Journal of Electrical and Computer Engineering Innovations
(JECEI)
Journal homepage: http://www.jecei.sru.ac.ir

Research paper

A New Hybrid NMF-based Infrastructure for Community Detection in Complex Networks

M. Ghadirian, N. Bigdeli*

Department of Control Engineering, Faculty of Technical and Engineering, Imam-Khomeini International University, Qazvin, Iran.

Article Info

Article History:
Received 09 January 2023
Reviewed 28 March 2023
Revised 21 April 2023
Accepted 07 May 2023

Keywords:
Complex networks
Nonnegative matrix factorization
Modularity
General modularity density
Graph clustering

*Corresponding Author's Email Address: *n.bigdeli@eng.ikiu.ac.ir*

Abstract

Background and Objectives: Community detection is a critical problem in investigating complex networks. Community detection based on modularity/general modularity density are the popular methods with the advantage of using complex network features and the disadvantage of being NP-hard problem for clustering. Moreover, Non-negative matrix factorization (NMF)-based community detection methods are a family of community detection tools that utilize network topology; but most of them cannot thoroughly exploit network features. In this paper, a hybrid NMF-based community detection infrastructure is developed, including modularity/ general modularity density as more comprehensive indices of networks. The proposed infrastructure enables to solve the challenges of combining the NMF method with modularity/general modularity density criteria and improves the community detection methods for complex networks.

Methods: First, new representations, similar to the model of symmetric NMF, are derived for the model of community detection based on modularity/general modularity density. Next, these indices are innovatively augmented to the proposed hybrid NMF-based model as two novel models called 'general modularity density NMF (GMDNMF) and mixed modularity NMF (MMNMF)'. In order to solve these two NP-hard problems, two iterative optimization algorithms are developed.

Results: it is proved that the modularity/general modularity density-based community detection can be consistently represented in the form of SNMF-based community detection. The performances of the proposed models are verified on various artificial and real-world networks of different sizes. It is shown that MMNMF and GMDNMF models outperform other community detection methods. Moreover, the GMDNMF model has better performance with higher computational complexity compared to the MMNMF model.

Conclusion: The results show that the proposed MMNMF model improves the performance of community detection based on NMF by employing the modularity index as a network feature for the NMF model, and the proposed GMDNMF model enhances NMF-based community detection by using the general modularity density index.

Introduction

Networks are used to model complex interconnected data. Common examples of this type of modeling are biological networks, social networks, and citation networks [1]. One standard method for representing a

network is using a graph data structure consisting of nodes and edges. Exploring and understanding network structures can provide useful information about interconnections. For instance, in social networks, the edges between the nodes depict the interaction between users.

Community detection is one of the powerful analysis methods that help understand the organization of network structures [1]. In social networks, a community (also called a cluster, a module, or a group) comprises users or nodes with close relations or connections but lost connections with others. Over the past years, various measurement criteria have been proposed to evaluate the quality of graph partitioning, including modularity and normalized mutual information (NMI) [2]. Some of these criteria such as modularity and general modularity density [3] have been applied to cluster complex networks as well. By the introduction of modularity, many modularity-based community detection algorithms have been suggested, including integer programming [4], genetic algorithm [5], greedy algorithm [6], and vector partition problem [7]. However, the modularity index has some restrictions in community detection [3], [8]. For example, the modularity depends on the total size of the links, and small communities tend to be merged into large communities. Therefore, new optimization criteria have been offered for community detection algorithms, including general modularity density maximization [3] and localized modularity optimization [8]. Detecting communities based on general modularity density is superior in cases such as resolving most modular networks, detecting communities of different sizes, and not dividing a clique into two parts. In recent years, various methods have been proposed for optimizing modularity density maximization as their index. For instance, one may refer to mixed integer linear programming [9], genetic algorithm [10], memetic algorithm [11], and linear mathematical programming, which consisted of two models based on mixed-integer linear programming and two models based on binary decomposition [12].

In the literature, other algorithms with different approaches have also been presented for clustering complex networks. They include the label propagation algorithm [13], [14], random walk algorithm [15], [16], greedy and weight-balanced algorithm [17], and nonnegative matrix factorization (NMF) algorithm [18]. Among them, NMF-based clustering methods have attracted much attention and wide applications in various fields such as speech separation [19], image processing [20], hyperspectral unmixing [21], document clustering [22], community detection in graph mining and data mining [23], and detection of fake news in social media [24]. On the other hand, one of the recent challenges in community detection is to use network features or its a' priori information to improve the performance of NMF. For instance, a' prior information was innovatively integrated into NMF and used for community detection in some studies [25], [26]. The GNMF (graph regularized NMF) model was improved using hypergraph regularization in previous research [27]. The modularized deep NMF (MDNMF) extended DNMF (Deep NMF) [28] by preserving topology information and instinct community structure properties [28], [29]. Community detection was developed by integrating the tri-NMF model with the modularized information called the 'Mtrinmf method' [30]. Community detection was developed by integrating the tri-NMF model with the modularized information called the 'Mtrinmf method' [30]. In the Mtrinmf method, the tri-NMF model is linearly combined with a modularity optimization. However, the linear combination method could not be generalized to the NMF model due to its modularity structure. This is a main drawback in the existing literature in this area, as, to the authors' best knowledge, the NMF model has not been customized using the modularity index or general modularity density index as network features for community detection. Using this customization can improve data clustering based on the NMF method to detect communities on complex networks.

Based on the provided discussions, NMF-based and modularity/general modularity density-based community detection methods are the most important community detection methods which have both advantages and limitations. That is, modularity/general modularity density is specific to the clustering of complex networks, while, it suffers from the NP-hard problem. On the other hand, the NMF-community detection benefits from the initial knowledge, having iterative solution and generality for all types of data, but, it is not specified for complex networks. These properties motivated the authors to consolidate the advantages and refine the properties of the NMF-based and modularity/general modularity density-based community detection methods via introducing a new hybrid NMF-based community detection infrastructure. The new hybrid NMF-based community detection infrastructure includes modularity or general modularity density as more comprehensive indices of the complex networks. In this way, various features such as prior information and community structure are simultaneously employed to cluster the networks. However, the model structures of the NMF-based community detection method and modularity/modularity density-based community detection methods are not consistent. Therefore, to develop the unified infrastructure, first, new representations would be derived for the model of community detection based on modularity/general

modularity density, which are similar to the model of community detection based on symmetric NMF (SNMF). SNMF clustering is one of the symmetric types of NMF methods that has been mentioned in a previous study [31]. These consistent representations would be later employed to extend NMF for community detection in complex networks. Next, two novel hybrid community detection methods are proposed due to the difference between community detection based on modularity/modularity density and the NMF and considering the derived equivalent representations of modularity/modularity density indices. These methods are called mixed modularity nonnegative matrix factorization (MMNMF) and general modularity density nonnegative matrix factorization (GMDNMF). MMNMF and GMDNMF would innovatively integrate modularity/general modularity density indices into the NMF model, respectively, to improve community detection performance by utilizing proper network features. However, these methods are NP-hard problems, which should be solved numerically. Therefore, iterative update rules would be derived and proved as optimal solutions for MMNMF and GMDNMF models. The performance of the two models is verified on two practical artificial networks and ten real-world networks. It is indicated that although MMNMF and GMDNMF have better performance respecting other community detection methods, these methods slightly differ in precision and computational complexity, which would be calculated and compared in this study. The final preference should be therefore performed based on problem requirements. The remaining sections of this paper are organized as follows:

Section 2 studies modularity and general modularity density optimizations and reviews related works on NMF-based community detection methods. Section 3 presents our proposed methods and analyzes iterative optimization algorithms. The computational complexity of our methods is calculated in Section 4, followed by studying the effect of parameters and presenting several comparative experimental results. Finally, Section 5 summarizes the proposed procedures, achievements, and discusses future works.

Related Works

Community detection based on the NMF method is one of the efficient methods for clustering various types of data such as audio, text, image, graph, and the like. This paper proposes a combination of NMF-based community detection and community detection based on network features. Therefore, first, a study of community detection based on modularity and general modularity density is presented, and then NMF-based community detection methods are reviewed and discussed in this section.

Modularity Maximization

A graph can represent a complex network without loss of network features. $G = (V, E)$ is a representation of directed (undirected) and unweighted graphs where V denotes the set of n nodes and E is a set of m edges between the two nodes. Modularity maximization is a popular community detection method for understanding the structure of networks. It considers the strength of the relationship density of each edge within each community and can regard the related nodes between communities. An algorithm based on modularity maximization clustering a network in two-by-two communities was first proposed by Newman [1]. Next, to cluster more than two-by-two communities, the generalized modularity index (Q) was presented as follows [1], [30]:

$$Q = \frac{1}{2m} tr(X^T B X) \tag{1a}$$

$$B = A - B_1 \tag{1b}$$

$$(B_1)_{ij} = \frac{k_i k_j}{2m} \tag{1c}$$

where A and B are adjacency and modularity matrices, respectively. In addition, k_i ,$X \in R^{n \times k}$ and m are the degree of the i^{th} nodes, the community membership matrix and the number of edges, respectively. Besides, k denotes the number of communities in the complex network. Moreover, the maximization problem can be rewritten as [30]:

$$\max_X Q = \max_X \frac{1}{2m} tr(X^T B X) \tag{2}$$

where $X^T X = I$ satisfies the orthogonal condition [32].

Since this method is a NP-hard problem, finding a way to solve it has been one of the challenges over the past years.

General Modularity Density Maximization

Community detection based on general modularity density maximization is efficient for clustering complex networks [3]. The general modularity density index considers each cluster's average inner degree and outer degree. The inner degree refers to the sum of the edges of interval nodes in each cluster, and the outer degree is the sum of edges between the nodes inside the cluster with the nodes of another cluster. It can be rewritten for k number of partitions ($\{V_r\}_{r=1}^k$) as follows [3]:

$$D_\lambda(\{V_r\}_{r=1}^k) = \sum_{r=1}^k \frac{2\lambda l(V_r, V_r) - 2(1-\lambda) l(V_r, \bar{V}_r)}{|V_r|} \tag{3}$$

where $l(V_1, V_2) = \sum_{i \in V_1, j \in V_2} A_{ij}$, $l(V_1, \bar{V}_1) = \sum_{i \in V_1, j \in \bar{V}_1} A_{ij}$, $\bar{V}_1 = V \backslash V_1$ and V_r is the set of vertices in the rth community. Furthermore, D_λ evaluates small and large clusters by using ratio association and ratio cut for $\lambda < 0.5$ and $\lambda > 0.5$, respectively. Therefore, D_λ equals modularity density when $\lambda = 0.5$. Advantages such as

selecting the best communities with different sizes, not dividing cliques, and resolving graph types are obtained by selecting different λ values.

Lemma 2.1 [33]. D_λ can be written as a trace of the similarity matrix of a complex network as follows:

$$D_\lambda(\{V_r\}_{r=1}^k) = tr\left(\tilde{X}^T(2A - 2(1-\lambda)C)\tilde{X}\right), \tilde{X} = XD \quad (4)$$

where D is a diagonal matrix with $D_{ii} = 1/\sqrt{\sum_{j=1}^n X_{ji}^2}$ values. Additionally, X and C represent a community relation matrix and a diagonal matrix with $C_{ii} = \sum_{j=1}^n A_{ii}$ values, respectively.

Proof: Given that X_{ir} indicates the existence probability of node i which belongs to the community and $X_{.r} = (X_{1r}, \dots, X_{nr})$ represents the probability of each node which belongs to V_r, D_λ in (3) can be rewritten according to $l(V_r, \bar{V}_r) = l(V_r, V) - l(V_r, V_r)$ as:

$$D_\lambda(\{V_r\}_{r=1}^k) = \sum_{r=1}^k \frac{2l(V_r,V_r) - 2(1-\lambda)l(V_r,V)}{|V_r|} \quad (5)$$

Moreover, $|V_r|$, $l(V_r, V_r)$ and $l(V_r, V)$ can be written based on $X_{.r}$ as $|V_r| = X_{.r}^T X_{.r}$, $l(V_r, V_r) = X_{.r}^T A X_{.r}$, $l(V_r, V) = X_{.r}^T C X_{.r}$ Thus (5) can be reformulated as:

$$D_\lambda(\{V_r\}_{r=1}^k) = \sum_{r=1}^k \frac{2X_{.r}^T A X_{.r} - 2(1-\lambda)X_{.r}^T C X_{.r}}{X_{.r}^T X_{.r}} = \sum_{r=1}^k \tilde{X}_{.r}(2A - 2(1-\lambda)C)\tilde{X} \quad (6)$$

where $\tilde{X}_{.r} = X_{.r}/\sqrt{X_{.r}^T X_{.r}}$ or $\tilde{X} = XD$ if $D_{ii} = 1/\sqrt{\sum_{j=1}^n X_{ji}^2}$. Therefore, (3) can be represented as (4), and the proof is completed accordingly.

Corollary: Considering that D_λ can be formulated as a trace form of (4), community detection based on general modularity density is a maximization problem with D_λ as its cost function, which can be rewritten as follows [12], [33]:

$$\max_{\tilde{X}} D_\lambda(\{V_r\}_{r=1}^k) = \max_{\tilde{X}} tr\left(\tilde{X}^T B_2 \tilde{X}\right) \quad (7)$$

$$s.t. \tilde{X} > 0, \tilde{X}^T\tilde{X} = I_k$$

where, $B_2 = 2A - 2(1-\lambda)C$, $C, \tilde{X}^T\tilde{X} = I$ are modularity density matrix, a diagonal matrix with $C_{ii} = \sum_{j=1}^n A_{ii}$ and a relaxing condition for orthogonality, respectively.

Nonnegative Matrix Factorization (NMF)

NMF models factorize a given similarity matrix $Y \in R^{n\times n}$ into two new matrices $W \in R^{n\times k}$ and $H \in R^{n\times k}$: $Y \simeq WH^T$ where W and H are called the community indicator feature matrix and called community relation matrix, respectively. The error between Y and WH^T is measured by a cost function of $J_{NMF}(W, H)$. W and H can be found by minimizing $J_{NMF}(W, H)$ as follows:

$$\min_{W,H} J_{nmf}(W, H) = \|Y - WH^T\|_F^2 \quad (8)$$

where $\|.\|_f$ stands for the Frobenius norm. If Y is assumed a symmetric matrix (such as undirected graph), then all the characteristics of the clustering index can be aggregated in one matrix ($W = H$). Therefore, as an extension to NMF, SNMF can drastically improve community detection. The objective function of SNMF model would be rewritten as follows [31]:

$$\min_H J_{SNMF}(H) = \|Y - HH^T\|_F^2 \quad (9)$$

Using network features or prior information in the NMF-based methods has been a challenging topic in recent years. Lu et al. [26] have recently used prior information to improve community detection and proposed two semi-supervised NMF-based methods named SVDCNMF and SVDCSNMF. In this method, the adjacency matrix A is considered as the similarity matrix (i.e., $Y = A$, $H = X$), and its objective function is represented as:

$$\min_X J_{\text{SVDCSNMF}}(X) = \|A - XX^T\|_F^2 + 2\lambda tr(X^T L X) \quad (10)$$

where L is the graph Laplacian of prior information. Since the Laplacian matrix is specific to the graph structure, this method will not provide the best clustering for other types of network features (such as modularity).

He et al. [25] suggested a robust semi-supervised NMF method named RSSNMF to enhance the robustness of semi-supervised NMF for uncertainties and errors in prior information. The cost function of the RSSNMF method is as follows:

$$\min_X J_{\text{RSSNMF}}(X) = \|A - XX^T\|_2 + \alpha tr(X^T P X Q) + \beta tr(X^T R X) \quad (11)$$

where α and β are semi-supervised tunning parameters and $Q = \begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix}$. Prior information is applied in the following matrix:

$$P_{ij} = \begin{cases} 1 & if\ \ x_i, x_j\ have\ same\ labels \text{ or } i = j \\ 0 & otherwise \end{cases} \quad (12)$$

$$R_{ij} = \begin{cases} 1 & if\ x_i, x_j\ have\ diffrent\ labels \\ 0 & otherwise \end{cases} \quad (13)$$

In addition, the terms $tr(X^T P X Q)$ and $tr(X^T R X)$ are derived from must- and cannot-link pairwise constraints among nodes, respectively. This method only uses the prior information and did not consider other features of the network structure.

Similarly, Wu et al. [27] proposed a mixed hypergraph NMF named MHGNMF by combining NMF with hypergraph regularization, which encodes the higher-order information into NMF by hypergraph. The objective function of MHGNMF is defined as:

$$\min_X J_{\text{MHGNMF}}(X) = \|A - XX^T\|_F^2 + \beta tr(X^T L_h X) \quad (14)$$

where L_h is hyperlaplacian matrix. Likewise, Yan et al. [30] combined modularity optimization and tri-NMF-based

community detection named Mtrinmf, which uses network features to enhance the performance of tri-NMF. The cost function of the Mtrinmf method can be written as (15):

$$\min_{U,X} J_{\text{Mtrinmf}}(U,X) = \|A - XUX^T\|_F^2 - \beta tr(X^T BX) \quad (15)$$

where, $tr(X^T BX)$,U and $\|A - XUX^T\|_F^2$ are the modularity optimization, the community indicator feature matrix and triNMF optimization terms, respectively. The Mtrinmf model innovatively combines the modularity criterion linearly with the triNMF model and has produced the best clustering ever.

Additionally, Huang et al. [29] suggested a new model named modularized deep nonnegative matrix factorization (MDNMF), which combined modularity and DNMF-based community detection. Deep NMF (DNMF) is another extension of NMF model that acquires additional levels of abstraction of the similarity between the nodes of each levels [29] and factorizes a given adjacency matrix A into $p+1$ nonnegative factors.

The MDNMF model has been composed as follows:

$$\min_{U_i,X,M,C} J_{\text{MDNMF}} = \left\|A - U_1U_2 \dots U_pX^T\right\|_F^2 + \alpha\|M - X^T C^T\|_F^2 - \beta tr(M^T BM) + \lambda tr(XLX^T) \quad (16)$$

$$st. U_i \geq 0, X \geq 0, \forall i = 1,2,\dots,p$$

where, L , M, C and λ denote the graph Laplacian matrix, the modularity cluster term, the final cluster term and the regularization parameter, respectively. $\lambda tr(XLX^T)$ utilizes a regularized graph and $\left\|A - U_1U_2 \dots U_pX^T\right\|_F^2$ refers to DNMF-based community detection model. This is one of the methods that has combined the graph features such as modularity criterion with DNMF-based community detection and illustrated a suitable clustering, but due to the DNMF-based community detection model, this method will have a high computational complexity and a high dependence on the correct selection of parameters in DNMF model.

From the above-mentioned discussion, it can be concluded that the modularity maximization in (1) suffers from the NP-hard problem [1], [3], [8]. Moreover, The NMF models have attracted much attention and have wide applications in various fields [9]-[12]. On the other hand, general modularity density maximization in (2) has not been used to improve NMF-based community detection, yet. Therefore, combining the NMF models and the graph features such as modularity/ general modularity density criterion will be considered, in this paper. The main achievements of the proposed methods are development of an iterative solution for modularity optimization with NMF model, specialization of NMF model for complex networks, and utilizing general modularity density criterion for NMF.

The Proposed Methods

According to the provided discussions about various NMF-based community detection methods, one could conclude that modularity and general modularity density indices as the network features have not been employed with NMF in a unified community detection method. To extend the NMF-based method to a hybrid method containing these indices, first, we derive the new representation of modularity-based and general modularity density-based community detection methods that are similar to the model of community detection based on SNMF. Next, these consistent representations help combine NMF-based community detection with modularity/general modularity density indices. It leads us to propose general modularity density nonnegative matrix factorization (GMDNMF) and mixed modularity nonnegative matrix factorization (MMNMF) models. Finally, proper iterative optimization methods for solving MMNMF and GMDNMF problems are developed accordingly.

New Representations of Modularity/General Modularity Density Maximization

In this section, new models of the modularity/general modularity density optimization are derived, which are similar to the model of symmetric nonnegative matrix factorization optimization problem summarized as:

Theorem 3.1. The modularity optimization in (2) and general modularity density optimization in (7) for complex networks can be represented in a similar form of the SNMF model, respectively, as follows:

$$\max_X Q = \min_X \|B - XX^T\|_F^2 \quad (17a)$$

and

$$\max_{\tilde{X}} D_\lambda(\{V_r\}_{r=1}^k) = \min_{W,\tilde{X}} \left\|B_2 - \tilde{X}\tilde{X}^T\right\|_F^2 \quad (17b)$$

As shown, the new formulation in (17) is similar to the model of SNMF in (9), while in (17a) and (17b), Y is replaced with B and B_2, respectively; in addition, H is replaced with X and $\tilde{X}$, respectively.

Proof: Modularity optimization of (2) is re-formulated as follows:

$$\max_X \frac{1}{2m} tr(X^T BX) \propto -\frac{1}{2m} \min_X tr(X^T BX) \propto -\min_X tr(X^T BX) \quad (18a)$$

If $X^T X = I$ and B is constant, (18a) is re-written as:

$$\max_X \frac{1}{2m} tr(X^T BX) \propto \min_X \left(tr(X^T XX^T X) - 2tr(X^T BX) + tr(BB^T)\right) \quad (18b)$$

According to trace properties, namely, $tr(X^T BX) = tr(BXX^T)$, $tr(X^T XX^T X) = tr(XX^T XX^T)$, (18b) is re-written as:

$$\max_X \frac{1}{2m} tr(X^T BX) \propto \min_X tr(XX^T XX^T - 2BXX^T + BB^T) \propto \min_X \|B - XX^T\|_F^2 \quad (19)$$

As a result, the new representation of modularity optimization is consistent with SNMF. The consistent representation of general modularity density optimization and SNMF can be similarly derived, completing the proof.

GMDNMF and MMNMF Models

The NMF model is an efficient method for clustering data types. However, it is not the best method for clustering complex networks because it may ignore some useful information and characteristics such as general modularity density and modularity indices. Motivated by this observation, in this section, general modularity density and modularity indices are augmented to NMF-based community detection to improve its performance. For this purpose, we refer to Theorem 3.1, when, it was shown that community detection based on modularity optimization (2) can be represented as an SNMF optimization problem with similarity matrix B, and community detection based on general modularity density optimization (7) is similar to community detection based on SNMF with similarity matrix B_2. Therefore, according to Theorem 3.1, for improving community detection based on NMF via modularity or general modularity density indices, it is necessary to combine SNMF-based community detection and NMF-based community detection methods. However, due to different structures of community detection based on NMF (NMF optimization with similarity matrix A) and community detection based on modularity/general modularity density optimization (SNMF optimization with similarity matrix B/B_2), NMF-based community detection cannot be linearly combined with modularity/general modularity density-based community detection as in previous research. Thus, multi-view clustering via joint NMF such as the methods presented in other studies [23] and [34] would be exploited in this paper. In this context, MMNMF and GMDNMF models are proposed to integrate modularity and general modularity density into the NMF model to improve community detection. These models are devised for complex networks as:

$$\min_{W,X,\tilde{X},X^*} J_{\text{MMNMF}} = \|A - WX^T\|_F^2 - tr(\tilde{X}^T B\tilde{X}) + \frac{1}{2}\|\tilde{X} - X^*\|_F^2 + \frac{1}{2}\|X - X^*\|_F^2 \quad (20)$$

$$s.t. W, X, \tilde{X}, X^* > 0\ , \sum_{r=1}^{k} X_{ir} = 1\ , \tilde{X}^T\tilde{X} = I_k$$

and

$$\min_{W,X,\tilde{X},X^*} J_{\text{GMDNMF}} = \|A - WX^T\|_F^2 - tr(\tilde{X}^T B_2\tilde{X}) + \frac{1}{2}\|\tilde{X}\tilde{D} - X^*\|_F^2 + \frac{1}{2}\|X - X^*\|_F^2 \quad (21)$$

$$s.t. W, X, \tilde{X}, X^* > 0\ , \sum_{r=1}^{k} X_{ir} = 1\ , \tilde{X}^T\tilde{X} = I_k$$

where X^* is the result of community detection models and $\tilde{D}$ denotes a diagonal matrix with $\tilde{D}_{ii} = \sqrt{\sum_{j=1}^{n} X_{ji}^2}$ or $\tilde{D} = D^{-1}$ values. In (20) and (21), $\|A - WX^T\|_F^2$ represents NMF-based community detection, and $tr(\tilde{X}^T B\tilde{X})$ and $tr(\tilde{X}^T B_2\tilde{X})$ refer to community detection based on modularity/general modularity density indices, respectively. Here, X^* is an interface parameter for combining NMF-based community detection with modularity-based and general modularity density-based community detection methods. GMDNMF and MMNMF models are NP-hard problems due to the orthogonal constraint. Many methods exist for extending the orthogonal constraint to a nonnegative term [2], [35]. For this purpose, we use the presented method in previous studies [35]. This method adds an orthogonal constraint to the objective model and can preserve clustering performance. Accordingly, the new objective functions for GMDNMF and MMNMF models are formulated as:

$$\min_{W,X,\tilde{X},X^*} J_{\text{MMNMF}} = \|A - WX^T\|_F^2 - tr(\tilde{X}^T B\tilde{X}) + \frac{1}{2}\|\tilde{X} - X^*\|_F^2 + \frac{1}{2}\|X - X^*\|_F^2 + \eta\|\tilde{X}^T\tilde{X} - I\| \quad (22)$$

$$s.t. W, X, \tilde{X}, X^* > 0\ , \sum_{r=1}^{k} X_{ir} = 1$$

$$\min_{W,X,\tilde{X},X^*} J_{\text{GMDNMF}} = \|A - WX^T\|_F^2 - tr(\tilde{X}^T B_2\tilde{X}) + \frac{1}{2}\|\tilde{X}\tilde{D} - X^*\|_F^2 + \frac{1}{2}\|X - X^*\|_F^2 + \eta\|\tilde{X}^T\tilde{X} - I\| \quad (23)$$

$$s.t. W, X, \tilde{X}, X^* > 0\ , \sum_{r=1}^{k} X^*{}_{ir} = 1$$

where η and $\|\tilde{X}^T\tilde{X} - I\|$ are the orthogonal condition control parameter and the orthogonal condition control cost, respectively. It is worth mentioning that the value of parameter λ in B_2, which was first introduced in (16), is chosen via a simple rule presented in [30]. That is, this parameter is selected in the interval of [0, 1] in small steps (e.g., 0.1). Then, the best community and its relating value of λ is selected by the best-obtained modularity index value.

Iterative Optimization Algorithms for MMNMF and GMDNMF Models

This section will develop an iterative method to solve the proposed MMNMF model of (22) and GMDNMF model of (23). This iterative method is performed via an alternative updating strategy (i.e., the model

updates one variable via the Lagrange method while the other variables are constant). Finally, the update variable process is repeated until the convergence or reaching the final iteration number.

Iterative Optimization Algorithm for the GMDNMF Model

The trace form (23) can be rewritten as:

$$\min_{W,X,\tilde{X},X^*} J_{\text{GMDNMF}} = \big(tr(AA^T) - 2tr(AXW^T) + Tr(WX^T XW^T)\big) - tr(\tilde{X}^T B\tilde{X}) + \frac{1}{2}\big(tr(\tilde{X}\tilde{D}\tilde{D}^T\tilde{X}^T) -$$

$$2tr\left(\tilde{X}\tilde{D}X^{*T}\right)+tr\left(X^{*}X^{*T}\right)\Big)+\frac{1}{2}\Big(tr(XX^{T})-2tr\left(XX^{*T}\right)+tr\left(X^{*}X^{*T}\right)\Big)+\eta\left(tr\left(\tilde{X}^{T}\tilde{X}\tilde{X}^{T}\tilde{X}\right)-2tr\left(\tilde{X}^{T}\tilde{X}\right)+k\right) \quad (24)$$

Given that $\tilde{D}_{ii}$ equals $\sqrt{\sum_{j=1}^{n}X_{ji}^{2}}$, calculating iterative updating rules for X is more complex compared to the other variables. Accordingly, first, updating rules are formulated for W, $\tilde{X}$, and X^{*}, followed by deriving the iterative rules for updating X.

Updating rules for W, $\tilde{X}$, and X^{*}:

the Lagrange cost function for (24) can be resulted as:

$$L_{\mathrm{GMDNMF}}=J_{\mathrm{GMDNMF}}+tr(\psi W)+tr\left(\phi\tilde{X}\right)+tr(\varphi X^{*}) \quad (25)$$

where ψ, ϕ, and φ are the Lagrangian multipliers for constraints $W>0$, $\tilde{X}>0$, and $X^{*}>0$, respectively. Then, the derivatives of L_{GMDNMF} would be then derived as follows:

$$\frac{\partial L_{\mathrm{GMDNMF}}}{\partial W}=\psi+2WX^{T}X-2AX$$
$$\frac{\partial L_{\mathrm{GMDNMF}}}{\partial\tilde{X}}=\phi-2B\tilde{X}^{T}+\tilde{X}\tilde{D}\tilde{D}^{T}-X^{*}\tilde{D}^{T}+4\eta\tilde{X}\tilde{X}^{T}\tilde{X}-4\eta\tilde{X}$$
$$\frac{\partial L_{\mathrm{GMDNMF}}}{\partial X^{*}}=\varphi-\tilde{X}\tilde{D}+X^{*}-X+X^{*} \quad (26)$$

where $B=2A-2(1-\lambda)C$. According to Karush-Kuhn-Tucker (KKT) conditions (i.e., $\psi_{ir}W_{ir}=0$, $\phi_{ir}\tilde{X}_{ir}=0$, and $\varphi_{ir}X_{ir}^{*}=0$), the solution can be formulated as follows:

$$(WX^{T}X)_{ir}W_{ir}-(AX)_{ir}W_{ir}=0$$
$$-4\left(A\tilde{X}^{T}\right)_{ir}\tilde{X}_{ir}+4(1-\lambda)\left(C\tilde{X}^{T}\right)_{ir}\tilde{X}_{ir}+\left(\tilde{X}\tilde{D}\tilde{D}^{T}\right)_{ir}\tilde{X}_{ir}-\left(X^{*}\tilde{D}^{T}\right)_{ir}\tilde{X}_{ir}+4\eta\left(\tilde{X}\tilde{X}^{T}\tilde{X}\right)_{ir}\tilde{X}_{ir}-4\eta\left(\tilde{X}\right)_{ir}\tilde{X}_{ir}=0$$
$$-\left(\tilde{X}\tilde{D}\right)_{ir}X_{ir}^{*}-(X)_{ir}X_{ir}^{*}+2(X^{*})_{ir}X_{ir}^{*}=0 \quad (27)$$

Finally, iterative updating rules are formulated as:

$$W_{ir}=W_{ir}.\frac{(AX)_{ir}}{(WX^{T}X)_{ir}}$$
$$\tilde{X}_{ir}=\tilde{X}_{ir}.\frac{4(A\tilde{X}^{T})_{ir}+(X^{*}\tilde{D}^{T})_{ir}+4\eta(\tilde{X})_{ir}}{4(1-\lambda)(C\tilde{X}^{T})_{ir}+(\tilde{X}\tilde{D}\tilde{D}^{T})_{ir}+4\eta(\tilde{X}\tilde{X}^{T}\tilde{X})_{ir}}$$
$$X_{ir}^{*}=X_{ir}^{*}.\frac{(\tilde{X}\tilde{D})_{ir}+(X)_{ir}}{2(X^{*})_{ir}} \quad (28)$$

Due to the probability of having zero values in the denominator of X^{*} according to updating rules in (28), a modified updating function is:

$$X_{ir}^{*}=X_{ir}^{*}.\frac{(\tilde{X}\tilde{D})_{ir}+(X)_{ir}+10^{-9}}{2(X^{*})_{ir}+10^{-9}} \quad (29)$$

Furthermore, to satisfy $\sum_{\mathrm{r=1}}^{\mathrm{k}}\mathrm{X^{*}_{ir}}=1$ condition, we use the same method in previous research [30] as:

$$X^{*}{}_{ir}:=\frac{X^{*}{}_{ir}}{\sum_{r=1}^{c}X^{*}{}_{ir}} \quad (30)$$

Updating rule for X:

The Lagrange cost function where Ω is the Lagrangian multiplier for constraint $X>0$ can be rewritten as:

$$L_{\mathrm{GMDNMF}}=\left(tr(WX^{T}XW^{T})-2tr(AXW^{T})\right)+\frac{1}{2}\Big(tr(XX^{T})-2tr\left(XX^{*T}\right)\Big)+\frac{1}{2}R+tr(\Omega X) \quad (31)$$

where $R=tr\left(\tilde{X}\tilde{D}\tilde{D}^{T}\tilde{X}^{T}\right)-2tr\left(\tilde{X}\tilde{D}X^{*T}\right)$ and $\tilde{D}$ is the diagonal matrix with $\tilde{D}_{ii}=\sqrt{\sum_{j=1}^{n}X_{ji}^{2}}$ values. One can rewrite R as follows [36]:

$$R=\sum_{j=1}^{n}\sum_{r=1}^{k}\tilde{X}_{jr}\sum_{i=1}^{n}X_{ir}^{2}\tilde{X}_{jr}-2\sum_{j=1}^{n}\sum_{r=1}^{k}\tilde{X}_{jr}\sqrt{\sum_{i=1}^{n}X_{ir}^{2}}X^{*}{}_{jr} \quad (32)$$

The partial derivative of R with respect to X_{ir} is as follows:

$$P_{ir}=\frac{\partial R}{\partial X_{ir}}=2\left(X_{ir}\sum_{j=1}^{n}\tilde{X}_{jr}^{2}-\frac{X_{ir}}{\sqrt{\sum_{i=1}^{n}X_{ir}^{2}}}\sum_{j=1}^{n}\tilde{X}_{jr}X^{*}{}_{jr}\right) \quad (33)$$

Therefore, the derivatives of L_{GMDNMF} can be derived as (34):

$$\frac{\partial L_{2}}{\partial X}=\Omega+g_{1}(2XW^{T}W-2A^{T}W)+\frac{1}{2}(2X-2X^{*})+\frac{1}{2}P \quad (34)$$

By using KKT conditions (i.e., $\Omega_{ir}X_{ir}=0$), one has:

$$(2XW^{T}W-2A^{T}W)_{ir}X_{ir}+(X-X^{*})_{ir}X_{ir}+\frac{1}{2}P_{ir}X_{ir}=0 \quad (35)$$

Considering (31) and (33), iterative updating rules would be as follows:

$$X_{ir}=X_{ir}.\frac{X_{ir}^{*}+2(A^{T}W)_{ir}+\frac{X_{ir}}{\sqrt{\sum_{i=1}^{n}X_{ir}^{2}}}\sum_{j=1}^{n}\tilde{X}_{jr}X^{*}{}_{jr}}{X_{ir}+2(XW^{T}W)_{ir}+X_{ir}\sum_{j=1}^{n}\tilde{X}_{jr}^{2}} \quad (36)$$

Eventually, according to (28), (29), (30), and (36), the GMDNMF model is proposed as Algorithm 1.

Algorithm1: GMDNMF model

Inputs:
- Adjacency matrix A
- Number of communities k
- General density parameters λ
- Orthogonal condition control parameter η
- Maximum number of iterations I_t

Output:
- Clustering label of each node

1: **Initialize $\boldsymbol{W}, \boldsymbol{X}, \tilde{\boldsymbol{X}}$ and $\boldsymbol{X}^{*}$**
2: **For $t=1$: I_t do**
3: $W_{ir}=W_{ir}.\frac{(AX)_{ir}}{(WX^{T}X)_{ir}}$
4: $\tilde{X}_{ir}=\tilde{X}_{ir}.\frac{4(A\tilde{X}^{T})_{ir}+(X^{*}\tilde{D}^{T})_{ir}+4\eta(\tilde{X})_{ir}}{4(1-\lambda)(C\tilde{X}^{T})_{ir}+(\tilde{X}\tilde{D}\tilde{D}^{T})_{ir}+4\eta(\tilde{X}\tilde{X}^{T}\tilde{X})_{ir}}$
5: $X_{ir}=X_{ir}.\frac{X_{ir}^{*}+2(A^{T}W)_{ir}+\frac{X_{ir}}{\sqrt{\sum_{i=1}^{n}X_{ir}^{2}}}\sum_{j=1}^{n}\tilde{X}_{jr}X^{*}{}_{jr}}{X_{ir}+2(XW^{T}W)_{ir}+X_{ir}\sum_{j=1}^{n}\tilde{X}_{jr}^{2}}$
6: $X_{ir}^{*}=X_{ir}^{*}.\frac{(\tilde{X}\tilde{D})_{ir}+(X)_{ir}+10^{-9}}{2(X^{*})_{ir}+10^{-9}}$
7: $X^{*}{}_{ir}:=\frac{X^{*}{}_{ir}}{\sum_{r=1}^{c}X^{*}{}_{ir}}$
8: **End for**
9: **Return $(\boldsymbol{v_i}, \boldsymbol{I_i})=\boldsymbol{argmax}_{r\leq k}\boldsymbol{X}_{ir}^{*}$**

Iterative Optimization Algorithm for the MMNMF Model

To the iterative optimization algorithm for the MMNMF model, first, one can define the trace form of (22) as follows:

$$\min_{W,X,\tilde{X},X^*} J_{\text{MMNMF}} = \left(tr(AA^T) - 2tr(AXW^T) + Tr(WX^TXW^T)\right) - tr\left(\tilde{X}^TB\tilde{X}\right) + \frac{1}{2}\left(tr\left(\tilde{X}\tilde{X}^T\right) - 2tr\left(\tilde{X}X^{*T}\right) + tr\left(X^*X^{*T}\right)\right) + \frac{1}{2}\left(tr(XX^T) - 2tr\left(XX^{*T}\right) + tr\left(X^*X^{*T}\right)\right) + \eta\left(tr\left(\tilde{X}^T\tilde{X}\tilde{X}^T\tilde{X}\right) - 2tr\left(\tilde{X}^T\tilde{X}\right) + k\right) \quad (37)$$

$$s.t.\, W, X, \tilde{X}, X^* > 0\,, \sum_{r=1}^{k} X^*{}_{ir} = 1$$

Following the same procedure of Section 3.3.1 (Appendix A), the iterative updating rules are derived (38):

$$W_{ir} = W_{ir}.\frac{(AX)_{ir}}{(WX^TX)_{ir}}$$

$$X_{ir} = X_{ir}.\frac{2(A^TW)_{ir}+(X^*)_{ir}}{2(XW^TW)_{ir}+(X)_{ir}}$$

$$\tilde{X}_{ir} = \tilde{X}_{ir}.\frac{2(A\tilde{X}^T)_{ir}+(X^*)_{ir}+4\eta(\tilde{X})_{ir}}{2(B_1\tilde{X}^T)_{ir}+(\tilde{X})_{ir}+4\eta(\tilde{X}\tilde{X}^T\tilde{X})_{ir}}$$

$$X^*_{ir} = X^*_{ir}.\frac{(\tilde{X})_{ir}+(X)_{ir}+10^{-9}}{2(X^*)_{ir}+10^{-9}}$$

$$X^*{}_{ir} := \frac{X^*{}_{ir}}{\sum_{r=1}^{c} X^*{}_{ir}} \quad (38)$$

Finally, the iterative optimization algorithm for the MMNMF model is suggested as Algorithm 2.

Algorithm2: MMNMF model

Inputs:

Adjacency matrix A

Number of communities k

Orthogonal condition control parameter η

Maximum number of iterations I_t

Output:

Clustering label of each node

Initialized $\boldsymbol{W}, \boldsymbol{X}, \tilde{\boldsymbol{X}}$ and $\boldsymbol{X}^*$

For $t = 1: I_t$ do

$$\boldsymbol{W}_{ir} := \boldsymbol{W}_{ir}.\frac{(\boldsymbol{AX})_{ir}}{(\boldsymbol{WX^TX})_{ir}}$$

$$\tilde{\boldsymbol{X}}_{ir} := \tilde{\boldsymbol{X}}_{ir}.\frac{2(\boldsymbol{A}\tilde{\boldsymbol{X}}^T)_{ir}+(\boldsymbol{X}^*)_{ir}+4\eta(\tilde{\boldsymbol{X}})_{ir}}{2(\boldsymbol{B_1}\tilde{\boldsymbol{X}}^T)_{ir}+(\tilde{\boldsymbol{X}})_{ir}+4\eta(\tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\tilde{\boldsymbol{X}})_{ir}}$$

$$\boldsymbol{X}_{ir} := \boldsymbol{X}_{ir}.\frac{2(\boldsymbol{A}^T\boldsymbol{W})_{ir}+(\boldsymbol{X}^*)_{ir}}{2(\boldsymbol{XW}^T\boldsymbol{W})_{ir}+(\boldsymbol{X})_{ir}}$$

$$\boldsymbol{X}^*_{ir} := \boldsymbol{X}^*_{ir}.\frac{(\tilde{\boldsymbol{X}})_{ir}+(\boldsymbol{X})_{ir}+\boldsymbol{10^{-9}}}{2(\boldsymbol{X}^*)_{ir}+\boldsymbol{10^{-9}}}$$

$$\boldsymbol{X}^*{}_{ir} := \frac{\boldsymbol{X}^*{}_{ir}}{\sum_{r=1}^{c}\boldsymbol{X}^*{}_{ir}}$$

End for

Return $(v_i, \boldsymbol{I_i}) = \boldsymbol{argmax}_{r\le k}\boldsymbol{X}^*_{ir}$

Experiments and Analysis

In this section, the computational complexities of the proposed models are computed and compared, followed by discussing assessment standards for performance evaluation. Finally, other community detection methods are introduced to compare some popular network sets, and the results and capabilities of the proposed methods will be demonstrated accordingly.

Assessment Standards

In this paper, NMI and modularity index (Q) are used to evaluate the performance of different community detection methods. NMI information is widely applied to compare the similarity between partition labels and the ground truth partition labels. NMI information was adopted as:

$$NMI(C, C') = \frac{-2\sum_{i=1}^{|C|}\sum_{j=1}^{|C'|} n_{C_i\cap C'_j}\log\left(\frac{n_{C_i\cap C'_j}n}{n_{C_i}n_{C'_j}}\right)}{\sum_{i=1}^{|C|} n_{C_i}\log\left(\frac{n_{C_i}}{n}\right)+\sum_{j=1}^{|C'|} n_{C'_j}\log\left(\frac{n_{C'_j}}{n}\right)} \quad (39)$$

where n_{C_i} and $|C|$ indicate the number of members in partitions C_i and number of partitions in C, respectively. If NMI tends to one, the partition labels will be closer to ground truth partition labels, and if it tends to zero, the partition labels will be dissimilar to these labels.

Performance Analysis

The employed real-world and artificial networks are described in this subsection. The comparative results of our GMDNMF and MMNMF methods with other methods such as Mtrinmf, NMF, LPA, CNM, Infomap, MHGNMF, and LRSCD are illustrated on the networks. For more explanation, the Mtrinmf is one of the efficient methods that has been able to exploit network features such as modularity index to improve the performance of community detection based on the tri-NMF method [35]. The LPAM method modified the LPA method by adding the modularity index [13]. The CNM method is a fast-greedy optimization for directly solving the modularity index [37]. Infomap is a popular community detection method based on flow running dynamic by random walk [18]. MHGNMF is a new mixed hypergraph regularized NMF method which makes use of structure similarity information and topological connection information [27]. The MHGNMF method is divided into MHGNMF_kl and MHGNMF_sq algorithms based on the type of the community detection function. The MHGNMF_sq algorithm is selected due to the use of the Frobenius norm in the optimization function. According to [38], LRSCD is a community detection method based on a low-rank decomposition strategy for decomposing each node vector in a new space (the geometric space). The NMF, CNM, and Infomap are common community detection methods, and MHGNMF, LRSCD, and Mtrinmf methods are the recent community detection algorithms that have been considered for comparison.

Performance Analysis on Ten Real-World Networks

Ten real-world networks have been chosen to evaluate different community detection methods. The information

of these real-world networks has been tabulated Table 1. Here, $\bar{c}$ is the number of ground-truth communities. The Karate, Jazz, Political books, Dolphins, Football, and Polbooks are small real-world networks, while the Polblogs, Cora, Citeseer, and Pubmed are large real-world networks [27]. Table 2 lists the best results of eight methods on ten real-word networks based on the modularity index (Q). The following results are concluded based on data in Table 2:

- MMNMF, GMDNMF, and GMDNMF ($\lambda = \frac{1}{2}$) methods have better clustering capability compared to the NMF method based on the modularity index.
- In the Pubmed network, fast methods such as CNM and LPAM have better clustering in comparison with other NMF-based methods. Due to the computational errors in large-scale networks, NMF-based community detection usually has clustering errors. However, the GMDNMF offers better clustering when compared to other methods.
- Due to different values of λ, the GMDNMF method and some methods can offer the best community detection compared to other ones in other networks. For example, GMDNMF in the Cora network, and GMDNMF and MHGNMF_sq in the Polbooks network are the best methods for community detection.
- In addition to the GMDNMF method, MMNMF and MHGNMF_sq methods outperform other methods.

Table 1: Real-world network information

Networks	**N**	**m**	$\bar{\mathbf{c}}$	Description
Karate	34	78	2	Zachary karate club network (Karate) [39]
Jazz	198	2742	4	Jazz network (Jazz) [40]
Political books	105	441	3	Political books network (Political books) [41]
Dolphins	62	159	4	Lusseau's bottlenose dolphins social network (Dolphins) [42]
Football	115	613	12	American college football network [43]
Polblogs	1490	16718	2	Blogs about politics [44]
Cora	2708	5429	7	A Cora citation network [45]
Citeseer	3312	4732	6	A Citeseer citation network [46]
Pubmed	19717	44338	3	A Pubmed citation network [47]

Table 2: Modularity index (Q) for different methods and real-world networks

	GMDNMF	MMNMF	GMDNMF ($\boldsymbol{\lambda = 0.5}$)	MHGNMF_sq [27]	LRSCD [38]	NMF [12]	Mtrinmf [30]	Infomap [19]	LPAM [11]	CNM [37]
Karate	**0.419** ($0.5 < \lambda < 0.74$)	**0.419**	**0.419**	**0.419**	**0.419**	0.142	**0.419**	0.403	0.397	0.383
Jazz	**0.444** ($\lambda = 0.57$)	**0.444**	0.439	**0.444**	0.442	0.436	0.442	0.442	**0.444**	**0.444**
Political books	**0.526** ($\lambda = 0.39$)	0.520	0.520	**0.526**	0.520	0.513	**0.526**	**0.526**	0.520	0.508
Dolphins	**0.528** ($\lambda = 0.32, 0.28$)	0.526	0.520	0.526	0.526	0.514	0.526	0.520	0.518	0.498
Football	**0.605** ($\lambda = 0.73, 0.79$)	0.603	0.600	**0.605**	0.603	0.588	0.603	0.603	0.603	0.556
Polblogs	**0.427** ($\lambda = 0.7$,0.9)	0.425	0.425	0.425	0.425	0.424	0.425	0.423	0.425	**0.427**
Cora	**0.604** ($\lambda = 0.2$)	0.564	0.582	0.601	0.564	0.548	0.590	0.231	0.526	0.600
Citeseer	**0.798** ($\lambda = 0.3$,0.2)	0.776	0.691	0.712	0.629	0.576	0.621	**0.798**	0.551	0.724
Pubmed	0.641 ($\lambda = 0.8$)	0.594	0.567	0.581	0.473	0.523	0.438	0.726	0.44	**0.751**

Analysis for Two Artificial Networks

As a further performance investigation, MMNMF and GMDNMF methods are applied to Lancichinetti-Fortunato-Radicchi (LFR) [26] and Girvan-Newman (GN) [1] networks. The GN network is divided into four non-overlapping communities with 32 nodes in each community. The average degree of each node equals $Z_{in} + Z_{out} = 16$ where Z_{in} and Z_{out} denote the internal and external degrees of the nodes, respectively. LFR networks have some essential characteristics of networks, including power, low distribution of node degrees, and community size. The parameters of the generated LFR network are defined as follows:

The number of nodes is 700, and the average degree and max degree of the network are 20 and 50, respectively. Power law exponent for degree distributions is considered -3, and power-law distribution of community size is -1.

In addition, the community size ranges from 20 to 60 nodes, and the mixing parameter μ varies from 0.1 to 0.9. Moreover, ten independent experiments were executed for comparison.

The GMDNMF method clusters the networks according to the λ parameter. Thus, one way to improve the performance of this method is to choose the correct λ value.

To show the effect of choosing λ on the performance of the GMDNMF method, the networks were formed in terms of various Z_{out} and μ values. Each time, keeping Z_{out} and μ constant for choosing the best λ, we ran the algorithms with various λ values (0-1) in the steps of 0.01 and steps of 0.1 for GN and LFR networks, respectively. The NMI and Q information are depicted in Figs. 1 and 2. As shown, for a given Z_{out} and μ, different λ values cause different NMI and Q information evaluations. This procedure was repeated for different values of Z_{out} and μ, where Z_{out} and μ changed in steps 1 and 0.1, respectively.

Additionally, to demonstrate the effect of λ more clearly, an epigraph of Figs. 1 and 2 has been shown in Figs. 3 and 4. Based on Fig. 3, $\lambda = 0.87$ and $\lambda = 0.98$ maximize both Q ($Q = 0.222$) and NMI ($NMI = 0.942$) information on the GN network with $Z_{out} = 8$. Additionally, in Fig. 4, parameters $\lambda = 0.4$ and $\lambda = 0.9$ maximize Q ($Q = 0.124$) and NMI ($NMI = 0.132$) information on the LFR network with $\mu = 0.8$, respectively. Finally, the experimental results of the mentioned methods on GN and LFR networks are provided in Tables 3 and 4. Based on the obtained data, the following conclusions could be drawn:

- Compared to the NMF and Mtrinmf methods on LFR and GN networks, the proposed GMDNMF outperforms other methods for all μ and Z_{out} values. Moreover, the MMNMF method beats the NMF and Mtrinmf methods for the upper and middle values of μ and Z_{out}.
- In general, for lower values of μ and Z_{out}, the Infomap method outperforms other methods, but GMDNMF and MMNMF would be better for the upper and middle values of μ and Z_{out} compared to other methods.

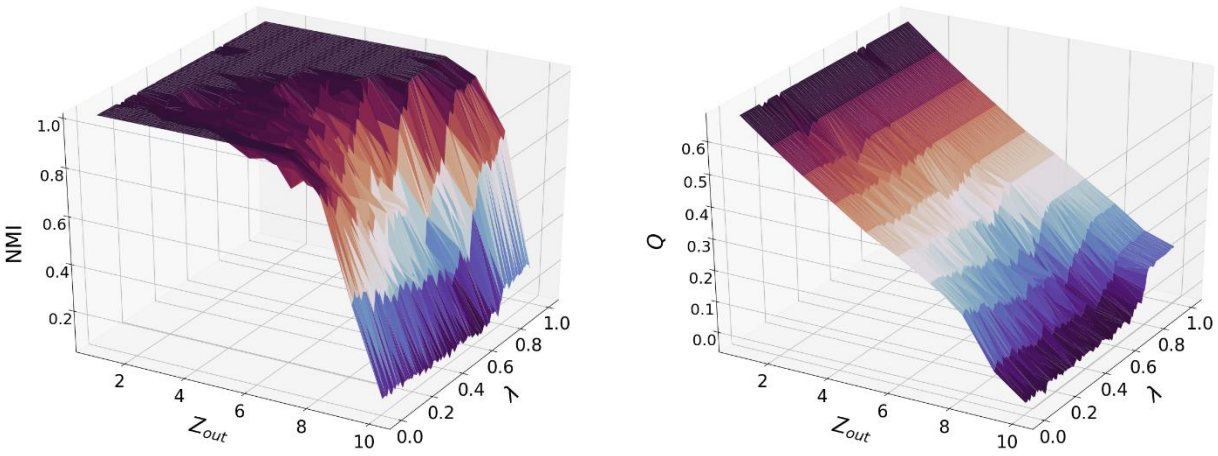

Fig. 1: The modularity index (Q) and NMI information for the GN network for different values of λ with the step of 0.01 and Z_{out} with the step of 1.

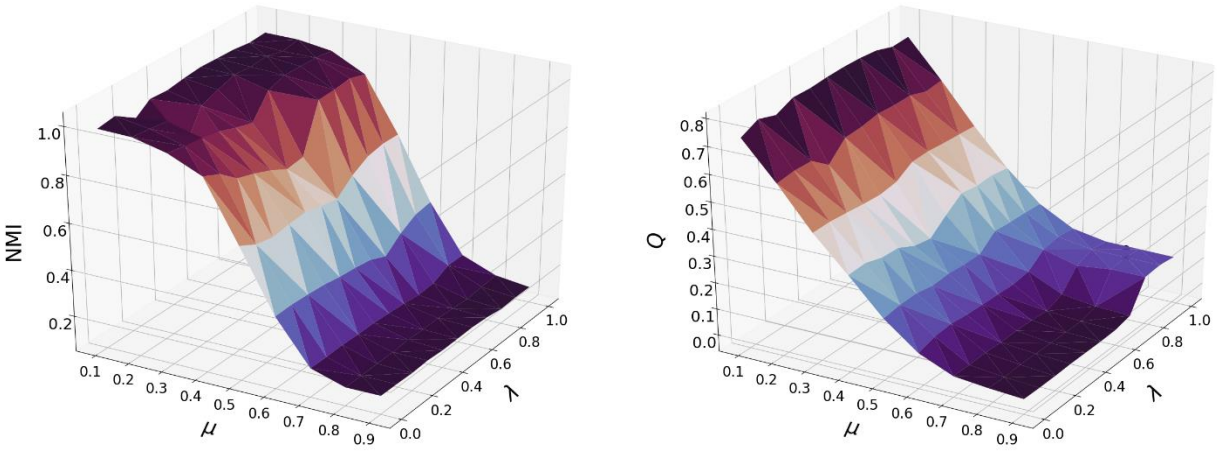

Fig. 2: The modularity index (Q) and NMI information for the LFR network for different values of λ and μ with the steps of 0.1.

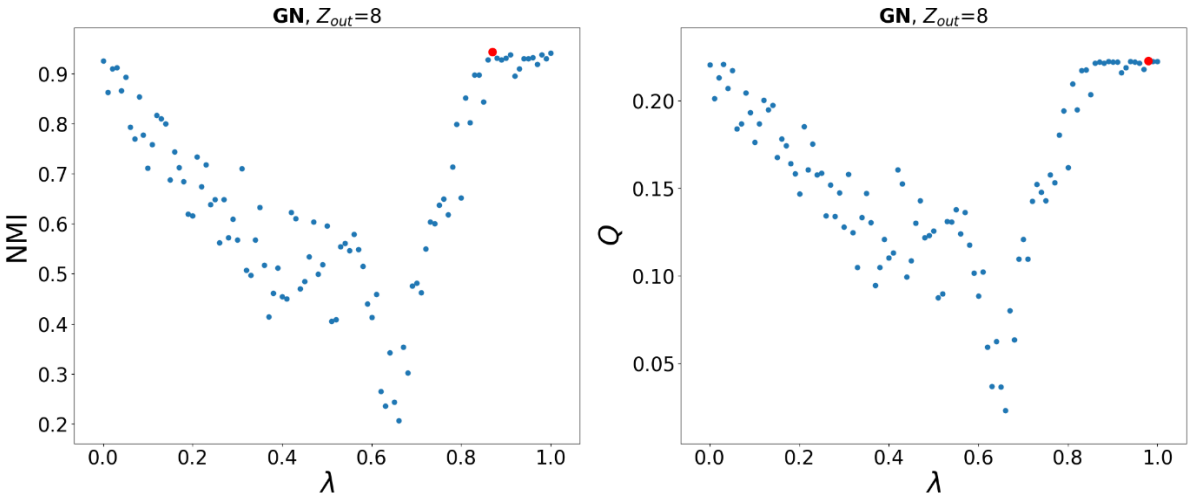

Fig. 3: The modularity index (Q) and NMI information for the GN network for different values of λ with the step of 0.01 and Z_{out} with the step of 1.

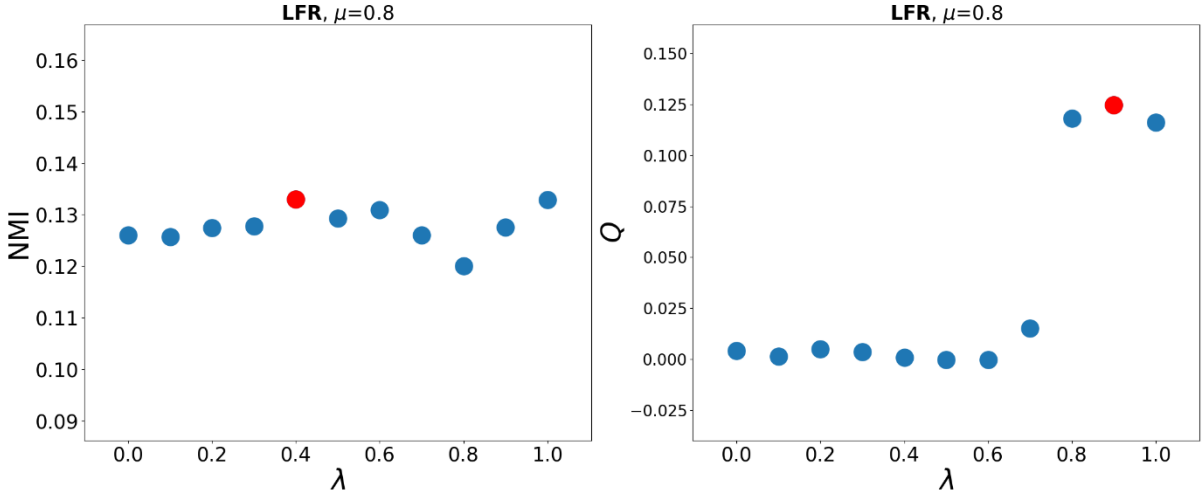

Fig. 4: The modularity index (Q) and NMI information for the LFR network for different values of λ and μ with the steps of 0.1.

Table 3: Comparison of different methods by mean of NMI information for 10 independent experiments on the GN network

	NMI								
Z_{out}	GMDNMF (variable λ)	MMNMF	GMDNMF ($\lambda = 0.5$)	MHGNMF_sq	LRSCD	NMF	Mtrinmf	Infomap	LPAM
4	**1**	**1**	0.98	**1**	**1**	**1**	**1**	**1**	**1**
5	**1**	**1**	**1**	0.99	**1**	**1**	**1**	**1**	**1**
6	**1**	**1**	0.99	**1**	**1**	**1**	**1**	0	1
7	**1**	**1**	0.925	0.99	**1**	**1**	**1**	0	0.98
8	0.97	**1**	0.925	0.64	0.89	0.89	0.64	0	0.96
9	**0.73**	0.704	0.47	0.41	0.47	0.41	0.33	0	0.38
10	**0.2**	0.1	0.18	0.054	0.16	0.04	0.054	0	0.16

Table 4: Comparison of different methods by mean of NMI information for 10 independent experiments on the LFR network

	NMI								
μ	GMDNMF	MMNMF	GMDNMF ($\lambda = 0.5$)	MHGNMF_sq	LRSCD	NMF	Mtrinmf	Infomap	LPAM
0.1	0.99	0.95	0.98	0.99	0.99	0.98	0.98	**1**	**1**
0.2	0.99	0.94	0.96	0.98	0.99	0.98	0.98	**1**	**1**
0.3	0.98	0.93	0.93	0.95	0.98	0.83	0.95	**1**	0.98
0.4	**0.92**	0.88	0.85	0.91	0.91	0.70	0.88	**0.92**	0.91
0.5	**0.66**	**0.66**	0.62	0.38	0.62	0.57	0.2	0	0.54
0.6	**0.39**	**0.39**	0.37	0.20	0.33	0.33	0.11	0	0.30
0.7	**0.19**	**0.19**	0.17	0.2	0.15	0.17	0.06	0	0.15
0.8	**0.13**	**0.13**	**0.13**	0.09	0.12	0.12	0.05	0	0.1
0.9	**0.12**	0.11	**0.12**	0.05	**0.12.**	**0.12**	0.05	0	0.09

Complexity Analysis and Comparison

In this subsection, the order of computational complexity would be computed for the proposed methods using O notation. For this purpose, first, the order of complexity of the main component would be evaluated, and the computational complexity of the two algorithms would be obtained accordingly.

In the MMNMF model, the computational complexity of updating rules is of $O(n^2k) + O(k^2n)$, $O(n^2k) + O(k^2n)$, $O(n^2k) + O(k^2n)$, and $O(n)$ for W, X, $\tilde{X}$ and X^*, respectively. Further, the computational complexity of MMNMF models is $O(n^2k) + O(k^2n) + O(n)$. Given that k is a small constant (i.e., $k \ll N$), the total complexity order is $O(n^2k)$.

The analysis of the GMDNMF model is similar to the MMNMF model. Therefore, the complexity orders of updating rules for W, $\tilde{X}$, and X^* are $O(n^2k) + O(k^2n)$, $O(n^2k) + O(kn) + O(k^2n)$, and $O(kn)$, respectively. Moreover, the updating rule for X has the order of $O(3n^2k) + O(k^2n) + O(k^2)$. Ultimately, the complexity order for all parameters in the GMDNMF model is $O(n^2k) + O(k^2n) + O(kn) + O(k^2)$. Considering that $k \ll N$, the total complexity for both models is $O(n^2k)$. If different λ values are examined seeking for the best results, the computational complexity will be $O(n^2kK_\lambda)$ where K_λ indicates the number of selected values for λ. The K_λ value can be selected with respect to a tradeoff between computational complexity and performance improvement. As the value of K_λ increases (e.g., $K_\lambda = 100$), better performance would be achieved, but with more computational complexity. Therefore, K_λ should be determined based on a compromise between performance and complexity.

Finally, if both algorithms converge after I_t iteration, the total complexity computation is $O(I_t n^2 k)$ in the MMNMF model and GMDNMF model with $\lambda = 0.5$. Furthermore, the total complexity for the K_λ values of different λ values in the GMDNMF model is $O(I_t n^2 k K_\lambda)$. Hence, the computational complexity of the GMDNMF model with respect to different λ values ($O(I_t n^2 k K_\lambda)$) is higher than that of the MMNMF model ($O(I_t n^2 k)$). Consequently, as pointed in previous studies [18], [25],

[4], [30] for other NMF-based community detection methods, the proposed methods are unsuitable for large-scale complex networks due to their complexity orders.

To compare the speed of different models, the run times of nine models on six real-world and two synthetic networks are recorded and brought in Tables 5 and 6. For this purpose, in Table 5 the proposed methods have been compared with other methods being NMF [12], LPAM [45], CNM [46] and Infomap [19]. From the results, it is clear that while the run times of all methods are of the same order of magnitude, but, the CNM model is faster than other models and LPAM model has more execution time. Hence, from the run time perspective, our algorithms are inferior to CNM but better than LPAM and Infomap models. Next, in Table 6 the run times of GMDNMF, MHGNMF_sq and Mtrinmf models are compared. For these methods, in contrast to methods in Table 5, parameter tuning is required, which is done generally by trial and error.

Therefore, Table 6 presents the run times of these models for the selection of 20 different values for the internal parameters.

The results in Table 6 show that GMDNMF is faster than the Mtrinmf, but slower than MHGNMF_sq models for 20 different values of the internal parameters. Hence, from NMF-based community detection models, the run time of MMNMF model is quite near to the NMF model and GMDNMF model is near to MHGNMF_sq and Mtrinmf methods. In summary, according to the results of Tables 1 to 6, our models are more flexible, less sensitive with better performance respecting to other models by utilizing the information and characteristics of the networks such as general modularity density and modularity indices, while their run times remain acceptable.

The machine used for the present study is powered Intel Core i7-6770 CPU and 16 GB RAM with 64-bit Windows 10, and Python (version 3.8) as the selected software.

Table 5: Comparison of run times (seconds) for different models and sets

	Karate	Jazz	Political books	Dolphins	Football	Polblogs	Cora	Citeseer	Pubmed	GN	LFR
MMNMF	0.181	0.93	0.85	0.27	0.83	3.78	13.76	15.09	3427	0.297	4.75
GMDNMF ($\boldsymbol{\lambda = 0.5}$)	0.196	0.87	0.83	0.27	0.89	4.02	13.89	18.09	3731	0.354	5.65
NMF [12]	0.175	0.89	0.82	0.27	0.73	3.32	13.06	13.29	2953	0.290	5.05
Infomap [19]	0.203	0.91	0.86	0.28	0.97	4.13	14.7	16.6	4030	0.449	5.10
LPAM [45]	0.428	2.18	0.98	0.71	2.08	13.12	66.8	76.2	5837	0.891	24.84
CNM [46]	0.118	0.80	0.63	0.18	0.67	2.56	9.3	11.7	1916	0.127	3.12

Table 6: Comparison of run times (seconds) of models for different sets with 20 various internal parameter values

	Karate	Jazz	Political books	Dolphins	Football	Polblogs	Cora	Citeseer	Pubmed	GN	LFR
GMDNMF	4.5	20.2	27.1	6.3	14.7	97.1	189.1	365.3	28875	7.3	102.9
MHGNMF_sq [27]	3.9	15.3	14.8	4.4	15.3	76.9	178.0	347.1	26901	6.8	93.8
Mtrinmf [30]	4.8	21.8	28.7	6.6	16.6	112.7	200.5	420.4	32510	8.2	120.4

Conclusion

In this paper, MMNMF and GMDNMF were presented as two novel NMF-based community detection methods to identify the best communities in complex networks. To this end, it was proved that the modularity/general modularity density-based community detection could be consistently represented in the form of SNMF-based community detection. This consistent representation helped combine NMF-based community detection with modularity/general modularity density-based community detection approaches. The proposed MMNMF model improved the performance of community detection

based on NMF by employing the modularity index as the network feature for the NMF model. The proposed GMDNMF model could enhance NMF-based community detection using the general modularity density index. Iterative update rules were derived as an optimal solution for solving MMNMF and GMDNMF optimization models. The performances of the two models were verified on various artificial and real-world networks of different sizes. According to the results, MMNMF and GMDNMF performed better than the other community detection methods. Additionally, the GMDNMF model had higher computational complexity compared to the MMNMF model, but it outperformed this model.

As future works, the proposed MMNMF and GMDNMF can be extended for NMF-based community detection in multi-layer networks. These proposed models may improve the performance of the multi-view clustering method for community detection by combining link and content information.

Author Contributions

M. ghadirian designed and simulated the proposed method and wrote the manuscript. N. Bigdeli chose strategies, analyzed the results, edited the manuscript, and managed the entire process.

Acknowledgment

The author would like to thank the editor and reviewers for their helpful comments

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues including plagiarism, informed consent, misconduct, data fabrication and, or falsification, double publication and, or submission, and redundancy have been completely witnessed by the authors.

Abbreviations

NMI	Normalized Mutual Information
IP	Integer Programming
NMF	Nonnegative Matrix Factorization
GNMF	Graph Regularized Nonnegative Matrix Factorization
DNMF	Deep Nonnegative Matrix Factorization
MDNMF	Modularized Deep Nonnegative Matrix Factorization
tri-NMF	Tri-factor Nonnegative Matrix Factorization
Mtrinmf	Modularized Tri-factor Nonnegative Matrix Factorization
SNMF	Symmetric Nonnegative Matrix Factorization
MMNMF	Mixed Modularity Nonnegative Matrix Factorization
GMDNMF	General Modularity Density Nonnegative Matrix Factorization
SVDCNMF	Singular-value Decomposition Community Detection Nonnegative Matrix Factorization
RSSNMF	Robust Semi-supervised Nonnegative Matrix Factorization
CNM	Clauset-Newman-Moore
MHGNMF	Mixed hypergraph Nonnegative Matrix Factorization
LRSCD	Low-rank Subspace Learning-based Network Community Detection
LFR	Lancichinetti–Fortunato–Radicchi
GN	Girvan–Newman

References

[1] M. E. J. Newman, “Networks,” OUP, 2018.

[2] P. Bedi, C. Sharma, “Community detection in social networks,” Wiley Interdiscip. Rev.: Data Min. Knowl. Discovery, 6 (3): 115-135, 2016.

[3] Z. Li, S. Zhang, R. S. Wang, X. S. Zhang, L. Chen, “Quantitative function for community detection,” Phys. Rev. E, 77 (3): 036109, 2008.

[4] L. H. N. Lorena, M. G. Quiles, L. A. N. Lorena, “Improving the performance of an integer linear programming community detection algorithm through clique filtering,” in Proc. International Conference on Computational Science and Its Applications (ICCSA): 757–769, 2019.

[5] M. Sathyakala, M. A. Sangeetha, “Weak clique based multi objective genetic algorithm for overlapping community detection in complex networks,” J. Ambient Intell. Humaniz. Comput. 12: 6761–6771, 2021.

[6] M. Mohammadi, M. Fazeli, M. Hosseinzadeh, "Parallel louvain community detection algorithm based on dynamic thread assignment on graphic processing unit", J. Electr. Comput. Eng. Innovations (JECEI), 10(1): 75-88, 2022.

[7] C. K. Tsung, S. L. Lee, H. J. Ho, S. Chou, “A modularity-maximization-based approach for detecting multi-communities in social networks,” Ann. Oper. Res., 303: 381–411, 2021.

[8] S. Muff, F. Rao, A. Caflisch, “Local modularity measure for network clusterizations”, Phys. Rev. E, 72: 056107, 2005.

[9] K. Sato, Y. Izunaga, “An enhanced MILP-based branch-and-price approach to modularity density maximization on graphs,” Comput. Oper. Res., 106: 236–245, 2019.

[10] J. Liu, J. Zeng, “Community detection based on modularity density and genetic algorithm,” in Proc. 2010 International Conference on Computational Aspects of Social Networks: 29-32, 2010.

[11] M. Li, J. Liu, “A link clustering based memetic algorithm for overlapping community detection,” Phys. A: Stat. Mech. Appl., 503: 410–423, 2018.

[12] A. Costa, "MILP formulations for the modularity density maximization problem," Eur. J. Oper. Res., 245(1): 14–21, 2015.

[13] M. J. Barber, J. W. Clark, "Detecting network communities by propagating labels under constraints," Phys. Rev. E, 80: 026129, 2009.

[14] Q. Wu, R. Chen, L. Wang, K. Guo, "A label propagation algorithm for community detection on high-mixed networks," Concurr. Comput. Pract. Exp., 33 (9): e6141, 2020.

[15] M. Rosvall, C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," Proc. Natl. Acad. Sci. U.S.A., 105 (4): 1118–1123, 2008.

[16] J. Zhou, L. Li, A. Zeng, Y. Fan, Z. Di, "Random walk on signed networks," Phys. A: Stat. Mech. Appl., 508: 558-556, 2018.

[17] C. Liu, F. Huang, R. Li, Q. Yang, Y. Li, S. Yu, "Community detection using multitopology and attributes in social networks," Concurr. Comput. Pract. Exp., 34 (12): e6028, 2020.

[18] R. S. Wang, S. Zhang, Y. Wang, X. Zhang, L. Chen, "Clustering complex networks and biological networks by nonnegative matrix factorization with various similarity measures," Neurocomputing 72 (1-3): 134-141, 2008.

[19] L. Xu, T. Ming, W. Xiaofei, W. Chao, F. Qiang, Y. Yonghong, "Single-channel speech separation based on non-negative matrix factorization and factorial conditional random field," Chin. J. Electron., 27 (5): 1063-1070, 2018.

[20] S. Peng, W. Ser, B. Chen, Z. Lin, "Robust semi-supervised nonnegative matrix factorization for image clustering," Pattern Recognit., 111: 107683, 2021.

[21] S. Zhang, G. Zhang, F. Li, C. Deng, S. Wang, A. Plaza, J. Li, "Spectral-spatial hyperspectral unmixing using nonnegative matrix factorization," IEEE Geosci. Remote. Sens., 60: 5505713, 2021.

[22] E. L. Lydia, P. K. Kumar, K. Kumar, S. K. Lakshmanaprabu, R. M. Vidhyavathi, "Charismatic Document clustering through novel K-Means Non-negative Matrix Factorization (KNMF) Algorithm using key phrase extraction," Int. J. Parallel Program., 48: 496–514, 2020.

[23] C. He, Y. Tang, K. Liu, H. Li, S. Liu, "A robust multi-view clustering method for community detection combining link and content information," Phys. A: Stat. Mech. Appl., 514: 396-411, 2018.

[24] K. Shu, S. Wang, H. Liu, "Beyond news contents: the role of social context for Fake news detection," in Proc. Twelfth ACM International Conference on Web Search and Data Mining: 312-320, 2019.

[25] C. He, Q. Z, Y. Tang, S. Liu, J. Zheng, "Community detection method based on robust semi-supervised nonnegative matrix factorization," Phys. A: Stat. Mech. Appl., 523: 279-291, 2019.

[26] H. Lu, X. Sang, Q. Zhoa, J. Lu, "Community detection algorithm based on nonnegative matrix factorization and pairwise constraints," Phys. A: Stat. Mech. Appl., 522: 205-214, 2019.

[27] W. Wu, S. Kwong, Y. Zhou, Y. Jia, W. Gao, "Nonnegative matrix factorization with mixed hypergraph regularization for community detection," Inf. Sci., 435: 263-281, 2018.

[28] M. Zhang, Z. Zhou, "Structural deep nonnegative matrix factorization for community detection," Appl. Soft Comput., 97: 106846, 2020.

[29] J. Huang, T. Zhang, W. Yu, J. Zhu, E. Cai, "Community detection based on modularized deep nonnegative matrix factorization," Int. J. Pattern Recognit. Artif. Intell., 2 (35): 2159006, 2021.

[30] C. Yan, Z. Chang, "Modularized tri-factor nonnegative matrix factorization for community detection enhancement," Phys. A: Stat. Mech. Appl., 533: 122050, 2019.

[31] X. Ma, L. Gao, L. Fu, X. Yong, "Semi-supervised clustering algorithm for community structure detection in complex networks," Phys. A: Stat. Mech. Appl., 389(1): 187–197, 2010.

[32] X. Wang, P. Cui, J. Wang, J. pei, W. Zhu, S. Yang, "Community preserving network embedding," in Proc. Thirty-First AAAI Conference on Artificial Intelligence: 203-209, 2017.

[33] X. Ma, D. Dong, Q. Wang, "Community detection in multi-layer networks using joint nonnegative matrix factorization," IEEE Trans. Knowl. Data. Eng., 31 (2): 273-286, 2019.

[34] L. Zong, Z. Zhang, L. Zhao, H. Yu, Q. Zhao, "Multi-view clustering via multi-manifold regularized non-negative matrix factorization," Neural Netw., 88: 74-89, 2017.

[35] S. Peng, W. Ser, B. Chen, Z. Lin, "Robust orthogonal nonnegative matrix tri-factorization for data representation," Knowl-Based Syst., 201-202: 106054, 2020.

[36] J. Liu, C. Wang, J. Gao, J. Han, "Multi-view clustering via Joint nonnegative matrix factorization," in Proc. the 2013 SIAM International Conference on Data Mining: 252-260, 2013.

[37] A. Clauset, M. E. J. Newman, C. Moore, "Finding community structure in very large networks," Phys. Rev. E, 70(6): 066111, 2004.

[38] Z. Ding, Z. Shang, D. Sun, B. Luo, "Low-rank subspace learning based network community detection," Knowl-Based Syst., 155: 71-82, 2018.

[39] W. W. Zachary, "An information flow model for conflict and fission in small groups," J. Anthropol. Res., 33 (4): 452–473, 1977.

[40] P. M. Gleiser, L. Danon, "Community structure in jazz," Adv. Compl. Syst., 6 (4): 565–573, 2003

[41] J. Kunegis, "KONECT: The koblenz network collection,"in Proc. 22nd International Conference on World Wide Web: 1343–1350, 2013.

[42] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, S. M. Dawson, "The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations," Behav. Ecol. Sociobiol., 54 (4): 396–405, 2003.

[43] A. Lancichinetti, S. Fortunato, F. Radicchi, "Benchmark graphs for testing community detection algorithms", Phys. Rev. E, 78 (4): 046110, 2008.

[44] L. Yang, X. Cao, D. Jin, X. Wang, D. Meng, "A unified semi-supervised community detection framework using latent space graph regularization," IEEE Trans. Cyber., 45(11): 2585–2598, 2015.

[45] L. A. Adamic, N. Glance, "The political blogosphere and the 2004 us election: divided they blog,"in Proc. 3 the 3rd international workshop on Link discovery: 36–43, 2005.

[46] D. He, Z. Feng, D. Jin, X. Wang, W. Zhang, "Joint identification of network communities and semantics via integrative modeling of network topologies and node contents," in Proc. the Thirty-First AAAI Conference on Artificial Intelligence: 116–124, 2017.

[47] G. Namata, B. London, L. Getoor, B. Huang, U. EDU, "Query-driven active surveying for collective classification," in Proc. 10th Workshop on Mining and Learning with Graphs, 8 ,2012.

Biographies

Mohammad Ghadirian was born in Iran, in 1992, He received M.Sc. degree in Electrical Engineering Majoring in Control from the Sharif University of Technology, Tehran, Iran in 2016. He is already Ph.D. candidate in Electrical Engineering Department of Imam Khomeini International University, Qazvin, Iran. His research interest includes in graph mining, data mining and medical image processing.

- Email: s956191004@edu.ikiu.ac.ir
- ORCID: ID: 0000-0002-4106-2406
- Web of Science Researcher ID: NA
- Scopus Author ID: NA
- Homepage: NA

Nooshin Bigdeli was born in 1977 in Iran, and completed her Ph.D. degree in Electrical Engineering majoring in Control at Sharif University of Technology, Tehran, Iran in 2007. She is currently professor of Electrical Engineering Department of Imam Khomeini International University, Qazvin, Iran. Her research interests include control systems, applied optimization, intelligent systems, model predictive control as well as model order reduction in high order systems.

- Email: n.bigdeli@eng.ikiu.ac.ir
- ORCID: 0000-0001-5536-4491
- Web of Science Researcher ID: AAT-8622-2021
- Scopus Author ID: 8528681600
- Homepage: http://www.ikiu.ac.ir/members/?id=23&lang=0

How to cite this paper:
M. Ghadirian, N. Bigdeli, "A new hybrid nmf-based infrastructure for community detection in complex networks," J. Electr. Comput. Eng. Innovations, 11(2): 443-458, 2023.

DOI: 10.22061/jecei.2023.9150.577

URL: http://jecei.sru.ac.ir/article_1879.html