PHYSA: 17524 Model 3Gsc pp. 1–19 (col. fig: NIL)

Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Q1 Evolutionary link community structure discovery in dynamic weighted networks

Q2 Qiang Liu [a,*], Caihong Liu [b,c], Jiajia Wang [d], Xiang Wang [a], Bin Zhou [a,e], Peng Zou [a]

[a] College of Computer, National University of Defense Technology, Changsha 410073, China
[b] College of Software, Dalian University of Foreign Languages, Dalian 116044, China
[c] Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, China
[d] Production Management Section, Dongfeng Nissan Dalian Branch Company, Dalian 116600, China
[e] State Key Laboratory of High Performance Computing, National University of Defense Technology, Changsha 410073, China

HIGHLIGHTS

- We propose a novel evolutionary link community structure discovery (ELCSD) algorithm in dynamic weighted networks based on link analysis.
- The ELCSD can detect evolutionary link community structure when the number of nodes and edges, the strength of edges and the number of link communities are changing with time passing.
- The ELCSD is based on a local link expansion strategy, which can be employed to deal with the condition that acquiring the global network structure information is difficult.

ARTICLE INFO

Article history:
Received 21 June 2015
Received in revised form 9 June 2016
Available online xxxx

Keywords:
Link community structure
Dynamic networks
Weighted networks
Community evolution

ABSTRACT

Traditional community detection methods are often restricted in static network analysis. In fact, most of networks in real world obviously show dynamic characteristics with time passing. In this paper, we design a link community structure discovery algorithm in dynamic weighted networks, which can not only reveal the evolutionary link community structure, but also detect overlapping communities by mapping link communities to node communities. Meanwhile, our algorithm can also get the hierarchical structure of link communities by tuning a parameter. The proposed algorithm is based on weighted edge fitness and weighted partition density so as to determine whether to add a link to a community and whether to merge two communities to form a new link community. Experiments on both synthetic and real world networks demonstrate the proposed algorithm can detect evolutionary link community structure in dynamic weighted networks effectively.

1. Introduction

Recently, an increasing number of researchers pay attention to community structure identification [1–5], such as social network, biology network, and World Wide Webs. A complex network is a representation of complex system in real life Q3

* Corresponding author.
E-mail address: liuqiang1981@nudt.edu.cn (Q. Liu).

http://dx.doi.org/10.1016/j.physa.2016.09.028
0378-4371/

Please cite this article in press as: Q. Liu, et al., Evolutionary link community structure discovery in dynamic weighted networks, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2016.09.028

in the form of nodes and edges, where nodes represent objects and edges represent their mutual interactions respectively. Finding relationship between these nodes and edges plays an important role in shedding light on the overall organization of complex systems and their functional properties. Therefore, detecting communities in networks has become a basic problem in the field of network science.

Conventional community detection methods are mainly restricted in analyzing the static networks, such as spectral analysis [6], GN [1], hierarchical clustering [7] and so on. However, the structure of network may evolve over time in real world. For instance, in scientific research cooperation networks, the cooperating relations between each node often change with time passing. Some nodes may construct new relations or increase strength of cooperation with other nodes, while relations of other nodes may disappear due to their cooperation finishing. Therefore, an increasing number of scholars engage in researches of dynamic network community detection so as to have a better understanding of network structure with time passing.

At present, there are mainly two categories for detecting community structure in dynamic networks. One category is based on a two-stage approach to analyze community evolutions. Firstly, communities are detected for each timestep, and then compared to determine correspondences for discovering steady community structure. For instance, Hopcroft et al. [8] proposed an agglomerative clustering method based on cosine similarity of nodes at different snapshots to tracking evolving communities in large linked networks. Palla et al. [9] developed a new algorithm based on clique percolation, that allows, for the first time, to investigate the time dependence of overlapping communities on a large scale and as such, to disclose basic relationships characterizing community evolution. Duan et al. [10] put forward a community mining algorithm including community discovery and change-point detection on dynamic weighted directed graphs. Takaffoli et al. [11] proposed a community matching algorithm to efficiently identify and track similar communities over time. All these studies have a common weak point that the historic community structure information, which contains valuable information related to current community structure, is not taken into account.

On the contrary, the other category takes into account of influence of historic community structure information. According to the fact that the dynamic network evolves slowly with time passing, Chakrabarti et al. [12] first proposed a framework for evolutionary clustering. In this framework, evolutionary clustering should simultaneously optimize two potentially conflicting criteria: one is that the clustering at any point in time should remain faithful to the current data as much as possible; the other is that the clustering should not shift dramatically from one timestep to the next. Based on this idea, some evolutionary clustering methods have been proposed. Chi et al. [13] proposed an evolutionary spectral clustering algorithm by incorporating temporal smoothness. Lin et al. [14] presented a novel framework for analyzing communities and their evolutions in dynamic networks, named FaceNet, based on non-negative matrix factorization. However, these above mentioned methods have some weaknesses. For instance, they assume that the number of nodes and clusters should not be changed over time, and the number of clusters should be known in advance. Moreover, Newman had pointed out [6] that spectral clustering is not good at detect natural community structure in real world networks due to the fact that cut size cannot reflect the intuitive concept of network communities.

Besides, some other optimizing methods based on evolutionary clustering had been proposed in recent years. For example, Folino et al. [15] proposed an evolutionary multiobjective optimizing approach for community discovery in dynamic networks utilizing genetic algorithm with two objectives: one objective used to maximize the snapshot cost while the other used to minimize the temporal cost in the framework of evolutionary clustering. Zhou et al. [16] put forward a multiobjective biogeography based optimization algorithm with decomposition for community detection in dynamic networks. A common weakness of evolutionary-based methods is its low efficiency in finding community partitions at different timesteps. That is because it needs to perform too many complicated operations, such as selection, crossover, mutation, iteration and so on, for each network slice, which may not be effective in real-world networks.

In our previous work [17], we had proposed an evolutionary community structure discovery algorithm (ECSD) in dynamic weighted networks based on node analysis. One weakness of the ECSD is that it cannot detect overlapping and hierarchical community structure in dynamic weighted networks, while these features play an important role in providing users to have a better understanding of dynamic community structure in evolving networks. Moreover, the ECSD adopts the change of modularity to check whether to expand a community and whether to merge two communities to form a new community, while Fortunato and Barthelemy pointed out [18] that the modularity has a resolution limit problem, and modularity-based methods cannot find communities smaller than a specialized size.

More recently, some scholars begin to focus on link communities' analysis [19,20], regarding communities as partitions of links rather than nodes. The link communities' analysis mainly concentrates on relations mining among nodes in the network. In real world, networks have communities with overlapping, where each node belongs to more than one group, whereas an edge between nodes always belongs to only one community. According to this feature of the network, Ahn et al. [19] proposed a novel link community detection method by utilizing link structure information. At first, it constructs a dendrogram by edges similarity. Then, it partitions this dendrogram based on a function named partition density, and then converts the link communities to nodes communities according to inherent relations between nodes and edges, which can find overlapping nodes in the network. One weak of this method is it needs to get the whole network structure to construct hierarchy structure of links, and assumes there are no unconnected nodes in the network. In dynamic networks, it is hard to acquire the entire network structure information at changing timesteps. Moreover, there also exist unconnected parts due to the relations between nodes evolving with time passing.

Please cite this article in press as: Q. Liu, et al., Evolutionary link community structure discovery in dynamic weighted networks, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2016.09.028