies to the single-layer network. For community detection in two-layer and more-than-two-layer networks, there are two basic models: one is a multilayer network, and the other is a multiplex network [27]. In the multilayer network, there are interlayer connections between the network nodes of the adjacent layers. In the multiplex network, each layer has the same type of entity, and there are no node connections between adjacent layers.

For multilayer networks, an existing model is Community Detection in Fully Connected Multilayer Networks Through Joint Non-negative Matrix Factorization [28,29]. The approach views a multilayer network as a hybrid of a bipartite graph network and a multiplex network. In the network, there are intralayer edges of a multiplex network and interlayer edges of a bipartite graph network, so an intralayer adjacency matrix and an interlayer adjacency matrix are introduced in this method. And considering the connectivity between intralayer communities and interlayer communities, its objective function is as follows:

$$\begin{aligned}
 \min \mathcal{J} = & \min_{\substack{\mathbf{B}_\alpha \in \mathbb{R}^{n^\alpha \times r_\alpha}, \mathbf{U}_\alpha \in \mathbb{R}^{n^\alpha \times r_\alpha}, \\ \mathbf{S}^{12} \in \mathbb{R}^{r_{\alpha\beta} \times r_{\alpha\beta}}} \underbrace{\sum_{\alpha=1}^2 \|\mathbf{A}^\alpha - \mathbf{U}_\alpha \mathbf{U}_\alpha^\top\|_{2,1}}_{\text{Intralayer low-rank approximation}} \\
 & + \mu_1 \underbrace{\|\mathbf{A}^{12} - \mathbf{B}_1 \mathbf{S}^{12} \mathbf{B}_2^\top\|_{2,1}}_{\text{Interlayer low-rank approximation}} \\
 & + \mu_2 \underbrace{\sum_{\alpha=1}^2 \|\mathbf{U}_\alpha \mathbf{U}_\alpha^\top - \mathbf{B}_\alpha \mathbf{B}_\alpha^\top\|_F^2}_{\text{Pairwise Similarity}} \\
 \text{s.t. } & \mathbf{U}_\alpha \geq 0, \quad \mathbf{B}_\alpha \geq 0, \quad \mathbf{S}^{12} \geq 0,
 \end{aligned} \tag{4}$$

where the coefficient matrices are \mathbf{B}_α and \mathbf{U}_α . The first one corresponds to intralayer adjacency matrices. The second one represents interlayer adjacency matrices. The \mathbf{A}^{12} low-rank embedding is denoted by \mathbf{S}^{12} .

Ref. [30] proposed an orthogonal non-negative matrix tri-factorization approach for community discovery in multiplex networks. The algorithm decomposes the adjacency matrix \mathbf{A} of each layer into two terms, each of which is an orthogonal non-negative matrix tri-factorization. The first item represents the common community of all levels, and the second item represents the private community of each level. The algorithm is as follows:

$$\begin{aligned}
 \min_{\mathbf{H} \geq 0, \mathbf{H}_l \geq 0, \mathbf{S}_l \geq 0, \mathbf{G}_l \geq 0} & \sum_{l=1}^L \|\mathbf{A}_l - \mathbf{H} \mathbf{S}_l \mathbf{H}^\top - \mathbf{H}_l \mathbf{G}_l \mathbf{H}_l^\top\|_F^2 \\
 \text{s.t. } & \mathbf{H}^\top \mathbf{H} = \mathbf{I}, \mathbf{H}_l^\top \mathbf{H}_l = \mathbf{I}, l \in \{1, 2, \dots, L\}
 \end{aligned} \tag{5}$$

assuming that $l \in \{1, 2, \dots, L\}$. The community membership matrices are $\mathbf{H} \in \mathbb{R}^{m \times k_c}$ and $\mathbf{H}_l \in \mathbb{R}^{m \times k_{pl}}$. One represents a common community, and the other corresponds to a private community. \mathbf{S}_l and \mathbf{G}_l are symmetric matrices. This method performs community detection on a multiplex network with L layers, where the number of private communities in each layer is denoted by k_{pl} . And the number of common communities is indicated by k_c .

Multilayer network connection is taken into account in the model presented in [29]. However, the approach is limited to networks with two layers. When the number of layers in the network increases, the adjacency matrix between the adjacent layers also increases. This will lead to the model becoming more and more complex. For a multiplex network, the model in [30] only divides the community in a multiplex network into common communities and private communities. However, it does not perform a separate study of the separate common and private communities.

3. Notations

Some notations are described in Table 1.

Table 1. Notations and definitions.

Notations	Meaning
\mathbf{A}_l	Layer l 's adjacency matrix
\mathbf{H}_{cl}	The shared community membership matrix of layer l
\mathbf{H}_{pl}	The private community membership matrix of layer l
\mathbf{H}_c	The shared community member matrix integrating the shared community member matrix of each layer
\mathbf{S}_l	The symmetric matrix of the shared community membership matrix

Table 1. Cont.

Notations	Meaning
G_l	The symmetric matrix of the private community membership matrix
α_l	The weight factor added to the shared community membership matrix for every stratum
$Tr(\cdot)$	The traces of a matrix

4. The Proposed Method

ONMTF breaks down the adjacency matrix of each layer A_l into the sum of the low-rank representations of exclusive communities and shared communities. Then, considering that the common communities of all layers are the same, the shared communities obtained in each layer are summed by weight. The private communities are independent of each other, and the independence between any two private communities is constrained by the HSIC. Finally, graph regularization is added for each layer’s private community.

For a given multiplex network $G_l = (V_l, E_l, A_l)$ [20] with l layers, the adjacency matrix for each layer is A_l . One way to formulate the resultant objective function is as follows:

$$\begin{aligned}
 & \min_{H_{cl} \geq 0, H_{pl} \geq 0, S_l \geq 0, G_l \geq 0} \sum_{l=1}^L \|A_l - H_{cl}S_lH_{cl}^T - H_{pl}G_lH_{pl}^T\|_F^2 \\
 & + \mu_1\alpha_l Tr(H_{cl}H_{cl}^T) + \sum_{l \neq m} \beta HSIC(H_{pl}, H_{pm}) + \lambda Tr(H_{pl}^T L H_{pl}), \tag{6} \\
 & \text{s.t. } H_{cl}^T H_{cl} = I, H_{pl}^T H_{pl} = I, l \in \{1, 2, \dots, L\}
 \end{aligned}$$

where $H_{cl} \in \mathbb{R}^{m \times k_c}$ and $H_{pl} \in \mathbb{R}^{m \times k_{pl}}$, $l \in \{1, 2, \dots, L\}$ are the community membership matrices. One represents a shared community, and the other corresponds to an exclusive community. S_l and G_l are symmetric matrices. The second term is used to sum the weights of the common communities of all the obtained layers, where α_l is the weight factor. The predefined private matrices are not related to each other, and the third part uses the HSIC to impose correlation constraints on the private matrices of each layer. λ denotes the regularization constant, and L is the Laplacian matrix. For each layer, $L = D - W$, where D is a diagonal matrix that is defined by $D_{ii} = \sum_l W_{il}$, and W is a weight matrix. The Laplacian matrix L depends on the definition of W . Should nodes i and j be linked, $W_{ij} = 1$. Otherwise, $W_{ij} = 0$. We have $W = A$. For each layer, due to the differences in the adjacent matrix, we have $L = D - A_l$.

5. Optimization

To resolve the updating rules for H_{cl} , H_{pl} , H_c , S_l , and G_l , Lagrange multipliers Λ and Λ_l are introduced. The Lagrangian function is minimized:

$$\begin{aligned}
 \mathcal{L}(H_{cl}, H_{pl}, S_l, G_l) &= \sum_{l=1}^L \|A_l - H_{cl}S_lH_{cl}^T - H_{pl}G_lH_{pl}^T\|_F^2 \\
 &+ \sum_{l=1}^L \text{tr}(\Lambda(H_{cl}^T H_{cl} - I)) + \sum_{l=1}^L \text{tr}(\Lambda_l(H_{pl}^T H_{pl} - I)) \tag{7} \\
 &+ \mu_1\alpha_l Tr(H_{cl}H_{cl}^T) + \beta \sum_{l \neq m} HSIC(H_{pl}, H_{pm}) + \lambda Tr(H_{pl}^T L H_{pl}).
 \end{aligned}$$

$$\alpha_l = \frac{1}{2\sqrt{Tr(H_{cl}H_{cl}^T) + \Gamma}}, \Gamma = \mu_2 Tr(H_{cl}^T L H_{cl}). \tag{8}$$

To update H_{cl} , $\nabla_{H_{cl}} \mathcal{L}$ is found as follows:

$$\nabla_{H_{cl}} \mathcal{L} = \sum_{l=1}^L (4H_{cl}S_l^T H_{cl}^T H_{cl}S_l + 4H_{pl}G_l^T H_{pl}^T H_{cl}S_l - 4A_l H_{cl}S_l) + 4H_{cl}\Lambda + 4\mu_1\alpha_l H_{cl}. \tag{9}$$

$\nabla_{H_{cl}} \mathcal{L} = 0$ and $\nabla_{\Lambda} \mathcal{L} = 0$ are applied. We acquire

$$\Lambda = \sum_{l=1}^L (-S_l^T S_l - H_{cl}^T H_{pl} G_l^T H_{pl}^T H_{cl} S_l + H_{cl}^T A_l H_{cl} S_l). \tag{10}$$

$$H_{cl}^T H_{cl} = I. \tag{11}$$

Substituting (10) and (11) into (9), we obtain

$$\begin{aligned} \nabla_{H_{cl}} \mathcal{L} = & \sum_{l=1}^L (4H_{pl}G_l^T H_{pl}^T H_{cl}S_l - 4A_l H_{cl}S_l + 4H_{cl}H_{cl}^T A_l H_{cl}S_l \\ & - 4H_{cl}H_{cl}^T H_{pl}G_l^T H_{pl}^T H_{cl}S_l) + 4\mu_1\alpha_l H_{cl}. \end{aligned} \tag{12}$$

Finally, the updated rule for H_{cl} is obtained:

$$H_{cl} \leftarrow H_{cl} * \frac{A_l H_{cl} S_l + H_{cl} H_{cl}^T H_{pl} G_l^T H_{pl}^T H_{cl} S_l}{H_{pl} G_l^T H_{pl}^T H_{cl} S_l + H_{cl} H_{cl}^T A_l H_{cl} S_l + \mu_1 \alpha_l H_{cl}}, \tag{13}$$

For every $l \in \{1, 2, \dots, L\}$, we also derive the update rules for H_{pl} , S_l , and G_l :

$$H_{pl} \leftarrow H_{pl} * \frac{A_l H_{pl} G_l + H_{pl} H_{pl}^T H_{cl} S_l H_{cl}^T H_{pl} G_l^T}{H_{pl} H_{pl}^T A_l H_{pl} G_l + H_{cl} S_l H_{cl}^T H_{pl} G_l^T + \lambda L H_{pl} + \beta H_{pl} H H_{pm}^T H_{pm} H + \lambda L H_{pl}}, \tag{14}$$

$$S_l \leftarrow S_l * \frac{H_{cl}^T A_l H_{cl}}{H_{cl}^T H_{cl} S_l H_{cl}^T H_{cl} + H_{cl}^T H_{pl} G_l H_{pl}^T H_{cl}}, \tag{15}$$

$$G_l \leftarrow G_l * \frac{H_{pl}^T A_l H_{pl}}{H_{pl}^T H_{pl} G_l H_{pl}^T H_{pl} + H_{pl}^T H_{cl} S_l H_{cl}^T H_{pl}}, \tag{16}$$

The updated H_{cl} of each layer is fused with the corresponding weights, and finally, the common community member matrix H_c is obtained:

$$H_c = \sum_{l=1}^L \frac{\alpha_l H_{cl}}{\sum_{l=1}^L \alpha_l}. \tag{17}$$

Since NMF algorithms are initialized with random matrices and different runs may return different results, we repeat the algorithm 100 times. As shown in Algorithm 1, for each random initialization of H_{cl} , H_{pl} , S_l and G_l , they all follow the update rules described in the corresponding formula. Finally, we choose the maximum performance value of different running results for calculation. According to the NMI value, the difference of community detection effect is judged.

Algorithm 1 MX-ONMTF based on graph regularization and diversity.

Require: $A_l, k_c, k_{pl}, l \in \{1, 2, \dots, L\}$;
 1: **for** $r = 1$ to 100 **do**
 2: Randomly initialize $H_{cl}, H_{pl}, S_l, G_l > 0$;
 3: **for** 1000 iterations or until convergence **do**
 4: for every $l \in \{1, 2, \dots, L\}$, update H_{cl} using Equation (13)
 5: for every $l \in \{1, 2, \dots, L\}$, update H_{pl} using Equation (14)
 6: for every $l \in \{1, 2, \dots, L\}$, update S_l using Equation (15)
 7: for every $l \in \{1, 2, \dots, L\}$, update G_l using Equation (16)
 8: for every $l \in \{1, 2, \dots, L\}$ update α_l using Equation (8)
 9: update H_c according to Equation (17)
 10: **end for**
 11: Compute NMI .
 12: **end for**

6. Experiments

6.1. Datasets

Real-World Multiplex Network:

Lazega Law Firm Multiplex Social Network [31]: The Lazega Law Firm is a complex social network consisting of 71 nodes. It has three layers. Each layer represents different types of relationships within the firm. Additionally, the dataset includes various attributes for each node. There is no ground-truth community in this network, but the detected community structure and node attributes of each type can be used to calculate the NMI and analyze the performance of the community detection by this metric. The calculation method of NMI comes from [8]. For each attribute, the nodes are divided into different communities based on specific attributes.

3sources: The 3sources dataset comprises data from three distinct network layers (BBC, Reuters, and *The Guardian*) that represent different types of relationships or interactions among entities. It provides valuable insights for studying complex networks [32].

BBCSport: The BBCSport dataset is a collection of data that encompass various sports articles published by the British Broadcasting Corporation (BBC). It includes articles covering a wide range of sports, such as football, cricket, rugby, and tennis [32].

Wikipedia: The Wikipedia dataset is a comprehensive collection of data extracted from the Wikipedia website. The Wikipedia website is a vast online encyclopedia covering a wide range of topics in multiple languages. It includes articles, images, metadata, and other types of content contributed by users from around the world. The dataset provides valuable resources for various research tasks, including natural language processing (NLP), information retrieval, knowledge extraction, and data mining [33].

Table 2 shows the basic information of the three datasets. For each dataset, N is the number of nodes, k is the number of layers, and c is the number of communities. The last line represents the size of each community.

Table 2. The information of real-word datasets.

	3Sources	BBCSport	Wikipedia
N	169	544	693
k	3	2	2
c	6	5	10
$ C_i $	54, 21, 11, 18, 51, 12	62, 104, 193, 124, 61	34, 88, 96, 85, 65, 58, 51, 41, 71, 104

Benchmark Multiplex Networks:

The generated multiplex networks based on the model in [34] are used. It suggests creating multilayer networks with community structures in two steps. First, it is necessary to manually define the parameters in the multilayer network, including the number of

layers, the number of nodes, and an interlayer dependency tensor. The interlayer dependency tensor outlines the interlayer dependency structure. Then, there will be a partition of the multilayer network. Next, according to a degree-corrected block model, edges are generated within each layer. They are generated by constraints on the distribution of expected degrees and the community mixing parameter $\mu \in [0, 1]$. The modularity of the network is governed by the mixing parameter μ . When $\mu = 0$, all edges are contained in communities. Therefore, the closer μ is to 1, the denser the distribution of nodes in the community, and the closer μ is to 0, the higher the independence of edges. For multiplex networks, the interlayer dependency tensors of all layers are uniform. Their range of values is from 0 to 1. When $p = 0$, the partitions of all layers are independent. $p = 1$ indicates identical cross-layer partitions.

6.2. Experimental Settings

6.2.1. Comparison Algorithm Models

In this study, five comparison models were used for community detection in multiplex networks. The methods include Generalized Louvain (GL), Co-Regularized Spectral Clustering (CoReg), Multi-view clustering via Adaptively Weighted Procrustes (AWP), and Multi-view Consensus Graph Clustering (MCGC).

MX-ONMTF [30] operates by simultaneously factorizing multiple layers of the network data into three non-negative matrices. A shared basis matrix captures common structural patterns across layers. Two layer-specific coefficient matrices reflect the participation of nodes in each layer's communities. By enforcing orthogonality constraints on the shared basis matrix, MX-ONMTF effectively disentangles the intertwined community structures present in multiplex networks.

Generalized Louvain (GL) [35] is an algorithm designed for community detection in complex networks. The core idea of GL is to iteratively optimize a quality function that measures the modularity of the network partitioning. GL efficiently identifies communities that exhibit strong internal connections. And it allows nodes to belong to multiple communities simultaneously. Overall, GL offers a flexible and effective approach for detecting communities in a wide range of network structures.

Co-Regularized Spectral Clustering (CoReg) [36] is a clustering algorithm designed to handle data with multiple views or modalities. Co-Regularized Spectral Clustering jointly clusters data from multiple views while leveraging the shared information across views.

Multi-view clustering via Adaptively Weighted Procrustes (AWP) [37] is an innovative clustering algorithm. AWP tackles the challenge of integrating information from diverse views by employing an Adaptively Weighted Procrustes analysis.

Multi-view Consensus Graph Clustering (MCGC) [38] is an algorithm designed to cluster data that come from multiple sources or views. Differing from traditional clustering methods, MCGC combines information from different views by constructing a consensus graph that captures the common structure across views.

Notably, all of these algorithms, except for GL and MX-ONMTF, require the user to specify the number of communities to look for a priori. It is usually denoted by k . This is a potential drawback in practice, as we usually do not have information about the community structure of the graph and would have to make some (possibly unjustified) assumptions about the number of clusters.

In experiments, the co-regularization parameter λ in CoReg [36] is set to 0.01. In AWP [37], the number of neighbors used in graph construction is fixed to 20. For MCGC [38], there is one parameter β in the objective function. We fix $\beta = 0.6$ for all the datasets. For MX-ONMTF [30], there are no parameters in the algorithm. The community can be detected directly.

6.2.2. Evaluation Metrics

The performance of community discovery techniques is analyzed in this study using Normalized Mutual Information (NMI). Let the ground-truth community label set be

represented by C_g . The one that a detector predicts is shown by C_r . Ultimately, the NMI is ascertained:

$$NMI_{C_g, C_r} = \frac{\sum_{i=1}^K \sum_{j=1}^K n_{ij} \log \frac{n_{ij}n}{n_i^{(1)}n_j^{(2)}}}{\sqrt{\left(\sum_{i=1}^K n_i^{(1)} \log \frac{n_i^{(1)}}{n}\right) \left(\sum_{j=1}^K n_j^{(2)} \log \frac{n_j^{(2)}}{n}\right)}} \quad (18)$$

where n represents the number of nodes, and K is the number of communities. n_{ij} is the number of nodes that are assigned to community j by a detector, but n_{ij} actually belongs to community i . $n_i^{(1)}$ is the node count in the ground-truth community i . $n_j^{(2)}$ is the number of nodes. It is allocated to community j by a detector. The larger the NMI value, the better the performance of the community detector [24].

6.3. Analysis

The proposed algorithm is first used to perform community detection on the generated multiplex reference network. Each layer of the resulting multiplex network has common communities. The proposed model is first used to detect two-layer generative networks. The interlayer dependence tensor p is fixed to 1. In cases where nodes = 64 and nodes = 128, respectively, the value of the mixed parameter μ increases from 0.1 to 0.8. The NMI values corresponding to different μ values are shown in Figure 1. When the mixing parameter μ is closer to 0, the NMI value is closer to 1. The results show that the algorithm has good performance for community detection under certain conditions. As the μ value increases, the corresponding NMI in both cases decreases.

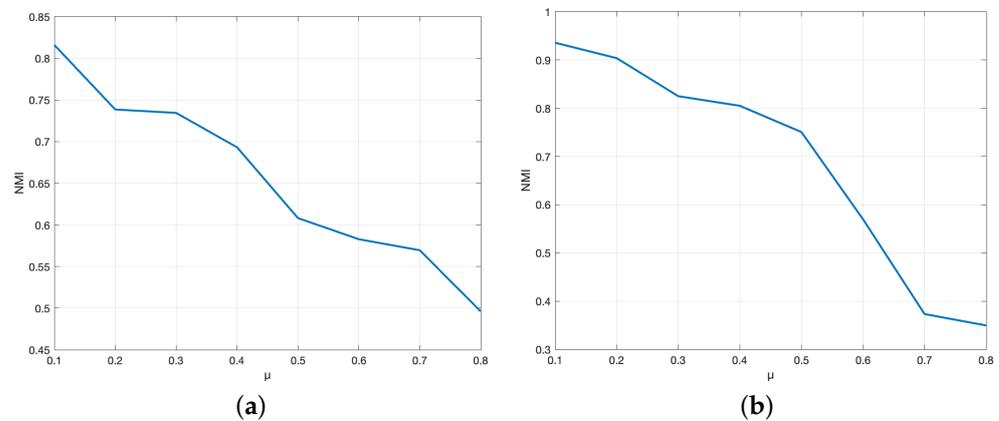


Figure 1. When nodes = 64 and nodes = 128, respectively, (a,b) depict the NMI corresponding to different μ values with $p = 1$ and $L = 2$.

Figure 2 shows the variation trend of the corresponding NMI value with μ in the cases of three-layer, four-layer, and five-layer generation networks. In all three cases, the NMI values decrease significantly with the increase in μ value. This shows that the effect of the mixing parameter μ on the experimental results is independent of the number of layers in the generated network. The smaller the mixing parameter, the more obvious the community division in the generating network. Therefore, the model can detect the generated network with small mixing parameters better.

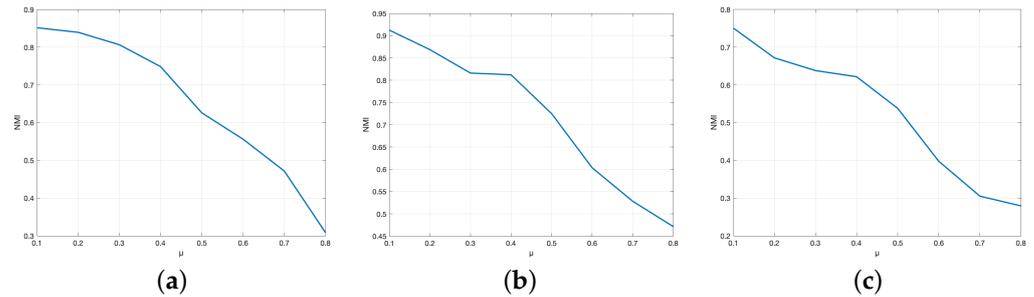


Figure 2. The graphs in (a–c) depict the NMI corresponding to different μ values in three-layer, four-layer, and five-layer networks, with $p = 1$, nodes = 128.

In order to investigate the impact of the interlayer dependency tensor on the detection outcomes, the number of produced network layers is kept at two. Then, the mixing parameter μ is set to 0.1, and the number of nodes is set to 64. The interlayer dependence tensor gradually increases from 0.3 to 1. The variation trend of the NMI value can be clearly seen in Figure 3.

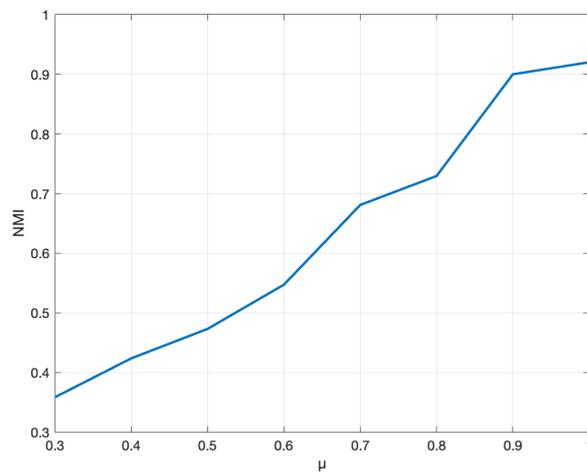


Figure 3. When $\mu = 0.1$ and nodes = 64, NMI corresponds to different p values in 2-layer network.

In Figure 3, when the interlayer dependence tensor p is closer to 1, the corresponding NMI value becomes larger and larger. This shows that the interlayer dependency tensor controls the similarity of the common communities in different layers. The greater the interlayer dependence tensor, the higher the degree of community integration between different layers. In this case, the algorithm is more accurate in detecting the shared communities and exclusive communities in the multiplex network.

In Figure 4, the interlayer dependence tensor is fixed at 1. The mixing parameter μ of the generated network is set to 0.3. In the case of a three-layer network, the number of nodes is gradually increased. When the number of nodes in the generated network increases, the corresponding NMI value remains basically unchanged. The experimental results show that the number of nodes in the generated network has little effect on the detection performance of the model. Therefore, the proposed algorithm shows excellent performance in both small- and large-scale networks. For the Lazega Law Firm Multiplex Social Network, the partitions detected by the proposed algorithm have a good effect on each property. The proposed algorithm performs better than most algorithms in the detection of this network.

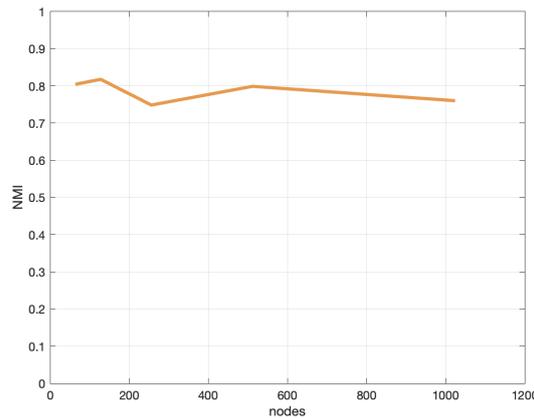


Figure 4. When $p = 1$ and $\mu = 0.3$, NMI corresponds to different values of nodes in 3-layer network.

In Table 3, the proposed algorithm outperforms other algorithms on both the Status and Wikipedia datasets. Under both the Gender and Law School datasets, the proposed algorithm performs second only to MX-ONMTF. On several other datasets, the proposed algorithm is unable to outperform the comparison algorithms. The GL model performs well on the 3sources and BBCSport datasets, while the CoReg model performs best on Seniority and Age. The experimental results show that no model performs optimally on several datasets at the same time. The AWP and MCGC models do not outperform other models on any of the datasets.

Table 3. NMI values for datasets based on six models.

Dataset	GL	CoReg	AWP	MCGC	MX-ONMTF	MX-GONMTF
Status	0.0350	0.0468	0.0115	0.0252	0.3902	0.4264
Gender	0.0312	0.0019	0.0042	0.0129	0.4100	0.1126
Office	0.5080	0.5828	0.3235	0.1029	0.7297	0.4216
Seniority	0.0727	0.6946	0.6368	NULL	0.2192	0.3565
Age	0.0395	0.7394	0.6997	NULL	0.3024	0.3363
Practice	0.5515	0.0676	0.1857	0.0207	0.5610	0.1143
Law School	0.0163	0.0051	0.0070	0.0444	0.3308	0.1249
3sources	0.7488	0.6584	0.6619	0.5948	0.5265	0.5414
BBCSport	0.7525	0.6168	0.7223	0.7951	0.5347	0.5082
Wikipedia	0.5435	0.4338	0.3249	0.1372	0.3912	0.5573

7. Conclusions

This study presents a multiplex network community discovery technique based on ONMTF. The suggested technique can identify shared and unique communities spread across many network tiers. An independence constraint and a graph regularity constraint are added to the original model. The results show that for synthetic networks and real-world networks, the algorithm can detect shared communities and exclusive communities with different layer community structures and performs well on some datasets. This is especially important for further exploring real networks with heterogeneity in cross-layer relationships.

Author Contributions: Conceptualization, Y.Y., S.Y. and B.P.; Methodology, Y.Y., S.Y. and C.L.; Software, Y.Y. and S.Y.; Validation, C.L. and M.-F.L.; Formal analysis, Y.Y., B.P. and C.L.; Investigation, Y.Y., S.Y. and B.P.; Resources, S.Y., B.P. and M.-F.L.; Writing—original draft, Y.Y., S.Y. and B.P.; Writing—review & editing, B.P., C.L. and M.-F.L.; Supervision, C.L. and M.-F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data from the experimental results are available from the first author. The author’s email address is y15572188780@163.com.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Kivela, M.; Arenas, A.; Barthelemy, M.; Gleeson, J.P.; Moreno, Y.; Porter, M.A. Multilayer networks. *J. Complex Netw.* **2014**, *2*, 203–271. [[CrossRef](#)]
2. Berlingerio, M.; Coscia, M.; Giannotti, F. Finding and characterizing communities in multidimensional networks. In Proceedings of the 2011 International Conference on Advances in Social Networks Analysis and Mining ASONAM 2011, Kaohsiung, Taiwan, 25–27 July 2011; IEEE Computer Society: Los Alamitos, CA, USA, 2011; pp. 490–494. [[CrossRef](#)]
3. Chen, P.Y.; Hero, A.O. Multilayer Spectral Graph Clustering via Convex Layer Aggregation: Theory and Algorithms. *arXiv* **2017**, arXiv:stat.ML/1708.02620.
4. Dong, X.; Frossard, P.; Vandergheynst, P.; Nefedov, N. Clustering on Multi-Layer Graphs via Subspace Analysis on Grassmann Manifolds. *IEEE Trans. Signal Process.* **2014**, *62*, 905–918. [[CrossRef](#)]
5. Kuncheva, Z.; Montana, G. Community detection in multiplex networks using Locally Adaptive Random walks. In Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2015), Paris, France, 25–28 August 2015; pp. 1308–1315. [[CrossRef](#)]
6. De Domenico, M.; Lancichinetti, A.; Arenas, A.; Rosvall, M. Identifying Modular Flows on Multilayer Networks Reveals Highly Overlapping Organization in Interconnected Systems. *Phys. Rev. X* **2015**, *5*, 011027. [[CrossRef](#)]
7. Amelio, A.; Pizzuti, C. Community Detection in Multidimensional Networks. In Proceedings of the 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, Limassol, Cyprus, 10–12 November 2014; pp. 352–359. [[CrossRef](#)]
8. Pamfil, A.R.; Howison, S.D.; Lambiotte, R.; Porter, M.A. Relating modularity maximization and stochastic block models in multilayer networks. *arXiv* **2018**, arXiv:cs.SI/1804.01964.
9. Mucha, P.J.; Richardson, T.; Macon, K.; Porter, M.A.; Onnela, J.P. Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science* **2010**, *328*, 876–878. [[CrossRef](#)] [[PubMed](#)]
10. Magnani, M.; Hanteer, O.; Interdonato, R.; Rossi, L.; Tagarelli, A. Community Detection in Multiplex Networks. *arXiv* **2021**, arXiv:cs.SI/1910.07646.
11. Wang, H.; Nie, F.; Huang, H.; Ding, C. Nonnegative Matrix Tri-factorization Based High-Order Co-clustering and Its Fast Implementation. In Proceedings of the 2011 IEEE 11th International Conference on Data Mining, Vancouver, BC, Canada, 11–14 December 2011; pp. 774–783. [[CrossRef](#)]
12. Sun, B.J.; Shen, H.; Gao, J.; Ouyang, W.; Cheng, X. A Non-negative Symmetric Encoder-Decoder Approach for Community Detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, Singapore, 6–10 November 2017; pp. 597–606. [[CrossRef](#)]
13. Gligorićević, V.; Panagakis, Y.; Zafeiriou, S. Non-Negative Matrix Factorizations for Multiplex Network Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 928–940. [[CrossRef](#)] [[PubMed](#)]
14. Yoo, J.; Choi, S. Orthogonal nonnegative matrix tri-factorization for co-clustering: Multiplicative updates on Stiefel manifolds. *Inf. Process. Manag.* **2010**, *46*, 559–570. [[CrossRef](#)]
15. Cai, D.; He, X.; Han, J.; Huang, T.S. Graph Regularized Nonnegative Matrix Factorization for Data Representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 1548–1560. [[CrossRef](#)] [[PubMed](#)]
16. Cao, X.; Zhang, C.; Fu, H.; Liu, S.; Zhang, H. Diversity-induced Multi-view Subspace Clustering. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA 7–12 June 2015; pp. 586–594. [[CrossRef](#)]
17. Che, H.; Pan, B.; Leung, M.F.; Cao, Y.; Yan, Z. Tensor Factorization With Sparse and Graph Regularization for Fake News Detection on Social Networks. *IEEE Trans. Comput. Soc. Syst.* **2023**, 1–11. [[CrossRef](#)]
18. Li, C.; Che, H.; Leung, M.F.; Liu, C.; Yan, Z. Robust multi-view non-negative matrix factorization with adaptive graph and diversity constraints. *Inf. Sci.* **2023**, *634*, 587–607. [[CrossRef](#)]
19. Pan, B.; Li, C.; Che, H. Nonconvex low-rank tensor approximation with graph and consistent regularizations for multi-view subspace learning. *Neural Netw.* **2023**, *161*, 638–658. [[CrossRef](#)] [[PubMed](#)]
20. Cozzo, E.; Kivelä, M.; Domenico, M.D.; Solé-Ribalta, A.; Arenas, A.; Gómez, S.; Porter, M.A.; Moreno, Y. Structure of triadic relations in multiplex networks. *New J. Phys.* **2015**, *17*, 073029. [[CrossRef](#)]
21. Gretton, A.; Bousquet, O.; Smola, A.; Schölkopf, B. Measuring Statistical Dependence with Hilbert-Schmidt Norms. In Proceedings of the International Conference on Algorithmic Learning Theory, Singapore, 8–11 October 2005.
22. Ding, C.; He, X.; Simon, H.D. On the Equivalence of Nonnegative Matrix Factorization and Spectral Clustering. In Proceedings of the 2005 SIAM International Conference on Data Mining (SDM), Newport Beach, CA, USA, 21–23 April 2005; pp. 606–610. [[CrossRef](#)]
23. Dong, Y.; Che, H.; Leung, M.F.; Liu, C.; Yan, Z. Centric graph regularized log-norm sparse non-negative matrix factorization for multi-view clustering. *Signal Process.* **2024**, *217*, 109341. [[CrossRef](#)]
24. Wu, W.; Kwong, S.; Zhou, Y.; Jia, Y.; Gao, W. Nonnegative matrix factorization with mixed hypergraph regularization for community detection. *Inf. Sci.* **2018**, *435*, 263–281. [[CrossRef](#)]
25. Lu, H.; Sang, X.; Zhao, Q.; Lu, J. Community detection algorithm based on nonnegative matrix factorization and pairwise constraints. *Phys. A Stat. Mech. Its Appl.* **2020**, *545*, 123491. [[CrossRef](#)]

26. Kong, Q.; Sun, J.; Xu, Z. Joint orthogonal symmetric non-negative matrix factorization for community detection in attribute network. *Knowl.-Based Syst.* **2024**, *283*, 111192. [[CrossRef](#)]
27. Smith Aguilar, S.; Aureli, F.; Busia, L.; Schaffner, C.; Ramos-Fernandez, G. Using multiplex networks to capture the multidimensional nature of social structure. *Primates* **2018**, *60*, 277–295. [[CrossRef](#)] [[PubMed](#)]
28. Ma, X.; Dong, D.; Wang, Q. Community Detection in Multi-Layer Networks Using Joint Nonnegative Matrix Factorization. *IEEE Trans. Knowl. Data Eng.* **2019**, *31*, 273–286. [[CrossRef](#)]
29. Al-Sharoha, E.M.; Aviyente, S. Community Detection in Fully-Connected Multi-layer Networks Through Joint Nonnegative Matrix Factorization. *IEEE Access* **2022**, *10*, 43022–43043. [[CrossRef](#)]
30. Ortiz-Bouza, M.; Aviyente, S. Community Detection in Multiplex Networks Based on Orthogonal Nonnegative Matrix Tri-Factorization. *IEEE Access* **2024**, *12*, 6423–6436. [[CrossRef](#)]
31. Lazega, E. *The Collegial Phenomenon: The Social Mechanisms of Cooperation Among Peers in a Corporate Law Partnership*; Oxford University Press: Oxford, UK, 2001. [[CrossRef](#)]
32. Greene, D.; Cunningham, P. A Matrix Factorization Approach for Integrating Multiple Data Views. In Proceedings of the ECML PKDD 2009: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Bled, Slovenia, 6–10 September 2009.
33. Rasiwasia, N.; Costa Pereira, J.; Coviello, E.; Doyle, G.; Lanckriet, G.R.; Levy, R.; Vasconcelos, N. A new approach to cross-modal multimedia retrieval. In Proceedings of the 18th ACM International Conference on Multimedia, Firenze, Italy, 25–29 October 2010; pp. 251–260. [[CrossRef](#)]
34. Bazzi, M.; Jeub, L.G.S.; Arenas, A.; Howison, S.D.; Porter, M.A. A framework for the construction of generative models for mesoscale structure in multilayer networks. *Phys. Rev. Res.* **2020**, *2*, 023100. [[CrossRef](#)]
35. Pramanik, S.; Tackx, R.; Navelkar, A.; Guillaume, J.L.; Mitra, B. Discovering Community Structure in Multilayer Networks. In Proceedings of the 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Tokyo, Japan, 19–21 October 2017; pp. 611–620. [[CrossRef](#)]
36. Kumar, A.; Rai, P.; Daumé, H. Co-regularized multi-view spectral clustering. In Proceedings of the 24th International Conference on Neural Information Processing Systems, Granada, Spain, 13–15 December 2011; pp. 1413–1421.
37. Nie, F.; Tian, L.; Li, X. Multiview Clustering via Adaptively Weighted Procrustes. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London, UK, 19–23 August 2018; pp. 2022–2030. [[CrossRef](#)]
38. Zhan, K.; Nie, F.; Wang, J.; Yang, Y. Multiview Consensus Graph Clustering. *IEEE Trans. Image Process.* **2019**, *28*, 1261–1270. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.