Physica A 445 (2016) 231–239

Contents lists available at ScienceDirect

Physica A

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

Structure properties of a doubly-stochastic process on a network

Rui-Jie Xu, Zhe He, Jia-Rong Xie, Bing-Hong Wang *

Department of Modern Physics, University of Science and Technology of China, Hefei 230026, People's Republic of China

HIGHLIGHTS

- We prove the convergence of a process to generate a doubly-stochastic matrix.
- The convergence depends on existence of certain patterns in a network.
- The existence probability depends on higher moments of degree distribution.
- The existence probability depends on whether vertices with degree 1 exist.
- We find the converge–diverge phase transition point of the process on BA networks.

ARTICLE INFO

Article history:
Received 30 January 2015
Received in revised form 15 September 2015
Available online 19 October 2015

Keywords:
Networks
Doubly-stochastic
Patterns

ABSTRACT

In this paper, we study how special patterns affect certain dynamic process on networks. The process we analyze is an iteration to generate a doubly-stochastic matrix consistent to the adjacent matrix of a network and the patterns can be described as h non-interconnected vertices only connect other g vertices ($h > g$). From the perspective of network structure, we prove that the necessary and sufficient condition when the iteration converges is that these patterns do not exist in the network. For BA networks, there is a phase transition. The diverge–converge transition point is that the average degree is about 8, which is theoretically proved. The existence of these patterns depends on two factors: first, higher moments of degree distribution of the network; second, the probability that vertices with degree 1 exist in the network. Simulation results also support our theory.

1. Introduction

As a method to describe the real word complex systems [1], complex network science pays much attention to the heterogeneity of systems. In the last decades, many statistical parameters have been proposed to distinguish networks with different properties, such as degree distribution, degree correlation [2,3], and clustering coefficient [4]. Present study finds that the heterogeneous structures make the network dynamics very different from ordinary cases. For example, the power law degree distribution makes the networks more robustness to random failure, but extremely vulnerable to targeted attack [5], and the threshold of epidemic spreading [6] becomes zero on certain networks. Besides, ranks and eigenvalues of the adjacent matrix [7] of networks are also being used to describes the heterogeneity of networks.

However, the local topology patterns, formed by several adjacent vertices and edges are not so popular while they may be the key to the study of some dynamics on networks. One of the most famous examples of the usage of patterns is network controllability [8]. Another example is feed-forward loop [9], which is discussed in engineering systems.

* Corresponding author.
E-mail address: bhwang@ustc.edu.cn (B.-H. Wang).

http://dx.doi.org/10.1016/j.physa.2015.10.002

In this paper, we analyze a special dynamic and try to discover the relationships between topology and statistical parameters of networks. The dynamic we focus on is the process to generate a doubly-stochastic matrix [10] which is useful in many fields [11,12]. The process is formed by row normalization and column normalization. It is a way to make a distribution 'more equal' and this is a majorization [12] method for all continuous Shur-convex functions. Early in the 20th century, economists used doubly-stochastic matrix to measure inequality of incomes or of wealth [13]. In the study of networks, it can be used to optimize load distribution in a transmission system [14]. The properties and the methods to generate a doubly-stochastic matrix [15–18] have been studied widely. However, the relationships between doubly-stochastic matrix and network structures have not been discussed. The study of how network structures affect the feasibility of the generating process is necessary for the application of a doubly-stochastic matrix and this is what we are going to discuss in this paper.

The paper is organized as follows: first, we will propose the iteration process on an undirected network. We will show what determines the convergence of the iteration process and prove the existence of a doubly-stochastic matrix of a network. Then, we will propose a method to judge the convergence and some special networks will be discussed.

2. The iteration process and the convergence condition

The iteration process is defined as follow:

In a network, two positive variables are assigned to each edge, p_{ij} and p_{ji}. In real systems, they can be regarded as the elements of the transition matrix of a routing strategy or the load of a road in a traffic system. The process is an iteration. The initial value p_{ij} can be arbitrary and we do column normalization and row normalization iteratively to update each p_{ij}. Here, normalization means that to sum all the elements in a column (or a row), and then divide each element by the corresponding sum. If the iterative process is convergent, the result is a doubly-stochastic matrix of the network.

A more comprehensive definition of this process can be described as some operations on the adjacent matrix of the network.

a. Randomly generate a matrix P which is consistent with network A.
b. Normalize each column of matrix P and then normalize each row of it.
c. Repeat step b until matrix P becomes a constant in operation b.

It is easy to see that for a network with isolate vertices, there exist some rows and columns in which all elements are zero. These rows and columns cannot be operated by this process. However, we may just ignore them because these columns and rows do not involve in the normalization and they do not affect the results of the normalization of other columns and rows. So for simplicity, the networks we discuss below do not contain isolate vertices.

The convergence of this iteration process differs with different networks. We can see that when the sum of elements in each row and the sum of elements in each column become 1, the value of all elements will not change in the iteration and the process converges. According to our process, after doing operation b, the sum of elements in a row is 1 while the sum of elements in a column varies. So what we want to do is to find out when the sum of elements in each column becomes 1.

We will show that some patterns are the key to the convergence of this iteration process. The patterns are formed by two groups of vertices. The first group contains h vertices and the second contains g vertices and there is a requirement that $h > g$. The vertices in the first group only connect vertices in the second group and there are no constraints for the vertices in the second group. The patterns can be described like this: h non-interconnected vertices only connect other g vertices ($h > g$). We abbreviate these patterns as ($h > g$) patterns.

2.1. Grouping process

Before the derivation, we first classify the rows of a matrix into several groups. If we classify the rows into different groups by the following regulation: for any row x and row y, whenever there is a j such that $p_{xj} > 0$ and $p_{yj} > 0$, we put row x and y in a group. Traverse all pairs of rows, then, we have classified these rows into several groups $A, B, C \dots$. Rows in a group A satisfies: for any row $x \in A,\ y \in A$, there exists a column j such that $p_{xj} > 0$ and $p_{yj} > 0$.

Considering the columns of this matrix, for any row $x \in A$, whenever there exist a j such that $p_{xj} > 0$, we let column j belong to A. We can see that a column can only belong to one group since our classification method guarantees that rows in different groups satisfy that for any j, any row $x \in A$ and any row $y \in B$ ($A \neq B$), $p_{xj} > 0$ and $p_{yj} > 0$ cannot be both satisfied. Thus, columns can also be classified into groups corresponding to the groups of rows.

In the following discussion, that row x belongs to group A is labeled as $x \in A_r$, that column j belongs to group A is labeled as $j \in A_c$.

After doing operation b n times, the (i, j) element of matrix $P_{(n)}$ is $p_{(n)ij}$. Do operation b again once and we have $P_{(n+1)}$. Then let the sum of elements in jth column of $P_{(n)}$ be $r^j_{(n)}$. And $r^j_{(n)A}$ is defined as for any group A:

$$r^j_{(n)A} = r^j_{(n)}, \quad j \in A_c. \tag{1}$$