



Empirical analysis of online social networks in the age of Web 2.0

Feng Fu, Lianghuan Liu, Long Wang*

Center for Systems and Control, College of Engineering, Peking University, Beijing 100871, China

Department of Industrial Engineering and Management, College of Engineering, Peking University, Beijing 100871, China

Received 17 July 2007; received in revised form 30 September 2007

Available online 9 October 2007

Abstract

Today the World Wide Web is undergoing a subtle but profound shift to Web 2.0, to become more of a social web. The use of collaborative technologies such as blogs and social networking site (SNS) leads to instant online community in which people communicate rapidly and conveniently with each other. Moreover, there are growing interest and concern regarding the topological structure of these new online social networks. In this paper, we present empirical analysis of statistical properties of two important Chinese online social networks—a blogging network and an SNS open to college students. They are both emerging in the age of Web 2.0. We demonstrate that both networks possess small-world and scale-free features already observed in real-world and artificial networks. In addition, we investigate the distribution of topological distance. Furthermore, we study the correlations between degree (in/out) and degree (in/out), clustering coefficient and degree, popularity (in terms of number of page views) and in-degree (for the blogging network), respectively. We find that the blogging network shows disassortative mixing pattern, whereas the SNS network is an assortative one. Our research may help us to elucidate the self-organizing structural characteristics of these online social networks embedded in technical forms.

© 2007 Elsevier B.V. All rights reserved.

PACS: 89.75.Hc; 89.75.Fb; 89.65.-s

Keywords: Social networks; Blogging networks; Social Networking Site; Topological analysis; (dis)Assortativity

1. Introduction

What a networking life! The development of social collaborative technologies, such as blogs, Wiki, and social networking sites (SNS), results in an extraordinarily fast growing online virtual community, in which people communicate, share information, and keep in touch with each other. Indeed, the concept of Web 2.0, characterized by blogs, SNS, and Wiki, is very popular. Today the World Wide Web (WWW) is undergoing a subtle but profound shift to Web 2.0, becoming more of a social web. In comparison with WWW one decade ago, individuals (grassroots) play more crucial role in the evolution of web today, that is, the active participation of individuals leads to a more diverse online world, regardless of their different colors, beliefs, and countries. Especially, these Web 2.0 sites provide an “Eden” for free minds and ideas to trigger sparks of inspiration. Recently, numerous websites designed in the spirit

* Corresponding address: Intelligent Control Laboratory, Center for Systems and Control, College of Engineering, Peking University, Beijing 100871, China. Fax: +86 10 62754388.

E-mail addresses: fufeng@pku.edu.cn (F. Fu), longwang@pku.edu.cn (L. Wang).

of Web 2.0 emerge like mushrooms after rain. For example, the SNS Facebook is quite popular among American college students. Noteworthy, these online social networks in the age of Web 2.0 provide an opportunity to analyze how trends, ideas and information travel through social communities. Thus empirical analysis of topological structure of the mentioned online social networks will be a reasonable first step to understand the various dynamic processes on top of them, since it is believed that network's topology significantly affects the dynamics taking place on it.

Let us briefly introduce the blog and SNS. Blog, which is short for “web log”, has gained its ground by online community as a new mechanism for communication in recent years [1]. It is often a personal journal maintained on the Web, which is easily and frequently updated by the blogger. In the past few years, blogs are the fastest growing part of the WWW. Social networking site provides service for messaging, sharing information, and communicating [2]. More importantly, it is a flexible and convenient platform for individuals to form and maintain online friendships. SNS sometimes takes advantage of the advancement of understanding of complex networks (see Refs. [3–5] and a recent review [6]), e.g., small-world effect (“six degree of separation”) [7], promoting the increase and consolidation of online acquaintanceships. It also in return enhances the popularity of some social networking sites. In blogging networks, the vertices represent the bloggers, and the favorite links pointing from one blog to another denote the directed edges. In SNS, the online users are connected to one another based upon bilateral agreement. Thus this network is a bidirectional one (for simplicity, we view it as undirected one). Interestingly, in blogosphere and SNS, users can conveniently create and join communities or groups who share the same interests and activities. In general, blogging networks and SNS can be viewed as representative online social networks in the age of Web 2.0. Moreover, to our best knowledge, empirical analysis of the structure of online social networks in the age of Web 2.0 is rarely considered in previous investigations. Furthermore, blogs and SNS, as emerging social media, will no doubt grow in importance and popularity. Thus it is meaningful and interesting to scrutinize the structural characteristics of these online social networks in Web 2.0 age. Such study will help us to understand the topological structure of the new-fashioned online social networks. Furthermore, it will aid web designers shape the social networking services in a more friendly and functional way.

In this paper, we will focus on two Chinese online social networks, emerging in the age of Web 2.0, namely, Sina blogs and Xiaonei SNS [8]. The former is the largest Chinese blog space provider and has more than 2 million registered users in the mainland of China. The latter is the largest and most popular social networking service provider in China. It is only open to college students. A detailed empirical analysis of the statistical properties of these two networks is presented. It is found that both the networks have small-world and scale-free features already observed in real-world and artificial networks. Further, the correlations between degree (in/out) and degree (in/out), clustering coefficient and degree, popularity (in terms of number of page views) and in-degree (for the blogging network), etc. are examined. It is also shown that the blogging network is a disassortative one, whereas the Xiaonei network an assortative one.

The rest of this paper is organized as follows. Section 2 will first introduce the data sets of these two social networks. And then empirical analysis of topological structure of the Sina blogging network and Xiaonei network is successively described in Section 3. Some discussions to the results are also provided in the end of this section. Finally, the conclusion remarks and future works are drawn in Section 4.

2. Data sets

The blogging network is abstracted from Sina blog, a Chinese blog space provider. Such network is composed of blogs (vertices) and favorite links directing from one blog to another (directed edges). We focus our eyesight in this sub-community of global blogosphere, and thus omit the links out of this community. Sina blogging network was obtained in May 2006, which was crawled down by our designed robot along the directed links. This connected network consists of 200,292 nodes and 901,607 edges. Besides, the number of page views of each blog was also collected. It can be used as an appropriate measure of the blog popularity.

The Xiaonei network is obtained from a Chinese social networking site, which is open to college students. Each registered user of Xiaonei has a profile, including her friend list. We are interested in friendships between users within Xiaonei. The friendship is constructed by bilateral agreement. Thus Xiaonei network is bidirectional one (we viewed this network as undirected one). This undirected network is composed of 396,836 nodes and 7,097,144 edges.

In what follows, we will present in detail structural analysis of these two online social networks. Some statistical properties, including degree distribution, average shortest path length, as well as degree–degree correlation will be used to reveal the connection characteristics of online social networks in the age of Web 2.0.

Download English Version:

<https://daneshyari.com/en/article/978047>

Download Persian Version:

<https://daneshyari.com/article/978047>

[Daneshyari.com](https://daneshyari.com)