Physica A 449 (2016) 336–347

Contents lists available at ScienceDirect

Physica A

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

A multi-agent genetic algorithm for community detection in complex networks

Zhangtao Li, Jing Liu *

Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University, Xi'an 710071, China

HIGHLIGHTS

- The multi-agent system integrating with genetic algorithm is first used to detect communities in complex networks.
- A series of effective neighborhood-based operators are designed.
- The good performance of the new algorithm is validated by various networks and the systematic comparisons with two representative algorithms.
- The new algorithm can detect communities with high speed, accuracy and stability.

ARTICLE INFO

Article history:
Received 11 May 2015
Received in revised form 17 November 2015
Available online 7 January 2016

Keywords:
Community detection
Multi-agent systems
Neighborhood-based operators
Modularity
Genetic algorithm

ABSTRACT

Complex networks are popularly used to represent a lot of practical systems in the domains of biology and sociology, and the structure of community is one of the most important network attributes which has received an enormous amount of attention. Community detection is the process of discovering the community structure hidden in complex networks, and modularity Q is one of the best known quality functions measuring the quality of communities of networks. In this paper, a multi-agent genetic algorithm, named as MAGA-Net, is proposed to optimize modularity value for the community detection. An agent, coded by a division of a network, represents a candidate solution. All agents live in a lattice-like environment, with each agent fixed on a lattice point. A series of operators are designed, namely split and merging based neighborhood competition operator, hybrid neighborhood crossover, adaptive mutation and self-learning operator, to increase modularity value. In the experiments, the performance of MAGA-Net is validated on both well-known real-world benchmark networks and large-scale synthetic LFR networks with 5000 nodes. The systematic comparisons with GA-Net and Meme-Net show that MAGA-Net outperforms these two algorithms, and can detect communities with high speed, accuracy and stability.

1. Introduction

In recent years, with the development of information technique, complex networks have been used in many domains, such as web, power grids, sensor networks, biological networks and social networks [1]. In these applications, networks can be modeled as graphs where nodes represent objects and edges represent relationships between objects. The community structure, as one of the most important properties of complex networks, has received a lot of attention [2,3]. A network has

* Corresponding author. Tel.: +86 29 88202661.
E-mail addresses: neouma@mail.xidian.edu.cn, neouma@163.com (J. Liu).

http://dx.doi.org/10.1016/j.physa.2015.12.126
0378-4371/

clear community structure if the nodes inside communities are densely connected, while the nodes between communities are loosely connected.

A good community structure implies important messages about relationships between network function and topology. Many community detection algorithms have been proposed, where modularity-based algorithms are the most popular ones [1]. The modularity Q is a quality function measuring the quality of partitions of networks proposed by Newman and Girvan [3]. An effective way to solve the community detection problems is to find the optimal partitions with high modularity [4].

Community detection is a NP-hard problem that traditional optimization methods cannot solve it effectively [5]. Genetic algorithms (GAs) play a very important role in solving this kind of complex problems. But the major problem of GAs is that they may be trapped in local optima and they have difficulties in addressing large-scale problems effectively. In our previous work, a multi-agent genetic algorithm (MAGA) was proposed in Ref. [6] to solve large scale global numerical optimization problems, which achieved a good result with the dimensions increasing from 20 to 10 000. MAGA was a combination of multi-agent systems and GAs and the experimental results indicated that this kind of combination was effective in solving large-scale problems. Recently, multi-agent systems have been integrated with evolutionary algorithms to solve constraint satisfaction problems and combinatorial optimization problems with satisfactory results [7,8]. Moreover, an improved agent based model (iABM) is applied to dynamic airspace sectorization [9]. A multi-objective evolutionary algorithm is used to optimize this model with a good performance. All the above evidences show that MAGA is suitable for handling large-scale complex problems.

In order to cluster large-scale networks with high accuracy, in this paper, a multi-agent genetic algorithm, named as MAGA-Net, is proposed to optimize modularity value for the community detection. Based on the locus-based adjacency representation, a split and merging based neighborhood competition operator is designed. To make full use of the two-point and uniform crossover operators, we design a hybrid neighborhood crossover operator. Moreover, we use an adaptive mutation operator from Ref. [10] to effectively explore the search space when the number of generations grows without improvement. At last, we conduct the self-learning operator on the best sl number of agents in each generation to further increase their energy.

To validate the performance of MAGA-Net, in the experiments, both well-known real-world benchmark networks and large-scale synthetic LFR networks are used. The results show that MAGA-Net has the ability to find correct partitions of large-scale networks with 5000 nodes. The systematic comparisons with GA-Net and Meme-Net show that MAGA-Net outperforms these two algorithms, and can detect communities with high speed, accuracy, and stability.

The rest of the paper is organized as follows. We review related work on community detection in Section 2. The details of MAGA-Net are described in Section 3. The experiments on well-known real-world benchmark networks and large-scale synthetic LFR networks are performed in Section 4. Finally, conclusions are given in Section 5.

2. Related work

The key objective of community detection problems is to find the hidden communities of networks. Various methods have been proposed to give reasonable partitions of networks [11–14]. The method proposed in this paper is a kind of evolutionary algorithms (EAs). So, in this section, we will give a brief introduction of existing EAs for community detection.

EAs are effective methods to deal with problems in complex networks. In our previous work, a memetic algorithm for enhancing the robustness of scale-free networks against malicious attacks was proposed by Zhou et al. in Ref. [15]. For detecting communities, the evolution usually starts from a random set of individuals, and then every individual in the population is evaluated, next, a series of evolutionary operators are conducted to form a new population which will be used in the next generation. Repeat the above steps until termination has reached. As a result, only individuals with large fitness survive.

Bui et al. in Ref. [16] proposed a genetic algorithm for graph partition with a schema preprocessing phase to improve GAs' space searching capability. Talbi et al. in Ref. [17] proposed a parallel genetic algorithm for the graph partition problem which showed a superlinear speed-up. Tasgin et al. in Ref. [18] used a genetic algorithm to detect communities based on modularity. Pizzuti in Ref. [19] proposed a genetic algorithm for community detection named as GA-Net using the locus-based adjacency representation and uniform crossover. It was efficient in reducing the invalid search when only the actual correlations of all nodes were considered in each operator. A new collaborative evolutionary algorithm was proposed by Gog et al. in Ref. [20] which was based on information sharing mechanism between individuals in a population. Gong et al. in Ref. [21] proposed a memetic algorithm to optimize the modularity density for community detection. A local search procedure named as high-climbing strategy was added to genetic algorithm which performed better than traditional GAs. Gong et al. in Ref. [22] also proposed a multi-objective evolutionary algorithm based on decomposition which optimized two contradictory objectives, negative ratio association and ratio cut. Liu et al. in Ref. [23] designed a representation method which could represent separated and overlapping communities at the same time and proposed a multi-objective evolutionary algorithm to solve community detection problems under the framework of NSGA-II. Li et al. in Ref. [24] made a comparative analysis of evolutionary and memetic algorithms for community detection from signed social networks. Zeng et al. in Ref. [25] and Liu et al. in Ref. [26] both proposed a multi-objective evolutionary algorithm for community detection from signed social networks. While the algorithm in Ref. [26] was based on similarity and a direct and indirect combined representation was designed to detect both separated and overlapping communities.