Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

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

Bridging the gap between different social networks

Q1 Jian Qiao *, Hong-Qiao Huang, Guo-Ying Li, Ying Fan

School of Management, Northwestern Polytechnical University, 127 West Youyi Road, Xi'an 710072, China

HIGHLIGHTS

- The mechanisms of different social networks were integrated into the model.
- Individuals' social identities were employed as one of the model's components.
- The model can reproduce social networks with different growth speed.
- The model can be used to simulate a wide range of social networks.
- Our study is helpful to realize the difference and similarity of social networks.

ARTICLE INFO

Article history:
Received 4 January 2012
Received in revised form 21 May 2014
Available online xxxx

Keywords:
Social network
Synthetic model
Random attachment
Preferential attachment
Anti-preferential attachment
Transitive attachment

ABSTRACT

The available social network models that exist today were designed primarily on the basis of the analysis of statistical properties and structural features, as well as the physical or social distances between individuals of social systems, which sometimes is not sufficient because the structure of some social networks is closely tied to individuals' social identities. In addition, the difference in growth speed between different social networks is also neglected in these models. We propose a synthetic model that involves social identity and adjustable growth speed factors to compensate for these limitations. The model features four types of node connection mechanisms: random attachment, transitive attachment, preferential attachment and anti-preferential attachment. Experimental results indicate that the model can not only produce rich topological structures but can also match real social networks well in both their macro properties and their micro foundations. Thus, the model is helpful in understanding both the evolution of social networks and the differences and similarities among different social networks.

1. Introduction

A social network is a set of social actors with certain interaction patterns among them [1]. Friendship between individuals, cooperation between companies, trade between countries and intermarriage between families are all the examples of such patterns. Social network structures are highly rich. For instance, the individual's degree of corporate partnership networks [2], scientific coauthorship networks [3], company director networks [4], film actor networks [5], sexual contact networks [6] and online social networks [7] mostly follows an exponential or power-law distribution, but that of Utah Mormon networks [8] and student friendship networks [9] approximately follows a Gaussian distribution [10]. In addition to degree distribution, social networks are also distinct from each other with respect to statistical properties, such as

* Corresponding author. Tel.: +86 13227712259.
E-mail addresses: qiaojian@nwpu.edu.cn, jian.qiao@163.com (J. Qiao), huanghongqiao1989@163.com (H.-Q. Huang), hliguoying@mail.nwpu.edu.cn (G.-Y. Li), fanying@mail.nwpu.edu.cn (Y. Fan).

http://dx.doi.org/10.1016/j.physa.2014.05.067
0378-4371/

assortativity [11] and the small-world effect [12,13], which implies the coexistence of high clustering and small average distances.

To date, many social network models have been proposed for studying the structure of social networks and social processes that occur within them, such as the diffusion of information and the spread of epidemics. Watts and Strogatz [5] proposed the famous WS model to explain acquaintance networks' small-world effect, and the distinguished BA model proposed by Barabási and Albert [14] reveals the cause of the formation of many social networks following a power-law degree distribution. Since these seminal works were published, the number of studies on modeling social networks has increased tremendously [10,11,15–27]. For example, Newman [11] proposed a model characterized by significant assortativity; a model proposed by Davidsen et al. [18] interpolates between networks with scale-free and exponential degree distribution; the degree distribution of Csányi and Szendrői's model [21] exhibits two power-scaling regimes separated by a critical degree; and other researchers [22–26] have endowed their models with more properties, such as the small-world effect, right-skewed degree distribution and distinct communities. Most models are mainly based on the analysis of the statistical properties and structural features of real social networks. For example, preferential attachment and transitivity derive from the discovery of the scale-free effect and local triangular structure in real social networks. Other models are rooted mainly in the analysis of the distance dependence of social relations, such as the ubiquity of short physical [24] or social [19] distances between acquaintances.

However, the aforementioned concepts are sometimes not sufficient for modeling some social networks, one cause of which lies in neglecting the close tie between network structure and individuals' social identities—sets of characteristics attributed to them by themselves and others by virtue of their association with, and participation in, social groups [19]. It has been confirmed that there exist a large number of distinct social groups in many social networks [1]. Densely connected individuals within a social group often belong to the same team, live in the same place, perform the same work, have the same interests, etc. Sparsely connected individuals between different social groups mostly differ from each other in these aspects. Team, place, work and interest are some examples of how social identities, which are referred to as "social tags" in our model, are established. A more ubiquitous example is friendship. Our friends at different stages of our lives often do not know each other even if we are their common friends because we seldomly meet our friends belonging to different social groups at the same time and the same place. A comparative study on the performance of eight social network models carried out by Toivonen et al. [28] demonstrated that these models match only some properties of the real social networks to which they were compared. We argue that neglecting the effect of individuals' social identities is one of the important causes of this disparity. In addition, we noticed that actual social networks often evolve at different growth speeds due to their position at different stages of their life cycles, but growing models that have been developed only grow constantly by adding a node at each time step. For the above reasons, we propose a synthetic model that involves social identity and adjustable growth speed factors to compensate for the limitations of available models.

The rest of this paper is organized as follows. The key features of social networks are analyzed in the next section. Section 3 describes our model in detail. Section 4 introduces the applied statistical indices. Experiment results and discussions are presented in Section 5. Finally, we summarize our research.

2. Analysis on social network features

Acquaintance networks are a class of social networks that have been researched for many years. Some key features of acquaintance networks are ubiquitous in social networks. Here, we summarize these networks based on the following properties. (1) Fixed size. Social networks with this property are considered to include a closed population with fixed size; (2) Limited degree. The number of neighbors an individual has is limited [15,24], just as the Dunbar's number indicates that the actual number of close acquaintances of a person is rather small [29]. However, the number may be very large in online social networks due to very low maintenance cost; (3) High clustering. The probability of two strangers becoming neighbors increases with the increasing number of their common neighbors [15,22,24,30]; (4) Low density. In networks with this property, the ratio of the number of actual connections to that of potential connections is very small; (5) Short distance. Starting from an individual, we can reach others through a small number of interconnected individuals; (6) Skewness of degree distribution. The number of neighbors of a few individuals is very large, but that of most individuals is very small; (7) Q2 Community. This property refers to a group of internally densely connected nodes that are sparsely connected to others [22,24,30]; (8) Assortativity. The degree of an individual equals or very nearly approaches that of the individual's neighbors. Positive assortativity is interpreted by Bruggeman [30] as a type of homophily, i.e., sociable people like other sociable people; (9) Transitivity. If individual A is a neighbor of individual B and B is a neighbor of individual C, then A very possibly is a neighbor of C [24,30].

The features of limited degree and low density reflect people's limited time or other resources [24]. In social networks, the diversity of degree distribution is rooted mainly in different preferences for and costs of constructing and maintaining different social relations [26]. For example, the cause of acquaintance networks following a Gaussian distribution lies in the weak preference for and high cost [10] of constructing and maintaining acquaintance relations, whereas the cause of corporate partnership networks [2], scientific coauthorship networks [3], company director networks [4], film actor networks [5], sexual contact networks [6] and online social networks [7] following exponential or power-law distributions lies in the strong preference for and low cost of constructing and maintaining these social relations. Additionally, we can also find a small-world effect in many social networks.