Physics Letters A ••• (••••) •••–•••

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Quantum inspired evolutionary algorithm for community detection in complex networks

Meng Yuanyuan [a,b], Liu Xiyu [a,*]

[a] School of Management Science and Engineering, Shandong Normal University, China
[b] School of Computer Science and Technology, Shandong University of Finance and Economics, China

ARTICLE INFO

Article history:
Received 2 January 2018
Received in revised form 24 May 2018
Accepted 25 May 2018
Available online xxxx
Communicated by M.G.A. Paris

Keywords:
Quantum
Complex networks
Community detection

ABSTRACT

Community structure is indispensable to discover the potential property of complex network systems. In this paper we propose two algorithms (QIEA-net and iQIEA-net) to discover communities in social networks by optimizing modularity. Unlike many existing methods, the proposed algorithms adopt quantum inspired evolutionary algorithm (QIEA) to optimize a population of solutions and do not need to give the number of community beforehand, which is determined by optimizing the value of modularity function and needs no human intervention. In order to accelerate the convergence speed, in iQIEA-net, we apply the result of classical partitioning algorithm as a guiding quantum individual, which can instruct other quantum individuals' evolution. We demonstrate the potential of two algorithms on five real social networks. The results of comparison with other community detection algorithms prove our approaches have very competitive performance.

1. Introduction

Many systems existing in real world can be characterized by complex networks, such as neural, biological, technological and social networks, food web, etc. In those networks, a vertex (or a node) represents an individual or a component, and an edge (or a link) represents natural or artificial relationships. Generally, a community can be defined as a set of nodes with more densely internal connections and relatively sparse connections than with the rest nodes of the network [1,2]. Some unexpected meanings and structural features of complex networks can be revealed by the community structure. For this reason, the structure of network is regarded as a common and significant property. In recent years, community discovery is becoming one of research hotspots in the field of biology, physics, sociology, climate and others. Many different approaches have been proposed [1,3,4]. In these approaches, modularity is an important benefit function for measuring the quality of community division in social networks, which is proposed by Newman and Girvan. Besides, many typical algorithms have been declared to community discovery in complex networks, such as LPA, BGLL algorithm etc, which can be found in the literature [5,6].

In the existing algorithms, there are a class of approaches based on evolutionary algorithm (EAs) and swarm intelligence algorithms (SIAs), which has a large volume of research, for example, see [7–10]. These metaheuristics are population-based metaheuristics, which are outstanding for excellent local and global search abilities and suitable for solving optimization problems. In general, finding an underlying community structure in a network can be considered as a problem of cluster analysis. In many clustering algorithms, clustering can be equivalent to an optimization problem, therefore the research of community discovery can be formally defined as an optimization problem. Over the past ten-odd years, many researchers aim to apply EAs and SIAs to community division. Recent surveys could be found in [11–15]. In [11] the authors proposed a memetic algorithm for community detection by optimizing modularity and used multi-level learning strategies to accelerate the optimization process. Pizzuti applied two objective functions to evaluate division result, which were community score and community fitness. Consequently, the community detection problem can be transformed into a multiobjective optimization problem [12]. In [13], authors investigated the GAs with a random walk based on distance measure to study the subgroups in social networks. In [14], authors used a multiobjective discrete particle swarm optimization (MODPSO) algorithm to detect community. These algorithms play a significant role in community detection. However, they still will be confronted with an obstacle in solution limitation. In [11], MLCD algorithm used a multi-level learning

* Corresponding author.
E-mail addresses: myy801@126.com (M. Yuanyuan), Liuxy@sdnu.edu.cn (L. Xiyu).

https://doi.org/10.1016/j.physleta.2018.05.044
0375-9601/© 2018 Published by Elsevier B.V.

strategies and the sub community was the basic unit for merging and splitting. However, if a vertex was misclassified into a sub community, it was hardly to jump out from the sub community in the subsequent stages. The proposed algorithm in [12], the individuals were initialized randomly. Although the individuals were corrected in later period, they failed to make full use of the effective information in the community network, so the initialized individuals were not of high quality. In [13], the proposed algorithms can only solve the community's division which the number of community were known. Thus the use of these algorithms had some limitations.

In 1980, Benioff [16] proposed Quantum mechanical computers. In 1985, Deutsch [17] formalized the description of quantum mechanical computers. Quantum inspired genetic algorithms was firstly introduced in [18]. In [19], a genetic quantum algorithm was proposed by Han et al. which used the concept of qubit (Q-bit) and superposition, the principles of quantum computing. In [20], Han et al. further proposed quantum inspired evolutionary algorithm (QIEA). The essence of the QIEA is to use the superposition of quantum states in quantum mechanics. QIEA uses the qubit as a probabilistic representation and applies it to the chromosome coding, so that one chromosome can express multiple superposition states. Quantum algorithm works in parallel and the encoding and decoding of quantum are nonlinear. Based on this reason, an essential advantage of QIEA over other conventional EAs and SIAs, is QIEA has quantum parallelism and can correspond to a huge number of search states by using a smaller scale population. In addition, the representation of qubit makes QIEA have the characteristics of avoiding premature convergence and keeping the balance between exploration and exploitation even with a smaller population. Ever since emergence, QIEA has been utilized to solve various optimization problems and many other domains [21–26]. Due to the success of QIEA, we attempt this evaluation algorithm to detect community structure.

Inspired by the researches above, we propose two complex network detection algorithms, named QIEA-net and iQIEA-net. Furthermore, our two algorithms have the ability to determine the number of cluster automatically, instead of setting the number in advance, which is crucial to analyze a new social network with unknown structure. The paper is organized as follows. We briefly give the related work of community detection in Section 2. In Section 3, two proposed algorithms are described in detailedly. In Section 4, the experimental results on 5 real-world networks are discussed. Finally, this paper is concluded in Section 5.

2. Related work

2.1. Network community definition

In general, a community can be grouped into sets of vertices (or nodes) if each set of vertices is densely connected internally and relatively sparse connections between groups. A network is denoted as $G = (V, E)$ comprising a set V of vertices together with a set of E of edges. Let A be the adjacency matrix of G. Its element A_{ij} is one, if there is an edge from vertex i to vertex j, otherwise, A_{ij} is zero when there is no edge between the two vertices; k_i is the degree of vertex i, with $k_i = \sum_{j=1}^{n} A_{ij}$, $i, j = 1, 2, ..., n$, where n is the size of vertices. Network modularity function proposed by Newman [27], also called Q function, which is widely used in the studies of community partition to evaluate a division of a network into communities. Formally, one of the definition of Q-function is

$$Q = \frac{1}{2m}\sum_{i,j}[A_{ij} - \frac{k_i k_j}{2m}]\delta(c_i, c_j) \tag{1}$$

where m is the size of edges and $\delta(c_i, c_j) = 1$, if vertex i and vertex j are in the same community, otherwise, $\delta(c_i, c_j) = 0$. The value of Q lies in the range between -1 and 1 which can measure the strength of links within communities, rather than the random distribution of links among all communities. Normally, the higher value of Q indicates the stronger community partition of G. The quality is generally fine when Q is between 0.3 and 0.7. In general, Q is difficult to exceed 0.7 in the real-life network. In fact, if we analyze from the perspective of improving modularity, community division will evolve into an optimal clustering problem, which goal is to get optimal or near optimal modularity value. Therefore, several metaheuristics algorithms for finding communities have been utilized to optimize Q-function to discover the community structure with the optimum Q value.

2.2. Quantum inspired evolutionary algorithm

Theoretically, QIEA belongs to the domain of evolutionary algorithms (EAs), which is based upon the concept and principle of the superposition states of quantum bits in quantum computing. Like most EAs, QIEA is also based on the representation of candidate individuals, the fitness function and a population of randomly generated individuals [20]. Unlike traditional EAs, QIEA represents a qubit as a basic information unit. Accordingly, a quantum individual can be defined by a qubit string, which can represent the probability of a linear superposition of states. Therefore, from the point of view of population diversity, the representation of qubit can be more abundant than other classical representations. Furthermore, to generate a next generation population of quantum individuals, QIEA also uses a quantum gate (Q-gate) as an evolutionary operator. In the process of evolution, the diversity property disappears gradually and the quantum individuals eventually converge to a single state. In each iteration, the probability of each qubit is shifting to 1 or 0 under the drive of quantum gate. By these inherent mechanisms, QIEA has the ability to balance exploration and exploitation.

A qubit is a linear superposition of the basis states, defined as a column vector

$$[\alpha \quad \beta]^T \tag{2}$$

where α and β are probability amplitudes, which satisfy the constrain by the equation $|\alpha|^2 + |\beta|^2 = 1$. In quantum theory, there is a fundamental property that a qubit can be expressed as a linear superposition of $|0\rangle$ and $|1\rangle$. When we measure a qubit, $|\alpha|^2$ is the probability in state '0' and $|\beta|^2$ is the probability in state '1'. Eq. (2) can be represented as $\alpha|0\rangle + \beta|1\rangle$.

Qubit can also be combined. A quantum individual can be considered to a string of n qubit, which have the following states

$$\left[\begin{array}{c|c|c|c} \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \beta_1 & \beta_2 & \cdots & \beta_n \end{array}\right] \tag{3}$$

Where $|\alpha_i|^2 + |\beta_i|^2 = 1$, $i = 1, 2, \cdots, n$. For example, a three qubit string with three pairs of amplitudes is as follows

$$\left[\begin{array}{c|c|c} \frac{1}{2} & \frac{\sqrt{5}}{3} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{2}{3} & \frac{\sqrt{2}}{2} \end{array}\right] \tag{4}$$

then the states of this system in the computational basis can be written as

$$\frac{\sqrt{10}}{12}|000\rangle + \frac{\sqrt{10}}{12}|001\rangle - \frac{\sqrt{8}}{12}|010\rangle - \frac{\sqrt{8}}{12}|011\rangle - \frac{\sqrt{30}}{12}|100\rangle - \frac{\sqrt{30}}{12}|101\rangle + \frac{\sqrt{24}}{12}|110\rangle + \frac{\sqrt{24}}{12}|111\rangle$$

Please cite this article in press as: M. Yuanyuan, L. Xiyu, Quantum inspired evolutionary algorithm for community detection in complex networks, Phys. Lett. A (2018), https://doi.org/10.1016/j.physleta.2018.05.044