Physica A 469 (2017) 334–341

Contents lists available at ScienceDirect

Physica A

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

Information spreading on mobile communication networks: A new model that incorporates human behaviors

Fei Ren [a,b,c], Sai-Ping Li [d], Chuang Liu [e,*]

[a] School of Business, East China University of Science and Technology, Shanghai 200237, China
[b] School of Science, East China University of Science and Technology, Shanghai 200237, China
[c] Research Center for Econophysics, East China University of Science and Technology, Shanghai 200237, China
[d] Institute of Physics, Academia Sinica, Taipei 115, Taiwan
[e] Alibaba Research Center for Complexity Sciences, Hangzhou Normal University, Hangzhou 311121, China

HIGHLIGHTS

- A new information spreading model incorporates two ingredients of human behaviors.
- The fraction of informed individuals is determined by information prevalence.
- The heterogeneity of link weights would slow down the information diffusion.
- The uninformed nodes are generally located at special positions in the network.

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form 17 August 2016
Available online 21 November 2016

Keywords:
Complex networks
Information spreading
Mobile phone data

ABSTRACT

Recently, there is a growing interest in the modeling and simulation based on real social networks among researchers in multi-disciplines. Using an empirical social network constructed from the calling records of a Chinese mobile service provider, we here propose a new model to simulate the information spreading process. This model takes into account two important ingredients that exist in real human behaviors: information prevalence and preferential spreading. The fraction of informed nodes when the system reaches an asymptotically stable state is primarily determined by information prevalence, and the heterogeneity of link weights would slow down the information diffusion. Moreover, the sizes of blind clusters which consist of connected uninformed nodes show a power-law distribution, and these uninformed nodes correspond to a particular portion of nodes which are located at special positions in the network, namely at the edges of large clusters or inside the clusters connected through weak links. Since the simulations are performed on a real world network, the results should be useful in the understanding of the influences of social network structures and human behaviors on information propagation.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The study of information spreading is of essential importance for controlling and managing the propagation of news and rumors in sociology. In recent years, much effort has been devoted to the study of information propagation among multi-agents in a variety of social networks [1]. Most of these social networks were theoretically designed and were generally

* Corresponding author.
E-mail address: liuchuang@hznu.edu.cn (C. Liu).

http://dx.doi.org/10.1016/j.physa.2016.11.027
0378-4371/© 2016 Elsevier B.V. All rights reserved.

limited in size, for instance small-world network [2], scale-free network [3], random network [4] and regular network like square lattice [5]. The recent advent of communication technology allows one to study social interactions among millions of mobile phone users and their topologies [6–8]. This arouses a growing interest to study information spreading on human communication networks [6,9–13].

It would be of interest to study the information spreading on real world communication networks, especially in China, which has one of the largest mobile phone markets in the world. It has been shown that both the topological structure and link weights of communication networks can significantly affect information diffusion [14,15]. A study of communication patterns of mobile phone users revealed that the coupling between interaction strength and local structure of the network indeed slowed down the diffusion process [6]. Further work found that the slowing down of information spreading in human communication networks is caused by the simultaneous effects of structural and temporal correlations with link weights [9].

Many studies of information spreading are based on epidemic spreading models of virus or diseases, e.g. *SI*, *SIS*, and *SIR* models [16]. In these models, information propagation that spreads from an informed agent to uninformed agents is analogous to the propagation of virus or diseases that spreads from an infected agent to others. However, information diffusion mechanisms are essentially different from those in disease spreading [17]. Recent studies reveal that information diffusion depends on the specific nature of human behaviors, for instance memory effects [18], social reinforcement [16], response time [19], temporal patterns [10], and other effects in social contagions [20–22]. A model was proposed to distinguish the differences between epidemic spreading and information spreading by simultaneously taking into account memory effects, social reinforcement and non-redundancy of contacts [2].

Motivated by the wide range of applications of information spreading and the availability of large volume of real world data, we propose an information spreading model based on the structure and link weights of an empirical communication network constructed from the calling records of mobile phone users provided by one of the top three Chinese mobile service providers. In particular, we want to study two important ingredients in human communication behaviors, namely information prevalence and preferential spreading. Based on the fact that there exist interest decays in human behaviors, it is likely that informed users would transmit information to only a certain fraction of their neighbors [23]. We would assume that this interest decay behavior is homogeneous for all users, which we call information prevalence. Similar phenomena observed in realistic cases include: the collective attention on new story of a web site tends to saturate since its novelty fades with time [24]; social signatures in human communication tend to persist over time [25]; the probability a message propagates between individuals decays with the length of time latency [26]. We further assume that the spreader has a preference on its communication targets, analogous to the user's preference on the news content [27] or the tendency of connecting to the second-order friends with more uninformed neighbors [28]. The informed users spread the information to their neighbors according to the intimacy of their relationship, which is measured by their call frequencies. The information spreading processes for information prevalence and preference spreading will be thoroughly examined in our model.

2. Network construction based on empirical data

We here construct the communication network based on the mobile phone data provided by a large mobile service provider in China. It contains the calling records of individual mobile phone users during two periods: from June 28, 2010 till July 24, 2010 and from October 1, 2010 till December 31, 2010. Each record contains the information about caller number, callee number, call starting time, call length and call status. In the network, each mobile phone user corresponds to a node, and the links are added between users who have reciprocal conversations. We believe that social members who communicate with each other have relatively stable relations, and this reciprocal communication network captures the main features of social networks. Since information spreading dynamics are generally simulated among the nodes in the largest connected component, we therefore retain the links within the largest connected component of the whole reciprocal network, which has 28, 534, 892 nodes and 65, 065, 857 links. The link weights are determined by the call frequency between the two sides of each link.

To get a general idea of the topological properties of the reciprocal communication network, we first calculate the probability density function (PDF) $P(k)$ of the node degree k. We equally divide the whole period of our sample dataset into four sub-periods: 28 Jun–24 Jul, 1 Oct–31 Oct, 1 Nov–30 Nov, and 1 Dec–31 Dec. $P(k)$ of the communication networks in the four sub-periods are plotted in Fig. 1(a), and the curves practically collapse onto the same curve, indicating that $P(k)$ is stable over time. Moreover, there exists a cross-over behavior in $P(k)$, which displays power-law distributions with different exponents in two regions, similar to the bimodal distributions in Ref. [29]. We here integrate the sample dataset in each of the four sub-periods, and fit the PDF in the two regions by using a maximum likelihood estimation (MLE) method based on the Kolmogorov–Smirnov (KS) statistic [30]. The power-law fits with exponents 2.24 and 5.67 can well approximate $P(k)$ in the regions $k < 100$ and $k > 100$ respectively.

The sequence of call frequencies $f = \{f_{ij}\}$ is used to measure the link weights in our model, where f_{ij} is the total number of communications between users i and j. We next calculate the PDFs $P(f)$ of the call frequencies f of the communication networks constructed from data in the four sub-periods, as shown in Fig. 1(b). The fact that the PDFs for different sub-periods collapse onto the same curve suggests that $P(f)$ is stable over time. Moreover, we find that a power-law with exponential cut-off function $cf^{-\alpha}e^{-\lambda f}$ can well fit the tails of $P(f)$, as shown in the figure. The parameters α and λ are estimated to be 1.51 and 0.002 by using the MLE method in Ref. [31].

Download English Version:

https://daneshyari.com/en/article/5103386

Download Persian Version:

https://daneshyari.com/article/5103386

Daneshyari.com