

基于综合决策的在线社交网络消息传播模型

陆 峰

(中国科学院软件研究所 北京 100190)

(中国电子信息产业发展研究院 北京 100846)

摘 要 经典的社交网络消息传播模型没有充分考虑网络拓扑结构、节点差异、消息时延、消息规模等诸多关键因素对消息传播的影响,且将节点差异简单固化为谣言传播者、潜伏者、免疫者等类型,存在较大的不科学性。针对上述情况,提出一种基于综合决策的在线社交网络消息传播模型。从节点社会关注度、消息来源可信度、消息时效性、接收消息规模等四个方面,分别构建影响节点消息转发概率的模型,并从四个方面综合决策得出节点消息转发概率。构建度数幂分布的在线社交网络,通过仿真实验,模拟在线概率、节点度差异、消息时延、接收消息规模等因素对消息传播扩散的影响。仿真结果显示,上述参数对消息传播扩散具有显著影响。该模型对加强在线社会网络治理具有重要参考价值。

关键词 社交网络 传播模型 复杂网络 综合决策 互联网治理

中图分类号 TP393.08

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2018.12.054

ONLINE SOCIAL NETWORK INFORMATION PROPAGATION MODEL BASED ON COMPREHENSIVE DECISION

Lu Feng

(Institute of Software Chinese Academy of Sciences, Beijing 100190, China)

(China Center for Information Industry Development, Beijing 100846, China)

Abstract The classic social network information propagation model does not take full account of the impact of many key factors on the spread of the information such as network topology, node differences, message delay, and message size. And the node differences are simply taken as rumors spreaders, lurkers, immunes. Such models are not scientific. In view of the above situation, we presented online social network information propagation model based on comprehensive decision. We constructed the model which influenced node information forwarding probability from four aspects: the social attention of nodes, the credibility of information sources, the timeliness of information, and the size of received information. The probability of node information forwarding was obtained from the comprehensive decision of four aspects. Online social network with degree-power distribution was constructed. We simulated the influence of online probability, node degree difference, information delay, received information size and other factors on the diffusion of information spread through simulation experiments. Simulation results show that the above parameters have a significant impact on message propagation and diffusion. The model has important reference value on how to strengthen online social network governance.

Keywords Social network Propagation model Complex network Comprehensive decision Internet governance

0 引 言

近年来,以微博、微信等代表的在线社交网络快速

发展,与经济社会活动日益融合,为人民生活提供了极大的便利,为企业生产注入了新的发展活力,对经济社会产生了巨大促进作用。技术是把双刃剑,在线社交网络普及应用造福大众生活、促进生产力发展、推动社

会进步的同时,也成为了谣言等不法信息传播的温床,对人民生活、企业生产和社会稳定都带来了极大负面影响。微博、微信等在线社交网络由于用户规模庞大、消息即时传播、社会动员能力强,已经成为了国家治理的重要内容,加强治理刻不容缓,为此许多学者都展开了研究^[1]。如何高效地加强在线社交网络治理,已经成为促进国家治理体系和治理能力现代化的重要内容之一。摸清在线社交网络信息内在传播规律,把握其信息传播特征,是有针对性高效地加强在线社交网络治理的重要前提,本文对此展开了深入研究。

1 研究进展

近年来,国内外许多学者都对在线社交网络展开了研究,提出了众多消息传播模型。目前大多数消息传播模型都是源于经典的 SIS 和 SIR 传染病模型,比如早前的 DK 和 MK 谣言传播模型^[2-3]。王晨旭等研究了意见领袖对微博消息传播的影响^[4]。王辉等提出了基于记忆机制的 CSR 谣言传播模型^[5]。顾亦然等研究在线社交网络中谣言消息的传播和抑制问题^[6]。Wang 等多次研究了复杂网络、同质网络、新媒体时代等多种情况下谣言传播模型^[7-10]。Ma 等提出了积极和消极两极社会状态下的谣言传播模型^[11]。文献[12]研究了谣言的传播控制机制。王金龙等提出一种考虑用户相对权重的谣言传播模型^[13]。文献[14]提出了一种考虑网络媒体的谣言传播模型。曹玖新等提出了一种基于 K-核的社会网络影响最大化算法^[15]。李国良等构建了多个社交网络情况下消息传播模型^[16]。王祯骏等考虑了社交内容对消息传播的影响^[17]。Zhou 等考虑了位置对社交网络消息传播最大的影响^[18]。

然而上述大部分模型都基于均质网络假设,且仅考虑网络拓扑结构、网络在线特性、节点拓扑差异、消息特性等其中单个因素对节点消息传播的影响,没有综合考虑多种因素叠加影响。另外,有些模型简单将节点固化为传播者、潜伏者和免疫者等类型,存在一定的不科学性。社交网络是综合复杂网络,准确分析社交网络消息传播特性,需要综合考虑多种因素叠加效应。

2 本文模型

现实中,在线社交网络是一张不规则、非均匀、无标度动态网络,不同节点的连接度相差巨大,网络中节点的度分布呈现幂分布现象,富者愈富。另外,不是所

有节点都是实时在线的,非在线节点接收到的消息,由于消息时效性等因素,会影响节点上线后被转发的概率。基于上述原因,用均质、稳态的在线网络模拟谣言扩散存在一定局限性。

此外,经典传播模型简单地将节点固化为传播者、潜伏者、免疫者等存在一定不科学性。任何节点不会对所有消息都传播,也不会对所有消息不传播。特定节点针对所有消息不可能始终是传播者、潜伏者或免疫者。一个节点是否会继续传播某个消息,是节点充分考虑节点自身特性、消息发送节点特定和消息本身特性等因素后作出的综合判断结果。

节点自身特性是指节点自身转发消息的积极性,根据日常经验判断,摒弃节点自身兴趣爱好会影响消息传播概率因素之外,仅从节点度数角度考虑,相对而言,度数越高的节点或越低的节点都不太愿意转发别人的消息。

消息本身特性包括消息来源权威性、消息来源密集度,消息时效性等三个方面。相比而言,来自度数越高节点的消息,消息可信度越高,该消息则越容易被邻居转发。频繁接收到来自不同邻居节点相同的消息,该消息则越容易被接收节点相信,节点转发率也就越高。消息时效性越强,则该消息越容易被转发;不在线节点接收到的缓存消息,等节点上线后该消息被转发概率会大大降低。

基于上述原因,本文摒弃了之前已有研究将节点分为传播者、潜伏者和免疫者等简单固化分类,提出了在非均质网络中基于节点连接特性和消息特性的网络谣言传播模型,该模型构建方法如下:

(1) 根据网络节点连接度呈现幂分布特性^[21],本文构建网络中节点连接度分布函数如下:

$$y = \frac{c}{x^r} \quad (c, r \text{ 均为大于零的常数})$$

(2) 根据实际经验判断,粉丝量越多和越少的节点都不容易转发别人消息^[1]。基于节点转发消息积极性与自身度数呈现正态相关特性,本文构建了节点自身转发消息的积极性函数如下:

$$y_1 = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad x \geq 0$$

式中: y_1 表示节点从自己积极性单个因素考虑愿意转发他人消息的概率, μ 表示最愿意转发其他节点消息的节点平均度数, σ 表示绝大部分愿意转发消息的节点,其度数分布区间长度。根据正态分布规律,大部分节点度数 x 都落在该区间: $\mu - 3\sigma < x < \mu + 3\sigma$ 。

(3) 消息可信度一定程度和转发节点度有很大相关性。根据实际经验判断,度数高的节点转发的消息

相对而言可信度越高,也越容易被邻居节点转发。基于上述,本文构建了基于消息可信度的消息转发增强函数如下:

$$y_2 = cp \lg(10 + x)$$

式中: y_2 表示消息因消息来源可信度而被转发增强的倍数。 x 表示消息发送方节点的连接度数, cp 表示因消息来源节点度数差异外界对消息认可度的影响参数,简称节点度数差异影响参数。

(4) 频繁接收到来自不同邻居节点相同消息,则接收节点越容易信任该消息,该消息被节点转发的概率也就越高。基于上述,文本构建了基于消息重复接收量的消息转发增强函数如下:

$$y_3 = rp \ln(e + z)$$

式中: y_3 表示因为消息重复接收而被转发增强的倍数, z 表示相同消息被重复接收的次数, rp 表示因消息重复接收次数对消息转发概率的影响参数,简称消息重复接收差异影响参数。

(5) 借鉴文献[19]传播时间对消息的传播影响,本文构建了基于消息迟延的消息转发概率函数如下:

$$y_4 = e^{-\frac{\lambda}{tp}}$$

式中: y_4 表示因为消息时效性对消息转发概率降低比例, tp 表示消息接收时延对消息转发概率的影响参数,简称消息时延差异影响参数。 λ 表示消息时延,实际仿真模式中,为了平滑消息时延对消息转发概率影响,采用对数函数对物理消息时延参数 T 进行平滑处理,具体处理方式如下:

$$\lambda = \lg(1 + T)$$

(6) 综合节点转发消息的积极性、消息来源权威性、消息来源密集度,消息时效性等四个因素。最后,本文构建了节点接收到谣言消息后,根据综合决策,传播转发概率函数如下:

$$Y = y_1 \times y_2 \times y_3 \times y_4$$

3 实验仿真

模拟环境搭建。本文在借鉴斯坦福大学网络模拟器的基础上^[20],结合本文构建的传播模型,采用 C# 语言编程,构建了一个规模 8 000 个节点、消息有向传播的在线社交网络,其节点的度服从幂分布。该网络模型节点度服从幂分布,主要是考虑了传统模型节点度分布采用均匀分布,不能实际反映真实在线传播网络特征。节点度采用幂分布,使得模拟环境更加契合真实在线传播网络特征。

网络主要特征指标如表 1 所示。

表 1 仿真网络主要特征指标

参数名称	节点数	边数	平均度	网络密度	平均路径	平均聚类系数
参数指标	8 000	515 546	128.89	0.016 1	5.76	0.02

从图 1 可以看出仿真网络中节点度分布近似满足幂率关系,从表 1 可以看到仿真网络拥有较短的平均路径长度 5.76,说明了仿真社交网络较好地满足了真实社交网络无尺度和小世界特性^[21]。

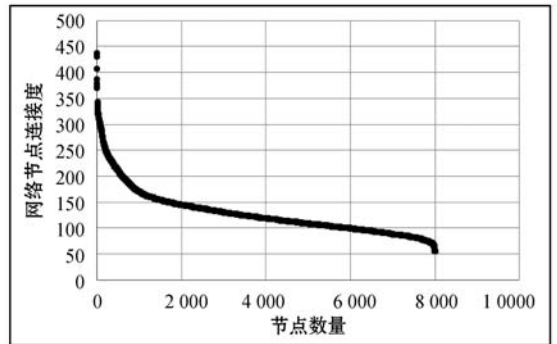


图 1 仿真网络中节点连接度分布

用 op 表示网络中节点平均在线概率,简称在线概率差异影响参数。 np 表示谣言消息包覆盖网络节点的比例。

定义:A1 事件表示谣言从网络中度数最大的节点首次向外传播。A2 表示表示谣言从网络中度数最小的节点首次向外传播。

本文模拟了 A1 和 A2 两种事件情况下,在线概率差异影响参数 op 、节点度数差异影响参数 cp 、消息重复接收差异影响参数 rp 、消息时延差异影响参数 tp 等四个参数变化对谣言消息传播扩散的影响,即谣言消息包不同传播周期内覆盖网络节点的比例 np 变化。

考虑到随机事件对仿真结果存在一定随机影响,为了让仿真结果尽可能逼近真实情况,本文所有仿真试验都模拟仿真 10 次,结果取 10 次仿真算术平均值。

(1) 仿真实验一:比较 A1 和 A2 两种事件情况下,在线概率差异影响参数 op 对谣言消息扩散的影响。 op 分别取 20%、30%、40%,结果如图 2 所示。其他参数设置如下: $cp = 1$ 、 $rp = 1$ 、 $tp = 1$ 。

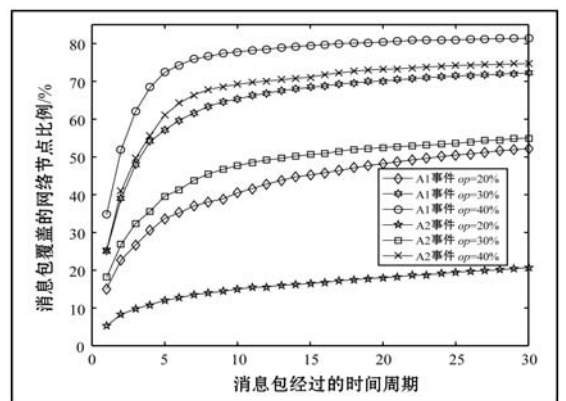


图 2 在线概率差异影响参数变化对谣言消息扩散影响

仿真结果显示:相同事件情况下,网络节点平均在线概率越高,谣言消息扩散越为迅速;在相同网络节点在线概率情况下,谣言消息从度数大的节点开始传播,要比度数小的节点开始传播,扩散快得多。此仿真实验说明,节点在线概率对谣言消息传播扩散速度具有显著影响。

对真实在线传播网络而言,仿真模拟情况反映了:针对相同谣言,在用户在线概率越高(例如,对于微信朋友圈而言,根据官方统计,晚上9点是在线浏览的用户数最高时候,凌晨3点是在线浏览的用户数最低时候),谣言传播快速越快;在线概率相同情况下,粉丝数量越多的用户(例如:微博大V)传播的谣言要比粉丝数量少的用户(例如:普通网民)传播的谣言扩散得快。

(2) 仿真实验二:比较 A1 和 A2 两种事件情况下,节点度数差异影响参数 cp 变化对谣言消息扩散的影响。 cp 分别取 0.5、1、2,其他参数设置如下: $op = 30\%$ 、 $rp = 1$ 、 $tp = 1$ 。结果如图 3 所示。

影响。 rp 分别取 0.5、1、2,其他参数设置如下: $op = 30\%$ 、 $cp = 1$ 、 $tp = 1$ 。结果如图 4 所示。

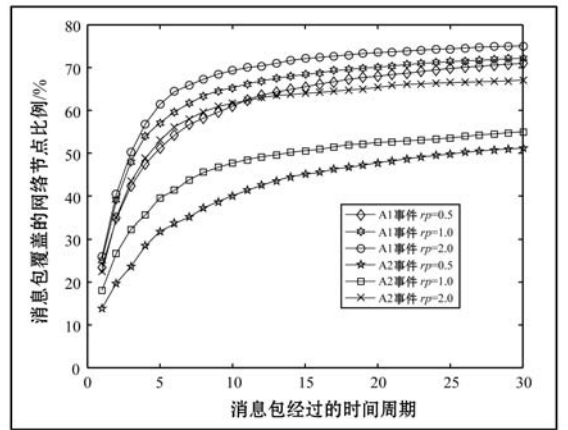


图 4 消息重复接收差异影响参数变化对谣言消息扩散影响

仿真结果显示:相同事件情况下,节点谣言消息重复接收度越高,谣言消息扩散越为迅速。在相同的消息重复接收影响节点信任情况下,谣言消息从度数大的节点开始传播,要比度数小的节点开始传播快得多。此仿真实验说明,相同谣言消息渠道来源数量,对谣言消息传播扩散速度具有显著影响。

对真实在线传播网络而言,仿真模拟情况反映了:用户接收到相同的谣言信息次数越多,越容易促成该用户转发该谣言,谣言传播也就越快;谣言消息重复接收次数相同的两个用户,粉丝数量越多的用户,其谣言传播扩散能力越强。

(4) 仿真实验四:比较 A1 和 A2 两种事件情况下,消息时延差异影响参数 tp 对谣言扩散的影响。 tp 分别取 0、1、2,其他参数设置如下: $op = 30\%$ 、 $cp = 1$ 、 $rp = 1$ 。结果如图 5 所示。

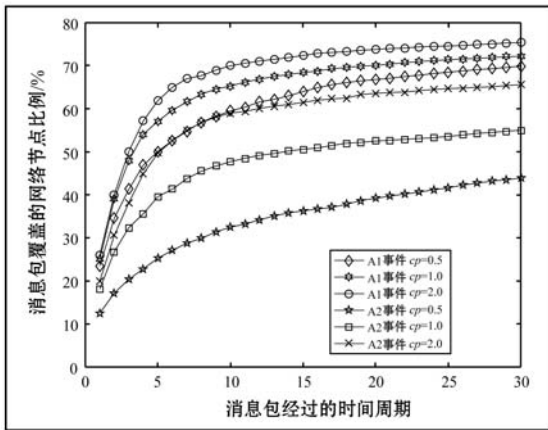


图 3 节点度数差异影响参数变化对谣言消息扩散影响

仿真结果显示:相同事件情况下,网络对节点对度数大的节点认可度越高,谣言消息扩散越为迅速。在相同的节点度数影响外部节点信任情况下,谣言消息从度数大的节点开始传播,要比度数小的节点开始传播,扩散得快得多。此仿真实验说明,大众对网络节点信任度高低,对谣言消息传播扩散速度具有显著影响。

对真实在线传播网络而言,仿真模拟情况反映了:社会知名度及公信力高的用户传播的谣言(例如:新闻媒体传播未经核实的新闻信息),要比社会知名度及公信力低的用户传播的谣言(例如:个人传播未经核实的信息),扩散得快得多;社会知名度及公信力相同的两用户,粉丝数量越多的用户,其谣言传播扩散能力越强。

(3) 仿真实验三:比较 A1 和 A2 两种事件情况下,消息重复接收差异影响参数 rp 对谣言消息扩散的

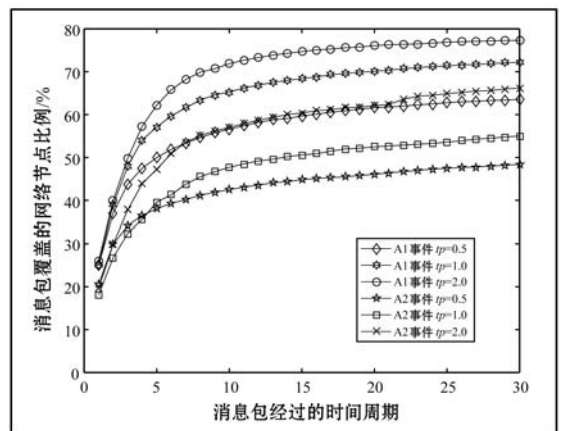


图 5 消息时延差异影响参数对谣言消息扩散影响

仿真结果显示:相同事件情况下,消息时延差异影响参数越高,谣言消息扩散越为迅速。在相同的消息时延差异影响参数情况下,谣言消息从度数大的节点开始传播,要比度数小的节点开始传播快得多。此仿真实验说明,谣言消息时间敏感度,对谣言消息传播扩

散速度具有显著影响。

对真实在线传播网络而言,仿真模拟情况反映了消息时延对谣言的传播能力有很大影响,时延越长的谣言,其传播能力越弱。接收到时延相同谣言的两个用户,粉丝数量越多的用户,该谣言越可能被传播。

另外从上述四个仿真实验中可以看到,A1 和 A2 两种事件情况下,度数越高的节点造谣,谣言消息传播扩散速度越快。仿真实验说明,造谣者自身社会关注度,对谣言消息传播扩散速度具有显著影响。

4 结 语

从上面四组对比实验中,可以清晰看到,造谣者的社会关注度、网络节点在线概率、大众对网络节点信任度的高低、相同谣言消息渠道来源数量、谣言消息时间敏感度五个因素,对谣言消息传播扩散速度具有显著影响。因此,加强网络空间谣言治理,需要综合根据上述五个影响因素分类施策,才能起到更好的网络谣言治理效果。

参 考 文 献

- [1] 李洋,陈毅恒,刘挺. 微博信息传播预测研究综述[J]. 软件学报,2016,27(2):247-263.
- [2] Moreno Y, Nekonec M, Pacheco A F. Dynamics of rumor spreading in complex networks [J]. Physical Review E, 2004, 69(6): 066130.
- [3] Nekovee M, Moreno Y, Bianconi G, et al. Theory of rumour spreading in complex social networks[J]. Physica A, 2007, 374(1): 457-470.
- [4] 王晨旭,管晓宏,秦涛,等. 微博消息传播中意见领袖影响力建模研究[J]. 软件学报,2015,26(6):1473-1485.
- [5] 王辉,韩江洪,邓林,等. 基于移动社交网络的谣言传播动力学研究[J]. 物理学报,2013,62(11):106-117.
- [6] 顾亦然,夏玲玲. 在线社交网络中谣言的传播与抑制[J]. 物理学报,2012,61(23): 544-550.
- [7] Wang J J, Zhao L J, Huang R B. SIRaRu: rumor spreading model in complex networks[J]. Physica A: Statistical Mechanics and Its Applications, 2014,398(15): 43-55.
- [8] Wang J J, Zhao L J, Huang R B. 2SI2R: rumor spreading model in homogeneous networks[J]. Physica A: Statistical Mechanics and Its Applications,2014,413(15):153-161.
- [9] Zhao L J, Wang J J, Chen Y C, et al. SIHR rumor spreading model in social networks[J]. Physica A:Statistical Mechanics and Its Application,2012,391(7):2444-2453.
- [10] Zhao L J, Cui H X, Qiu X Y, et al. SIR rumor spreading

model in the new media age[J]. Physica A: Statistical Mechanics and Its Applications, 2013,392(4): 995-1003.

- [11] Ma J, Li D D, Tian Z H. Rumor spreading in online social networks by considering the bipolar social reinforcement[J]. Physica A: Statistical Mechanics and Its Applications,2016, 447: 108-115.
 - [12] Wang Y Q, Yuan X Y, Wang J. A rumor spreading model with control mechanism on social networks [J]. Chinese journal of physics,2014,52(2):816-829.
 - [13] 王金龙,刘方爱,朱振方. 一种基于用户相对权重的在线社交网络信息传播模型[J]. 物理学报,2015,64(5): 71-81.
 - [14] Wang J, Wang Y Q. SIR rumor spreading model with network medium in complex social networks[J]. Chinese journal of physics,2015,53(1):83-93.
 - [15] 曹玖新,董丹,徐顺,等. 一种基于 K-核的社会网络影响最大化算法[J]. 计算机学报,2015,38(2):238-248.
 - [16] 李国良,楚娅萍,冯建华,等. 多社交网络的影响力最大化分析[J]. 计算机学报,2016,39(4):643-656.
 - [17] 王祯骏,王树徽,张维刚. 基于社交内容的潜在影响力传播模型[J]. 计算机学报, 2016,39(8):1528-1540.
 - [18] Zhou T, Cao J X, Liu B, et al. Location-based influence maximization in social networks [C]//Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. Melbourne, Australia, 2015:1211-1220.
 - [19] Qiu X, Zhao L, Wang J, et al. Effects of time-dependent diffusion behaviors on the rumor spreading in social networks [J]. Physics Letters A, 2016, 380(24):2054-2063.
 - [20] Stanford Network Analysis Project [EB/OL]. [2018-01-29]. <http://snap.stanford.edu/index.html>.
 - [21] 李勇军. 在线社交网络的拓扑特性分析[J]. 复杂系统与复杂性科学,2012,9(3):22-37.
- ~~~~~
- (上接第 236 页)
- [12] Ohtsu N. A Threshold Selection Method from Gray-Level Histograms[J]. IEEE Transactions on Systems Man & Cybernetics, 1979, 9(1):62-66.
 - [13] 百度百科. 数学形态学 [EB/OL]. <http://baike.baidu.com/view/852283.htm>.
 - [14] 刘松涛,殷福亮. 基于图割的图像分割方法及其新进展 [J]. 自动化学报,2012,38(6): 911-922.
 - [15] Mouelhi A, Sayadi M, Fnaiech F. Automatic segmentation of clustered breast cancer cells using watershed and concave vertex graph[C]//2011 International Conference on Communications, Computing and Control Applications (CCCA). IEEE, 2011:1-6.