Physica A 407 (2014) 1–6

Contents lists available at ScienceDirect

Physica A

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

Impact of degree mixing pattern on consensus formation in social networks

Xiao Fan Liu [a,*], Chi Kong Tse [b]

[a] *Southeast University, Nanjing, China*

[b] *The Hong Kong Polytechnic University, Hung Hom, Hong Kong*

HIGHLIGHTS

- A social network with more than 50,000 nodes is sampled from online social service "*Twitter*".
- A *degree mixing correlation* is proposed to measure the randomness of the mixing pattern in detail.
- We show that specific structure property slows down the consensus formation process in social networks.

ARTICLE INFO

Article history:
Received 4 November 2013
Received in revised form 3 March 2014
Available online 31 March 2014

Keywords:
Social network
Degree mixing pattern
Consensus
Twitter

ABSTRACT

The consensus formation process in a social network is affected by a number of factors. This paper studies how the degree mixing pattern of a social network affects the consensus formation process. A social network of more than 50,000 nodes was sampled from the online social services website *Twitter*. Nodes in the *Twitter* user network are grouped by their in-degrees and out-degrees. A *degree mixing correlation* is proposed to measure the randomness of the mixing pattern for each degree group. The DeGroot model is used to simulate the consensus formation processes in the network. Simulation suggests that the non-random degree mixing pattern of social networks can slow down the rate of consensus.

1. Introduction

The study of consensus formation in social networks has been an interesting topic in physics and social science. Understanding the process that consensus emerges in society is vital to organization management and decision making process. Numerical study of the consensus formation process usually considers a social network as a collection of agents interacting with each other. Existing study has focused on examining the consensus process on lattices [1], regular graphs [2] or random networks [3]. However, the structural properties of actual social networks are far more complex than regular network models and random network models. Therefore, it is very important to understand how the topological properties can affect the consensus process in an actual social network.

Online social services websites such as *Facebook*, *Twitter*, *Weibo* and *Renren* host hundreds of millions of active users everyday. These users map their offline social relations on to the website, by "friending" offline contacts and making new social connections through the websites. The social relations of hundreds of millions of people in these websites have formed a very detailed map of the structure of human society. In this paper, a sample network of more than 50,000 nodes is extracted from *Twitter* website. Topological properties, specifically, the degree mixing pattern, of the sample network is studied. Compared to regular and random network models, the sample social network shows an intrinsic non-randomness

* Corresponding author.
E-mail address: xfliu@seu.edu.cn (X.F. Liu).

http://dx.doi.org/10.1016/j.physa.2014.03.086

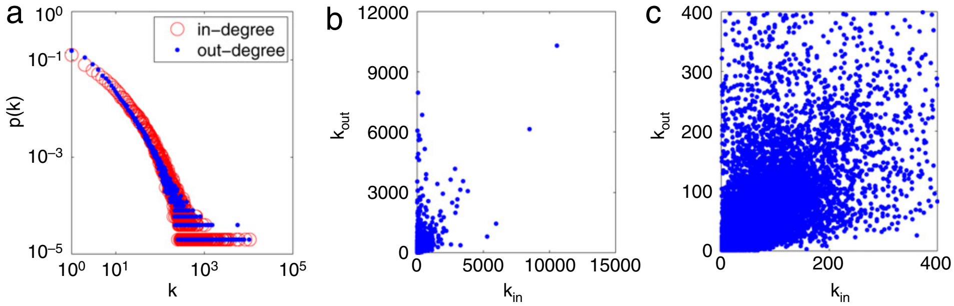

Fig. 1. (a) The in-degree and out-degree distributions of the sample *Twitter* user network. (b) For each point on the figure, the coordinate corresponds to the (in-degree k_{in}, out-degree k_{out}) of a specific node in *Twitter* network. (c) Same as (b) but only with nodes of in-degree and out-degree less than 400.

in degree mixing pattern. Using a simple consensus protocol, the consensus formation processes are simulated in both the actual social network and network models. The results indicate that the specific degree mixing pattern of social network can actually slow down the consensus process. Furthermore, an asymmetry of in-degree and out-degree of nodes can be observed in the sample social network. The impact of such in-/out-degree mismatch to the consensus formation process is also discussed in this paper.

2. Degree mixing pattern of an online social network

In order to obtain a social network sample, the earliest 60,000 users and their online relations are obtained from the *Twitter* website. On *Twitter*, users can choose to subscribe to other users and receive their updates by "following" them. These "following" relations can be modeled by edges directing from the users being followed to the subscribers, resembling the directions of information flow. The sampled network has 63,299 nodes and 1,648,267 directed edges. The average path length of *Twitter* network is 3.6 and the clustering coefficient is 0.27. These two values agree with other social networks sampled from *Twitter* [4]. To further refine the dataset, nodes of in-degree, i.e., number of incoming edges, $k_{\text{in}} = 0$ are removed from the sample network, as these nodes cannot receive information from other nodes. The remaining network has 50,414 nodes and 1,580,772 directed edges, the mean degree of nodes is 31. The in-degree and out-degree distributions of the nodes are shown in Fig. 1(a). Both distributions are tested for scale-free properties [5]. The in-degree distribution has a power-law exponent of 2.87, a lower bound of power-law behavior of 184 and a goodness-of-fit p-value of 0.675. The out-degree distribution has a power-law exponent of 2.38, a lower bound of power-law behavior of 162 and a p-value of 0.587. The distribution of all-link degree of the network also exhibits scale free properties, with a power-law exponent of 2.63, a lower bound of power-law behavior of 488 and a p-value of 0.611. Furthermore, we define the reciprocal of directed edge e directed from node i to node j as the edge e' directed from node j to i, only 861,240 (52.5%) of the edges in the *Twitter* user network have their reciprocals. Fig. 1(b) and (c) show the mismatch of in-degree and out-degree of the nodes. In the figure, each point represents a node in the network. The asymmetry of nodes' in-degree and out-degree can also be found in other social networks [6].

In a social network, people can be grouped according to certain properties, for example, number of friends, occupations, etc. Assortativity measures whether nodes in the same group have preference linking to nodes in the same group. In this paper, nodes are grouped by their in-degree and out-degree. The way in which these groups of nodes connect to other groups, i.e., the degree mixing pattern, is of particular interest. A common measure of the degree mixing pattern in a network is the assortativity coefficient, which calculates Pearson's correlation coefficient of degree between pairs of linked nodes [7]. For a directed network such as the *Twitter* user network, four types of assortativity coefficients have to be measured, i.e., $r(\text{out}, \text{in})$, $r(\text{in}, \text{out})$, $r(\text{out}, \text{out})$ and $r(\text{in}, \text{in})$. Each type of assortativity coefficient stands for the tendency of nodes with degree of a particular direction to connect to nodes with degree of another particular direction [8]. In the *Twitter* user network, $r(\text{out}, \text{in}) = -0.10$, $r(\text{in}, \text{out}) = -0.04$, $r(\text{out}, \text{out}) = -0.09$ and $r(\text{in}, \text{in}) = -0.05$. The four types of assortativity coefficients in the *Twitter* user network are all close to 0, meaning that the network does not show a particular degree mixing pattern in general. However, the assortativity coefficient is a loose-grain measure of degree mixing pattern of the network and ignores the details of how each group of nodes connects to others. Therefore, in this paper, a detailed measure of the degree mixing pattern is proposed.

In an undirected network, let $P_e(k)$ be the probability that any edge in the network is connected to a node of degree k, i.e.,

$$P_e(k) = \frac{M_k}{M}, \tag{1}$$

where M_k is the total number of edges connected to nodes of degree k and M is the total number of edges in the network. Also, let $P_e(k'|k)$ be the probability that an edge connected to a node of degree k is also connected to a node of degree k', i.e.,

$$P_e(k'|k) = \frac{M_{k'k}}{M_k}. \tag{2}$$