Available online at www.sciencedirect.com
ScienceDirect
Physica A 382 (2007) 693–703
PHYSICA A
www.elsevier.com/locate/physa

Empirical analysis of a scale-free railway network in China☆

W. Li[a,b,*], X. Cai[a]

[a]Institute of Particle Physics, Hua-Zhong Normal University, Wuhan 430079, PR China
[b]Max Planck Institute for the Physics of Complex Systems, Noethnitzer Str. 38, 01187 Dresden, Germany

Received 8 December 2006; received in revised form 3 April 2007
Available online 20 April 2007

Abstract

We present a detailed, empirical analysis of the statistical properties of the China Railway Network (CRN) consisting of 3915 nodes (train stations) and 22 259 edges (railways). Based on this, CRN displays two explicit features already observed in numerous real-world and artificial networks. One feature, the small-world property, has the fingerprint of a small characteristic shortest-path length, 3.5, accompanied by a high degree of clustering, 0.835. Another feature is characterized by the scale-free distributions of both degrees and weighted degrees, namely strengths. Correlations between strength and degree, degree and degree, and clustering coefficient and degree have been studied and the forms of such behaviors have been identified. In addition, we investigate distributions of clustering coefficients, topological distances, and spatial distances.

Keywords: Scaling law; Transportation network; Complex system

1. Introduction

In recent years there has been rapidly growing interest in investigating the statistical and dynamical properties of network systems that consist of a number of nodes and the edges representing interactions between them. Real-world network examples such as the Internet [1], food webs [2], coauthorship networks [3], chemical reaction networks [4], and the graph of human languages [5] are ubiquitous in nature and may span many diverse research fields. Real-world and artificial networks can be classified into different groups in various ways. For example, two typical types of networks are featured by small-world [6] and scale-free [7] properties, respectively. The small-world property refers to a small, average shortest-path length and a high degree of clustering simultaneously. The scale-free property addresses the issue that the degree distribution of the network is a scale-free power law.

Different kinds of transportation networks (air, railway, highway, and local public transportation networks) play crucial roles in people's daily lives and in world economies. Massive numbers of people and

☆Supported in part by the National Natural Science Foundation of China, and the Visitors' Program of the Max Planck Institute for the Physics of Complex Systems.
*Corresponding author. Institute of Particle Physics, Hua-Zhong Normal University, Wuhan 430079, PR China.
E-mail addresses: liw@mpipks-dresden.mpg.de, liw@pks.mpg.de (W. Li).

0378-4371/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2007.04.031

quantities of goods move through these networks. It is interesting to note that transportation networks exhibit some characteristics that are not necessarily shared by other types of networks, such as small sizes (compared to the large networks such as the Internet), relatively stable spatial–temporal structures, and weighted links. There have been some empirical studies of transportation networks, for instance, the China air network [8], U.S. air network [9], Indian Railway network [10], Polish public transportation network [11], worldwide air network [12], and the Boston subway [13]. These studies have revealed an important commonality: most transportation networks are endowed with small-world properties. The degree distributions of these networks, however, may differ greatly from each other. For instance, a two-regime power law has been found for degree distributions in both the China and U.S. air networks. For the Indian Railway and Polish public transportation networks, the degree distributions are exponential laws. These observations suggest that although these different transportation networks have similar characteristics, each has its own idiosyncracies, which may be related to the economic and geographical factors of certain regions or countries. It is also worthwhile to note that the sizes (numbers of nodes) of these transportation networks are quite small, on the order of a few hundred nodes at most. For example, the China and U.S. air networks have 128 and 215 nodes, respectively, and the Indian Railway network has only 587. The difficulty in dealing with such small systems is that small-scale data provide poor statistics [14].

Strictly speaking, railway networks show quite different architectures than air networks. In an air network, each airport is relatively independent and if two of them are connected, it is rare for there to be an airport in between (except on non-direct flights). However, in a railway network, it is common for there to be numerous smaller stations between two major ones. Hence, the railway network looks more like a tree with several trunks and many branches. The question is, will this difference in the organization of two types of networks lead to any difference in their statistical features?

This paper presents an empirical analysis of the statistical properties of the China Railway Network (CRN), which is composed of 3915 train stations and 22 259 railways (lengths of track between stations). Apparently, CRN holds a larger data set than nearly all the transportation networks previously studied. Therefore, it is reasonable to expect that CRN may provide good statistics in the measurement. Section 2 gives the results of degree distribution, strength (weighted degree) distribution, and the nonlinear dependence of strength on degree. Section 3 deals with weight distribution and degree correlations. Section 4 is mainly devoted to the clustering coefficient, its mean value, distribution, and its relation with degree. In Section 5 we calculate the diameter of CRN, the distribution of shortest-path lengths, and the distribution of spatial distances. The last section is a brief conclusion.

2. Degree- and strength distribution and strength–degree correlation

CRN is a big network which connects most cities and towns in China and handles a large portion of traffic flow within the country. The entire network is made up of 26 174 train stations and railways. The structure of CRN is relatively stable after several decades of construction, even though the total length of the railway network is still growing slightly each year due to the growth of the economy and population expansion.

In the railway network, a train station is equivalent to a node in terms of the network terminology. If there is a train running through two train stations i and j, then these two stations are connected. The number of different trains passing through a certain train station i, say k_i, is called the degree of that specific station. Since CRN is bi-directional, we may also distinguish the degrees by in- and out-degree, denoted by k_{in} and k_{out}, respectively. It is rather straightforward that the in-degree of a given station is the number of different trains arriving, while the out-degree of that same station is the number of trains departing. The distribution of degrees (officially termed degree distribution) is a key quantity in the network study because it reflects, to some extent, the network topology. In other words, the degree distribution contains some basic information about how the network is organized. The degree distribution follows the well-known Poisson, scale-free, and exponential laws. Poisson degree distribution may appear in the random network or small-world network, where the most probable degree occurs at the value equal to the average degree of the whole network. Alternatively speaking, the probability that two nodes are connected is nearly the same for all the node pairs within the same network. For the network with scale-free degree distribution, each node within it is not treated