Physica A 392 (2013) 4160–4171

Contents lists available at SciVerse ScienceDirect

Physica A

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

Measuring the dimension of partially embedded networks

Dániel Kondor [a,*], Péter Mátray [b], István Csabai [a], Gábor Vattay [a]

[a] Department of Physics of Complex Systems, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter Sétány 1/A, Hungary

[b] Department of Information Systems, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter Sétány 1/C, Hungary

HIGHLIGHTS

- We propose a method for measuring the spectral dimension of partially embedded networks.
- Our method takes into account the perceived distance traveled by a random walker.
- Only link lengths or 'travel times' are needed, not the full embedding information.
- We apply our method to both synthetic and various real-world networks.
- We also compare our method to previously proposed measures.

ARTICLE INFO

Article history:
Received 24 January 2013
Received in revised form 10 April 2013
Available online 29 April 2013

Keywords:
Dimension measurement
Complex network
Spectral dimension
Diffusion on networks

ABSTRACT

Scaling phenomena have been intensively studied during the past decade in the context of complex networks. As part of these works, recently novel methods have appeared to measure the dimension of abstract and spatially embedded networks. In this paper we propose a new dimension measurement method for networks, which does not require global knowledge on the embedding of the nodes, instead it exploits link-wise information (link lengths, link delays or other physical quantities). Our method can be regarded as a generalization of the spectral dimension, that grasps the network's large-scale structure through local observations made by a random walker while traversing the links. We apply the presented method to synthetic and real-world networks, including road maps, the Internet infrastructure and the Gowalla geosocial network. We analyze the theoretically and empirically designated case when the length distribution of the links has the form $P(\rho) \sim 1/\rho$. We show that while previous dimension concepts are not applicable in this case, the new dimension measure still exhibits scaling with two distinct scaling regimes. Our observations suggest that the link length distribution is not sufficient in itself to entirely control the dimensionality of complex networks, and we show that the proposed measure provides information that complements other known measures.

1. Introduction

The dimensionality of a physical system is an essential parameter reflecting its spatial scaling properties. The dimension influences the behavior near a critical point, affecting the scaling of various static and dynamic physical quantities [1,2]. Magnetic systems can be considered as a classic example, but in the last decades the concepts and methodology of critical behavior have been successfully applied to macroscopic and real-world inspired systems too [3,4]. Recently, the connection between criticality and dimensionality has also been extensively studied in the context of complex networks [5,6].

For general abstract networks however, it is not straightforward to obtain a proper definition of dimensionality. During the last several years, there appeared a number of methods that were successful in identifying scaling laws in complex

* Corresponding author. Tel.: +36 705045819.
E-mail addresses: kdani88@elte.hu, kondor.dani@gmail.com (D. Kondor).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.04.046

networks, giving suitable generalizations of existing concepts of dimensionality [7–12]. Recently, these methods were also adapted to include global spatial information present in spatially embedded networks [13].

In this paper we investigate the feasibility of a dimensionality measurement for complex networks that relies merely on *local* information. To do so, we give a generalization of the spectral dimension [7] to record what a random walker "perceives" while traversing the network. While a random walk is not a realistic model for every possible process taking place on a network, it is suitable to gain some information about the structure of the network. Both the random walk process and the concept of spectral dimension have been successfully applied to networks previously [9,14]. Now, we refine the concept of spectral dimension to include the information present in distances or delays associated with the links, gaining a more complete measure of network dimensionality. Our new approach is readily applicable to spatially embedded networks, and additionally it allows the treatment of partially embedded networks, where the embedding information is only present as link-wise properties (e.g. distances or delays). These networks can be considered a special kind of weighted networks, where the link weights are related to the time needed to traverse the link. Note that we use the weights as generalized distances, which is fundamentally different from the treatment of networks with arbitrary weighted links [8]. The motivation for considering such networks is twofold.

First, for many real-world networks the spatial embedding is only partially feasible in practice. For instance, let us suppose a traversal or transport process taking place on a complex network lacking spatial information. As a result, we can generally obtain some sort of local, link-wise information (e.g. "delivery times" along the links), but without gaining any global knowledge on the physical layout of the network. In such situations, we can assign the measured link-wise property to each link to obtain a *partial embedding* of the network. A straightforward illustration is the Internet, where delays are relatively easy to measure, but reliably determining the geographic position of the nodes is not feasible on a large scale [15]. For similar partially embedded networks, a random walk processes can effectively utilize local knowledge to characterize the large-scale structure of the system.

Another motivation is for "fully" embedded networks. Even if there exists a natural (2 or 3D) embedding space for the network in question, it may still be relevant to study the network's scaling behavior via local, link-wise properties. As an example, imagine a typical road network, where different types of roads have different speed limits. In such a network, travel times are not in a simple relationship with the metric distance of the embedding space (the length of road segments between two intersections). It is meaningful to investigate scaling in the light of the "overlay" property (travel time), instead of the metric distance.

Our method can be considered as an extension of the classical diffusion problem. On an abstract network the diffusion process is usually considered as the function of the discrete time steps taken (i.e. the number of hops). In the case of an embedded network however, there is a natural time-scale and the delays suffered while traversing the links are directly related to the structural properties of the network. If the link delays have a broad distribution, the dynamics of the resulting process will significantly differ from the standard diffusion where each step takes the same amount of time.

The rest of the paper is organized as follows. In Section 2 we give a general overview on network dimension measurements and in Section 2.1 we review those that utilize a random walk process. We introduce our method for partially embedded networks in Section 3. In Section 4, we present simulation results for synthetic and real-world networks, and compare the findings for the presented methods. Finally, we conclude the paper in Section 5.

2. Related work

In the last several years, a number of methods have appeared in the literature that were successful in identifying scale-invariant properties in small-world complex networks. Probably, the earliest such concept is the *spectral dimension* [7] which originates from random walks on the network (see Section 2.1), and was applied to both theoretical models of networks [8] and empirical datasets [9]. Additionally, the application of the box-counting dimension [16] to networks was also proposed [11], and further generalized to reveal fractal properties of complex networks [10,12]. The scaling exponents arising from these methods are usually interpreted as a special type of network dimension.

Beyond these methods, which handle networks as abstract graphs, there have been an increasing interest in including the spatial properties of networks, too (see e.g. Refs. [17–23]). Particularly, in the context of dimension measurements, the presence of spatial information enables the application of well-known approaches [24,16] to determine the fractal dimension of the point set of network nodes [18]. A shortcoming of these approaches may be, that while they take into account the geometric layout of the network, they entirely neglect its connectivity information. Recently, in Ref. [13] Daqing et al. have proposed more suitable methods to overcome this limitation. The authors combine *both* metric and topological knowledge to yield more comprehensive measures of dimensionality.

2.1. Random walks and dimensionality of graphs

A principal method to define the dimensionality of an abstract graph is performed by examining the properties of a random walk process. We define a graph G with a set of nodes V and a set of edges E between the nodes. An edge is an unordered pair (i, j) where i and j are distinct nodes from V (the edges are undirected, and we do not allow self-edges). The graph can be represented by its adjacency matrix $\mathbf{A}$, where $A_{ij} = A_{ji} = 1$ if there is an edge between i and j, and otherwise 0. Let $\mathbf{D}$ denote the degree matrix of G, the diagonal matrix with entries $D_{ii} = \sum_j A_{ij}$.