Community Detection in Temporal Networks Using Triple Nonnegative Matrix Factorization

Hai-fu LIU¹,* and Li-meng-zi YUAN²

¹Tianjin Modern Innovative TCM Technology Co., Ltd., Huayuan Industrial Park, Tianjin, China
²School of Computer Science and Technology, Tianjin University, Tianjin, China
*Corresponding author

Keywords: Community detection, Temporal networks, Node Weight Matrices.

Abstract. Analyzing temporal networks can uncover dynamic evolution and characterize the properties of the networks. This paper proposes a novel temporal community detection model using triple nonnegative matrix factorization. Node weight matrices are introduced for targeting central nodes of communities and reducing number of nodes that have unobvious propensities of belonging to communities, which improves the algorithm performance of community detection. Community membership temporal smoothness constraint is added to discover latent structure and evolutionary behaviors of temporal networks. We then propose a gradient descent algorithm to optimize objective function. Experimental results on synthetic and real benchmarked networks show the effectiveness of detecting communities and finding their temporal changes.

Introduction

Community structure is one of the most significant properties that reflect exactly the essences of a network. A network is said to have community structure if the vertices in networks often cluster into tightly knit groups with a high density of within-group edges and a lower density of between-group edges [1]. Particularly in the temporal networks, the structure of a real network is the result of the continuous evolution of the forces that formed it [2], which simulates number of researchers to make great efforts on analyzing community structure and temporal dynamics.

Despite a large arsenal of powerful community detection methods have been proposed for static networks, almost all networks change over time, that motivates a body of new work of dynamic community detection. Compared with static techniques, dynamic community detection methods aim at identifying how communities emerge, grow, combine and decay over time. Among these methods, some of them study communities and their evolutions separately, some of them are not able to maintain the quality of community detection, or others are not capable of guaranteeing robustness and effectiveness of the algorithm, which are inappropriate in applications with noisy temporal networks.

In this paper, we propose a novel model which is capable of detecting communities and analyze their temporal evolutions using triple nonnegative matrix factorization (Tri-NMF). The main contributions that we make are as follows:

- We propose a unified TTNMF which can detect and track dynamic communities. With the added temporal smoothness constraint of community membership, the inherent structure and dynamic units of temporal networks are uncovered. The evolutionary behaviors of dynamic communities that captured by our model provide useful information for predicting evolution tendency of temporal networks.
- By introducing node weight matrices, our model can target central nodes of communities and reduce number of nodes that have unobvious propensities of belonging to communities, which improves the algorithm performance of community detection.
- An optimal gradient descent algorithm is proposed to solve the obtained objective function. Extensive experimental studies on several synthetic benchmarks and real-world networks
demonstrate that our model maintain the quality of community detection, track the temporal evolutions better and considerably improve the computational efficiency as a result.

The rest of the paper is organized as follows. We review related work of dynamic community detection methods of low rank approximation in Section of Related Work. In Section of Model Formulation, we describe the proposed model. Experimental results performed on synthetic and real-world data are presented in Section of Experiments and Results. The conclusions and discussions follow in Section of Conclusions and Discussions.

Related Work

In this section, we provide some essential background on dynamic community detection methods of low rank approximation. Methods concerned constructing structured low rank approximation include matrix factorization, spectral clustering and stochastic block model (SBM), and we will mainly review related work of the three methods in the literature in the following.

The methods of matrix factorizations are widely applied for exploration and time-varying community detection in time-evolving graph sequences. The most common factorization is the Singular Value Decomposition (SVD), which has important connections to community detection, graph drawing, and areas of statistics and signal processing [3]. For instance in classical spectral layout, the coordinates of each node are given by the SVD of graph related matrices, and can be calculated efficiently using algorithms in [4]. Recently, there has been extensive interest in spectral clustering [5], which aims to discover community structure in eigenvectors of the graph Laplacian matrix. Low-rank approximations which composed of nonnegative entries, referred to as NMF, have been shown to be advantageous for visualization of non-negative data [6]. Non-negativity is typically satisfied with networks, as edges commonly correspond to flows, capacity, or binary relationships, most of which are non-negative. In addition, theoretical connections between NMF and important problems in data mining have been developed [7], and accordingly, NMF has been proposed for overlapping community detection on static [8] and dynamic [9] networks.

There has been significant research of dynamic extensions of stochastic block models dedicated to statistical modeling of dynamic networks, mostly in the past several years. Xing et al. [10] and Ho et al. [11] proposed dynamic extensions of a mixed-membership version of the SBM. Ishiguro et al. [12] proposed a dynamic extension of the infinite relation model, which is a nonparametric version of the SBM. Yang et al. [13] proposed an HM-SBM that posits a Markov model on the class membership vectors parameterized by a transition matrix. Xu and Hero [14] proposed an HM-SBM that places a state-space model on the block probability matrices.

Model Formulation

In this section, we first provide some notations and definitions on nonnegative matrix factorization used in this work, and then describe the new model.

Notation

In this paper, bold uppercase letters will donate matrices, e.g. $\mathbf{X}$, bold lowercase letters will donate column vectors, e.g. $\mathbf{x}$, while operators $(\cdot)^T$ will stand for matrix transposition, e.g. $\mathbf{X}^T$. Both $x_{ij}$ and $(\mathbf{X})_{ij}$ represent the Entry $(i, j)$ of the matrix $\mathbf{X}$. The Frobenius norms will be represented by $\|\cdot\|_F$. Considering a dynamic $N$-node network whose time-varying structures are captured by the time-series adjacency matrices $[\mathbf{A}^t \in \mathbb{R}^{N \times N} ]_{t=1}^T$. $d_{ij}^t$ is one if there is an edge from node $i$ to node $j$ at time $t$, and is zero otherwise. We assume that the dynamic network is undirected, i.e. $d_{ij}^t = d_{ji}^t$, and there are no self-edges, i.e. $d_{ii}^t = 0$. 
The Unified TTNMF Model Formulation

This section describes the Temporal Triple Nonnegative Matrix Factorization (TTNMF) model. We introduce the snapshot cost of modeling network topologies and temporal cost of smoothness constraint. In this model, the communities not only generate evolutions but also are regularized so that dramatic change is unlikely.

Considering the observed network at time $t$, denoted by $A_t^t$, the nonnegative data matrix $A_t^i$ can be factorized into matrices $F^G T$, i.e., $A_t^i \approx F^G T$, with the constraints that $F^G$ and $G^T$ are nonnegative. In the factorization, $F^G$ can be considered to be a centroid matrix as each column represent a community central node, while $G^T$ can be considered to be a community membership matrix with $G_{ij}^T$ denoting the probability that the node $i$ belongs to the community $j$. Then we define the central nodes of all communities in dynamic network as $\{F^G_t \in \mathbb{R}^{N \times K}_t\}^T_{t=1}$ and the community memberships of all nodes in dynamic network as $\{G^T_t \in \mathbb{R}^{N \times K}_t\}^T_{t=1}$ where $K$ represents the number of communities of the dynamic networks.

While, in the above factorization, there are no constraints on the central node column vectors. To achieve good interpretability of obtained matrices, consider the dynamic network at time $t$, we impose the constraint that the column vectors defining $F^G_t$ lie within the column space of data matrix $A_t^i$:

$$F^G_t = \sum_{n} w_{n} a_n^i = A_t^i W^i$$

Moreover, for reasons of obtaining good interpretability, we then restrict to convex combinations of the columns of $A_t^i$. The advantage of imposing the convex constraint is that the columns can be interpreted as weighted sums of certain nodes; these columns would capture the notion of central nodes in communities; in addition, number of nodes that have unobvious propensities of belonging to communities would decrease. In practice, movements of central nodes often influence the movements of nodes who have close relationship with them. Detecting central nodes in communities accurately is critical for analyzing dynamic evolutions of temporal networks. We define the node weight matrices of all nodes in dynamic network as $\{W^i \in \mathbb{R}^{N \times K}_t\}^T_{t=1}$. As a result, we can reconstruct the topology of network at time $t$ as $A_t^i \approx A_t^i W^i G^T_t$.

To regularize the community structure so that it is less likely for unreasonably dramatic changes in terms of the community memberships from time $t-1$ to $t$, we impose the temporal smoothness constraints on community memberships matrices. We define the temporal cost as the difference between the community membership matrices at time $t-1$ and that at time $t$.

Considering the snapshot cost of modeling network topologies and temporal cost of smoothness constraint, we define the cost function as the sum of community detection quality and historical cost. To achieve smooth temporal community detection, we solve this by maximizing the community detection quality of current time-stamp and minimizing the historical cost, then we have following function:

$$\min_{A^t \geq 0, G^t \geq 0} \|A^t - A_t^i W^i G^T_t\|_F^2 + \alpha \|G^t - G^t_{t-1}\|_F^2 .$$

where $\alpha$ is a temporal smoothness parameter that trades off between the first and second term of the objective function.

To solve objective function in (2), we propose a gradient algorithm using following update rules obtained by auxiliary functions.

The update rule for $W^i_t$ is as follows:

$$w_{ij}^t \leftarrow w_{ij}^t \left( \frac{A^T_t A^i G^t_t)_{ij}}{A^T_t A^i W^i_t G^T_t G^t_t)_{ij}} \right) .$$

The update rule for $G^t_t$ is as follows:
\[
g_{ij}^t \leftarrow \frac{(A^T A')^t W^t + \alpha A_{-1}^T A_{-1}^t W_{-1}^t)_{ij}}{(G^T W^t A^T A' W^t + \alpha G_{-1}^T W_{-1}^t A_{-1}^t W_{-1}^t)_{ij}}.
\] (4)

Moreover, our model is easily extended to the networks whose number of nodes and communities may change over time, which is a very common phenomenon in temporal networks. The emergence, death, split or merger of communities can be detected by analyzing community membership matrices after we modify our model.

**Experiments and Results**

In this section, we present the experiment results of our proposed algorithm running on one synthetic dataset and one real-world dataset. We compare the results of our methods with three popular methods: SNMF [15], Facenet [9] and genLouvain algorithm [16]. First, we introduce the two evaluation measures we used and then present experiment results of different datasets. As our algorithm is not sensitive to the parameter, we set the temporal smoothness parameter \( \alpha \) as 0.2. In addition, we ran each method 20 times and report the average results.

**Evaluation Measures**

Two metrics including Fscore and Normal Mutual Information (NMI) are used to evaluate the performance of our algorithm. We briefly introduce the computational process of the two metrics.

Fscore combines the information of precision and recall which is extensively applied in evaluating the community detection result [17]. The precision and recall are calculated as:

\[
\text{Precision}(C_q, C'_p) = \frac{n_{p,q}}{C'_p}.
\] (5)

\[
\text{Recall}(C_q, C'_p) = \frac{n_{p,q}}{C_q}.
\] (6)

Then the Fscore of the detected community \( C'_p \) and the real community \( C_q \) can be computed as:

\[
F(C_q, C'_p) = \frac{2 \times P(C_q, C'_p) \times R(C_q, C'_p)}{P(C_q, C'_p) + R(C_q, C'_p)}.
\] (7)

Normalized mutual information (NMI) is an increasingly popular measure of clustering quality [17], which can be formulated as:

\[
\text{NMI} = \frac{\sum_{p=1}^{K} \sum_{q=1}^{K} n_{p,q} \log \frac{n_{p,q}}{n_{p}n_{q}}}{\sqrt{\left(\sum_{p=1}^{K} n_{p} \log \frac{n_{p}}{n}\right)\left(\sum_{q=1}^{K} n_{q} \log \frac{n_{q}}{n}\right)}}.
\] (8)

**Synthetic Dataset: Grow-shrink benchmark**

We generate two networks with 21 time steps according to the benchmark model [18] and evaluate our results on these two dynamic networks. This benchmark models the movement of nodes from one community to another. \( \mu \) is the ration of each community’s nodes switching to other communities, which influences the degree of dynamics for networks. We generate each network with N nodes in 10 runs and for \( \mu = 0.25 \) and \( \mu = 0.26 \), respectively; \( N = 256 \). It can be known from figure 1 and figure 2 that our proposed algorithms outperforms other three algorithms.
Real-world Dataset: KIT-Email

To evaluate the performance of our proposed method, we ran our algorithm on the e-mail communication network in the Department of Informatics at KIT. The network include e-mail contacts of the department which changes from September 2006 to August 2010. The edges are sparse, so we construct the adjacency matrices among 210 active members. In this network, the ground truth of communities are the departments of computer science at KIT. We process the snapshots and get two networks whose number of communities is 25 and 27, for the snapshots of 4 and 6 months, respectively. Figure 3 shows the average values of Fscore for the snapshots of 4 months, and figure 4 shows the average values of Fscore for the snapshots of 6 months.
Conclusions and Discussions
In this paper, we present a unified model named TTNMF which can detect communities and track their evolutions. Because of introducing node weight matrices and imposing temporal smoothness constraint, the performance of community detection is significantly improved. Compared with state-of-art methods, our method employs the history information more effectively in the process of analyzing dynamics. The experimental results on the synthetic and real data show that TTNMF outperforms other popular methods in community evolution. The automatic determination of the number of communities at each time step in a temporal network becomes a difficult problem to solve recently, and it will be our future work. In addition, we plan to extend our model to networks whose number of nodes and communities may change over time.

References

