Physica A 486 (2017) 251–261

Contents lists available at ScienceDirect

Physica A

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

Key node selection in minimum-cost control of complex networks

Jie Ding [a], Changyun Wen [a], Guoqi Li [b,*]

[a] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 639798, Singapore
[b] Center for Brain Inspired Computing Research, Department of Precision Instrument, Tsinghua University, Beijing, 100084, China

HIGHLIGHTS

- The revisited projected gradient method extension (R-PGME) is proposed.
- The importance index of a node is strongly related to its occurrence rate.
- Key nodes distribute as evenly as possible in elementary stems and circles.
- Importance indices of nodes in a stem/dilation show a quasi-periodic distribution.

ARTICLE INFO

Article history:
Received 22 January 2017
Received in revised form 6 April 2017
Available online 31 May 2017

Keywords:
Key node
Minimum-cost control problem
Monte-Carlo realizations
Complex networks

ABSTRACT

Finding the key node set that is connected with a given number of external control sources for driving complex networks from initial state to any predefined state with minimum cost, known as minimum-cost control problem, is critically important but remains largely open. By defining an importance index for each node, we propose revisited projected gradient method extension (R-PGME) in Monte-Carlo scenario to determine key node set. It is found that the importance index of a node is strongly correlated to occurrence rate of that node to be selected as a key node in Monte-Carlo realizations for three elementary topologies, Erdős–Rényi and scale-free networks. We also discover the distribution patterns of key nodes when the control cost reaches its minimum. Specifically, the importance indices of all nodes in an elementary stem show a quasi-periodic distribution with high peak values in the beginning and end of a quasi-period while they approach to a uniform distribution in an elementary cycle. We further point out that an elementary dilation can be regarded as two elementary stems whose lengths are the closest, and the importance indices in each stem present similar distribution as in an elementary stem. Our results provide a better understanding and deep insight of locating the key nodes in different topologies with minimum control cost.

1. Introduction

In the past decades, a great interest has been shown in studying and analyzing complex networks ranging from physical to biological models, such as social networks [1,2], economics [3,4], smart grid [5,6], sensor networks [7,8], gene regulatory network [9,10] and so on. In many of these applications, one of the most important problems is to fully control the whole network, i.e. driving the network from any initial state to a desired final state. An interesting approach is to select a subset

* Corresponding author.
E-mail addresses: jding002@e.ntu.edu.sg (J. Ding), ecywen@ntu.edu.sg (C. Wen), liguoqi@mail.tsinghua.edu.cn (G. Li).

http://dx.doi.org/10.1016/j.physa.2017.05.090
0378-4371/

of nodes directly connected to external control sources, to achieve controllability of the whole network based on the characteristic of local couplings between nodes [11–14].

One target of controlling the complex networks is to choose the minimum number of nodes that ensure controllability, which is called minimal input problem. A prominent work to solve such a problem is presented in [11], in which Liu et al. propose maximum matching algorithm based on the concept of structural controllability [15]. It is also shown that the minimum number of driver nodes required for complete control mainly depends on network's degree distribution. Shortly after [11], another objective of designing the input to control the network with minimum state variables, denoted as actuated nodes, has emerged [16,17]. This is called minimal controllability problem and has been studied extensively, see for examples [18–20] and references therein.

However, there is no consideration of cost when addressing above control problems, while their solutions ensure controllability. Actually when the Gramian controllability matrix is nearly singular, the control cost becomes excessively high. Therefore, another problem is raised, that is to minimize the control cost by locating a subset of nodes connected with a given number of external control sources that ensure controllability. We term it as minimum-cost control problem and call such nodes as key nodes. Actually, given a number of external control sources to drive the network, there could be multiple sets of drive nodes. This motivates researchers starts to investigate how to locate the "key node set" connected with external controllers such that the network is driven with minimum energy cost. Yan et al. solve this problem only by analyzing a single external input which is connected with one node [21]. In [22], the smallest eigenvalue of controllability Gramian is adopted as a metric to derive a lower bound on control cost. But no general algorithm is provided to identify the set of key nodes. Such a problem is also addressed in [23,24] by considering a submodular function of some input energy metrics proposed in [25]. To find the node set with minimum control cost, a set of nodes that provides controllability of whole network should be given first. Then some other nodes are added to the given node set to minimize control cost. In [26], with a given number of key nodes, an optimization model is built to minimize an energy cost function on a sphere surface determined by trace boundary conditions. An algorithm named projected gradient method extension (PGME) is proposed to determine associated optimal input matrix with key nodes pinpointed.

In [26], it is claimed that key nodes tend to divide the elementary stem and circle equally. But evidences or quantitative analysis is not provided to support the claim. This motivates us to further explore the characteristics of key node distribution in networks based on the approach in [26]. On the other hand, for the optimization model in [26], a nonzero parameter ϵ is introduced to avoid a supporting function being zero on the boundary surface. But this actually makes the model become hard to understand and more complicated. To overcome this, we improve the model by redefining the supporting function without the parameter ϵ and propose revisited projected gradient method extension (R-PGME) based on definition of importance index in Monte-Carlo scenario so that key nodes can be determined. With proposed R-PGME, a great number of simulations has been done. Each node of networks is quantitatively analyzed based on its importance index and occurrence rate. Besides, we elaborate the characteristics of key nodes distribution on elementary dilations, apart from elementary stems and circles. With regard to elementary stems, an interpretation for importance index distribution is made from the perspective of memory patterns observed in psychological and social science. Furthermore, we apply R-PGME in Erdős–Rényi [27,28] and scale-free [29,30] networks to show the effectiveness of R-PGME and high correlation between importance index and occurrence rate. More specifically, three important findings are obtained by applying R-PGME in Monte-Carlo scenario on three typical topologies of networks. Firstly, key nodes distribute as evenly as possible in elementary stems and circles. That is, the importance indices of nodes in an elementary stem show quasi-periodic distribution with highest peak values at both ends and in an elementary cycle they distribute uniformly. Secondly, for an elementary dilation, it can be divided into two elementary stems and importance index distribution of each stem shows similar pattern as elementary stems. Finally, the distribution of importance index is strongly correlated with the occurrence rate in Monte-Carlo realizations. These observations are of importance, as they can help us understand key nodes distribution in more depth and provide a more detailed technique for proposing heuristic algorithms to analyze minimum-cost control problems in complex networks.

The rest of paper is organized as follows. In Section 2, minimum-cost control model and R-PGME algorithm are introduced. This facilitates the definition of importance indices which enables us to determine key nodes of a complex network. In Section 3, we present some interesting results based on Monte-Carlo simulations for three elementary topologies, Erdős–Rényi and scale-free networks in complex networks. Section 4 concludes this paper.

2. Determination of key nodes based on importance index

In this section, we propose a method to determine key nodes based on their important indices. The objective is to find a set of key nodes for minimum cost control. Firstly, a model of minimum-cost control problem is formulated. Then we propose an algorithm called R-PGM that provably converges to a dense solution B^d. To obtain a sparse input matrix, we introduce importance indices based on B^d in Monte-Carlo scenarios. Finally an extension method named R-PGME is proposed such that key node set is pinpointed.

2.1. Minimum-cost control problem

In real-life complex network, with a given number of external control sources that ensures controllability, selecting a subset of the whole node set to control the network at minimum cost is called minimum-cost control problem. Here, we