Physica A 392 (2013) 3260–3272

Contents lists available at SciVerse ScienceDirect

Physica A

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

Ant colony clustering with fitness perception and pheromone diffusion for community detection in complex networks

Junzhong Ji [a,*], Xiangjing Song [a], Chunnian Liu [a], Xiuzhen Zhang [b]

[a] College of Computer Science and Technology, Beijing University of Technology, Beijing Municipal Key Laboratory of Multimedia and Intelligent Software Technology, Beijing, 100124, China

[b] School of Computer Science and Information Technology, RMIT University, Australia

HIGHLIGHTS

- We propose an ant colony clustering approach to discover network communities.
- Each ant uses a new fitness function to percept local environment in the grid.
- Ants employ a pheromone diffusion model to realize information exchange.
- The method combines the local optimization with global intelligence.

ARTICLE INFO

Article history:
Received 10 October 2012
Received in revised form 31 March 2013
Available online 11 April 2013

Keywords:
Complex network
Community structure detection
Ant colony clustering
Fitness perception
Pheromone diffusion model

ABSTRACT

Community structure detection in complex networks has been intensively investigated in recent years. In this paper, we propose an adaptive approach based on ant colony clustering to discover communities in a complex network. The focus of the method is the clustering process of an ant colony in a virtual grid, where each ant represents a node in the complex network. During the ant colony search, the method uses a new fitness function to percept local environment and employs a pheromone diffusion model as a global information feedback mechanism to realize information exchange among ants. A significant advantage of our method is that the locations in the grid environment and the connections of the complex network structure are simultaneously taken into account in ants moving. Experimental results on computer-generated and real-world networks show the capability of our method to successfully detect community structures.

1. Introduction

Nature, society, and many technologies can be modeled by numerous networks that involve many important and complicated interactions among individuals [1], such as Biological networks, Social networks, Collaboration networks, and Web networks. These networks are often called complex networks. These complex networks are found to divide naturally into communities, which are groups of nodes such that the nodes within a group are much more connected to each other than to the rest of the network. In recent years, community structure discovery has been one of the most popular research areas along with its applicability to a wide range of disciplines [2–5]. For example, cell biologists use the community structure in protein interaction networks to make sense of signal transduction cascades and metabolism, to research the inherent relationships between cellular functions and biochemical events in this area. Computer scientists are mapping the Internet and the WWW into different communities to discover topic information from Web pages and potential relationships in

* Corresponding author. Tel.: +86 010 67396568.
E-mail address: jjz01@bjut.edu.cn (J. Ji).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.04.001

hyperlinks. Epidemiologists follow transmission networks through which viruses spread and try to find how to stop the spread of the virus by analyzing the transmission network structure, and so on. That is, communities are interesting and fundamental because they often correspond to functional units and reflect the in-homogeneity or locality of the topological relationships between the nodes of target systems. Thus, community structure mining is important for analyzing topological structures, comprehending network functions, recognizing hidden patterns, and forecasting the future behaviors of complex systems.

With the development and popularity of complex networks, various network clustering algorithms to mine community structures have been proposed [4,6–20]. In Ref. [21], Fortunato gave a comprehensive and in-depth summary for the methods and algorithms of community detection in graphs from different aspects. Although special strategies adopted are different, most of the algorithms are mainly divided into two basic categories from the view of the search mechanism: the global optimization approach and local search approach [20]. The first approach poses clustering as a global optimization problem, which employs an objective function to evaluate the network modularity quality and tries to find an optimal clustering result in the whole solution space [6–8,14–17]. In contrast, there are no global optimization objectives in the second approach, and they carry out clustering based on some local heuristic rules [4,9–13], such as edge betweenness, clique percolation, random walk, etc. Moreover, there are some algorithms that use a combination of these two basic methods [18–20], which can get better clustering results and higher time efficiency than each basic approach. For all that, however, how to further improve the detection accuracy, especially how to discover reasonable community structure without prior knowledge, is still a challenging problem. In this paper, we propose an Ant Colony Clustering algorithm based on Fitness perception and Pheromone diffusion for community detection in complex networks (called as ACC-FP), which is also a combination method. ACC-FP first uses a new fitness function to percept local environment and move to more comfortable locations in the grid. Then, in terms of the ants clustering quality at each iteration, it employs a pheromone diffusion model, which can faithfully simulate the volatilization character of pheromone, to realize information exchange among ants and guide ant colony to perform global searches. Experimental results and relevant comparative analyses show that the combination of local perception and global information feedback is effective to achieve high-quality clustering results.

The rest of this paper is organized as follows. Section 2 introduces related research on community detection. Section 3 presents the motivation and the details of the proposed algorithm. Section 4 presents and analyzes the experimental results. Section 5 concludes this paper.

2. Related work

In the past decade, the problem of community detection has attracted much attention in various scientific fields including physics, sociology, and computer science. Many methods employing different strategies have been proposed and applied successfully to some specific complex networks. For instance, Girvan and Newman proposed GN algorithm, which is a method for detecting a community structure by using information about edge betweenness [4]. Subsequently, Newman proposed a fast algorithm for detecting a community structure based on the Q metric [6]. Based on a q-state Potts model, Reichardt and Bornholdt presented a fast community detection algorithm, which can detect overlapping communities without prior knowledge about the number of communities [7]. By means of the eigenvectors of a modularity matrix for the network, Newman proposed a spectral algorithm for community detection, which can acquire higher quality than some competing methods in shorter running times [8]. Radicchi et al. gave two quantitative definitions of community and presented a local algorithm to detect communities, which can efficiently deal with large-scale networks [9]. Gergely et al. used clique percolation to identify densely connected subgraphs in complex networks [10]. Frey et al. employed affinity propagation to pass messages and iteratively partitioned data points into the exemplar clusters [11]. Yang et al. developed a new algorithm called FEC for identifying communities from signed social networks whose idea rested on an agent-based random walk model [12]. In 2012, Faqeeh et al. put forward a new algorithm (called FA in this paper), which employed the eigenvectors of the clumpiness matrix to construct a "projection space" and used some kind of angular distance in this space to define a border line, and then applied this borderline and hierarchical clustering methods to identify the community structure of a complex network [13]. Gong et al. proposed the improved memetic algorithm called iMeme-Net for solving community detection problems, which combined Label Propagation (PGLP) tactic, an Elitism Strategy (ES), and an Improved Simulated Annealing Combined Local Search (ISACLS) strategy to carry out population generation [14].

It is worth noting that there has been an increasing development trend which employed swarm intelligence to detect the community structure of complex networks in recent years. Pizzuti proposed a new algorithm to discover communities in networks by employing Genetic Algorithms (named GA-Net) [15]. The approach introduced the concept of community score to measure the quality of a partition in communities of a network and tried to optimize this quantity by running the genetic algorithm. Liu et al. proposed a communicating community discovery algorithm based on an ant colony clustering model, which employed movement, picking-up and dropping-down operators to perform node clustering in email networks [16]. To reduce clustering computational costs without loss of solution quality, Sadi et al. used ant colony optimization (ACO) techniques to find cliques in a network and assigned these cliques as meta-nodes, and then employed a traditional algorithm to find community memberships on the reduced graph [17]. To further optimize the approach, authors subsequently used the snowball sampling method to generate these subgraphs and ran the ACO-based clique finding technique on each one in parallel [18]. Taking the Q metric as an objective function, Zhang et al. proposed a mining community method in dynamic social networks, which used a clustering center initialization, the pheromone updating, and heuristic function guiding