Physica A 478 (2017) 49–62

Contents lists available at ScienceDirect

Physica A

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

Network-based landscape of research strengths of universities in Mainland China

Zihua Liu [a], Qin Xiao [b], Qian Zhan [c], Changgui Gu [a], Huijie Yang [a,*]

[a] Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
[b] College of Sciences, Shanghai Institute of Technology, Shanghai 201418, China
[c] Ordnance Engineering College, Shijiazhuang, Hebei Province 050003, China

HIGHLIGHTS

- We proposed a network based method to extract landscape of a complex system.
- We illustrated the landscape of research strength in Mainland.
- An old infrastructure domains the educational system in Mainland China.

ARTICLE INFO

Article history:
Received 1 November 2016
Received in revised form 10 February 2017
Available online 28 February 2017

Keywords:
Complex network
Landscape of complex system
Research strength in Mainland China

ABSTRACT

A landscape of a complex system presents a quantitative measure of its global state. The profile of research strength in Mainland China is investigated in detail, by which we illustrate a complex network based framework to extract a landscape from detailed records. First, a measure analogous to the Jaccard similarity is proposed to calculate from the presided funds similarities between the top-ranked universities. The neighbor threshold method is employed to reconstruct the similarity network of the universities. Second, the network is divided into communities. In each community the node with the largest degree and the smallest average shortest path length is taken as the representative of the community, called central node. The node bridging each pair of communities is defined to be a boundary. The central nodes and boundaries cooperatively give us a picture of the research strength landscape. Third, the evolutionary behavior is monitored by the fission and fusion probability matrices, elements of which are the percentage of a community at present time that joins into every community at the next time, and the percentage of a community at next time that comes from every present community, respectively. The landscapes in three successive 4-year durations are identified. It was found that some types of universities, such as the medicine&pharmacy and the finance&economy, conserve in single communities in the more than ten years, respectively. The agriculture&forest universities tend to cluster into one community. Meanwhile the engineering type distributes in different communities and tends to mix with the comprehension type. This framework can be used straightforwardly to analyze temporal networks. It provides also a new network-based method for multivariate time series analysis.

* Corresponding author.
E-mail address: hjyang@ustc.edu.cn (H. Yang).

http://dx.doi.org/10.1016/j.physa.2017.02.058
0378-4371/

1. Introduction

A landscape of research strength is beneficial for us in multi-ways [1,2]. It can provide an evaluation of the academic level for the government to monitor the scientific development, a guidance for fund providers and high-educational administrators to adjust supporting policies, and a direction for researchers to find collaborators.

In Mainland China, most of the researchers working in basic disciplinary fields distribute in the top universities. Research activities of the universities can be regarded as representative of the national academic level. In the present paper by using the database of National Science Foundation of China (NSFC) [3] we reconstruct a similarity network of top universities [4]. The structural behaviors of the network and their changes in successive durations are used to illustrate the landscape of the research strength and its evolution.

Technically, we select the top 116 universities to represent the institutions working on basic researches. All the NSFC-supported projects presided by these universities in a duration of twelve years are collected, which are catalogued according to the applying system into 101 sub-disciplinary fields [3]. The overlapping of distributions of projects is used to measure similarities between the universities. We construct subsequently the similarity network by linking every university with the two universities most similar with it.

It was found that the universities cluster into several communities. In each community the university with the largest number of neighbors and the smallest average shortest path length is taken as the representative of the community, called central node. The universities connecting different communities are defined to be boundaries. The central nodes and boundaries are identified, which cooperatively give us a profile of the research strength. The profiles in successive durations show us a picture of global evolution. Some types of universities, such as the medicine&pharmacy and the finance&economy, stay in single communities in the more than ten years, respectively. The agriculture&forest universities tend to cluster into one community. The engineering type distributes in different communities and tends to mix with the comprehension type. Though great progress has been achieved, the main infrastructure of the educational system remains unchanged. The proposed framework can be used to display a picture of a temporal network. It provides also a new method to analyze multivariate time series from the perspective of complex network theory.

2. Materials and methods

2.1. Data

In 1986, China's State Council approved the establishment of the National Science Foundation of China (NSFC) [3]. The NSFC Committee is in charge of its running. The aim is to support in the nationwide free explorations of basic scientific problems, and play consequently a guiding role in the scientific development. Researchers in universities and research institutions can freely apply the support by providing their proposals. The support will be given to the applications with competitive strengths and innovative ideas. Researchers working in basic scientific fields in Mainland China receive fund supports mainly from NSFC, though it is not the sole way.

A typical record of a NSFC project contains the name of institution, sub-catalogue identification number, title, amount, duration, and researcher's name, as shown in Table 1 several examples. Herein, we are interested in the sub-catalogue identification number. All the research fields are divided into 101 sub-catalogues (see Table A.1 in Appendix).

Each year several organizations will independently issue ranks for all the universities in Mainland China. Though the details of the rank lists are different, the listed top universities are almost the same. In the present work we select the top 124 universities issued by Prof. Shulian Wu's group [4]. Because the records of eight universities are bad in quality, these universities are deleted in our consideration, which results into a set of 116 universities (see Table A.2 in Appendix). Here the ranks are used as ID numbers of the universities. In the duration from 2004 to 2015, the selected universities applied successfully a total of 134,064 projects (51.07 billion RMB) from the NSFC, while the total number of projects supported by the NSFC is 252,168 (108.58 billion RMB). The 116 universities have covered a ratio of $\frac{134064}{252168} \sim 53.16\%$ of the supported researches. Hence, the selected universities can be used as representative of the communities working in basic disciplinary fields. In calculations, some records are cleared out due to low qualities and the final number of projects is 131,298, covering $\frac{131298}{134064} \sim 97.94\%$ of the total records.

About sixty years ago, the present government conducted a plan of reuniting universities, namely, the existing universities were broken into parts according to specialties and the parts belonging to identical or similar specialties were subsequently re-integrated into new universities [5]. From then on the universities are catalogued into different types [5], including comprehension (COM), science&technology (ENG), agriculture&forest (AGR), medicine&pharmacy (MED), normal (NOR), and finance&economy (FIN), etc. About thirty years ago, the opening Mainland China realized the limitations of the catalogues and all the universities began to develop independently. But due to the unbalance of social developments in different regions and the persistence of the government's policies, not every university can develop as its expect. The government is still issuing for each university a catalogue type by using the traditional evaluation system, as shown in Table A.2 in Appendix.