

社交网络上基于信息驱动的行为传播研究

陈玟宇¹, 贾 贞^{1,2}, 祝光湖³

(1. 桂林理工大学理学院 广西 桂林 541004; 2. 广西空间信息与测绘重点实验室 广西 桂林 541004;

3. 桂林电子科技大学数学与计算科学学院 广西 桂林 541004)

【摘要】针对社交网络上信息与行为的交互作用与协同传播的特点,建立了基于信息驱动的行为传播模型。该模型综合反映了社交网络上信息与行为传播的交互性,以及信息传播的记忆性、时效性与社会加强效果等多种因素对传播的影响。研究发现,信息的时效性、社会加强效果、信息的转播率等模型参量,以及源节点的度与核数、节点的度及与源节点距离等网络参量,都在一定程度上影响行为传播的速度和范围,但各参量在不同环境下发挥不同的作用。

关键词 行为传播; 信息时效; 社交网络; 社会加强

中图分类号 O41; TN92

文献标志码 A

doi:10.3969/j.issn.1001-0548.2015.02.002

Information-Driven Behavior Spread on Social Networks

CHEN Wen-yu¹, JIA Zhen^{1,2} and ZHU Guang-hu³

(1. College of science, Guilin University of Technology Guilin Guangxi 541004;

2. Guangxi Key Laboratory of Space Information and Mapping Guilin Guangxi 541004;

3. School of Mathematics and Computing Science, Guilin University of Electronic Technology Guilin Guangxi 541004)

Abstract In this paper, a novel behavior spreading model is proposed, which incorporates the reciprocal interaction and collaborative spreading of information and behavior on social networks. Various network parameters and transmission factors, describing the interaction of information and behavior, the memory and timeliness of information, the social reinforcement and so on, are incorporated into the spreading process. The simulation results indicate that the involved parameters have important influence on the transmission range and speed, but their functions are varied according to different circumstances. Our research shows that the behavior spreading on social networks not only possesses some common laws similar to the transmission of information and diseases, but also exhibits its unique characteristics.

Key words behavior spreading; information timeliness; social network; social reinforcement

现实生活中广泛存在各种传播现象,如疾病在人群中的传播、计算机病毒在因特网上的蔓延、信息和舆论在各种社交网络上的扩散等,可见复杂网络上的传播动力学已成为许多领域学者共同关注的研究热点。近年来的研究表明,除了疾病、信息、舆论等传播主体可以在各种网络上传播以外,人的观念、行为、甚至情绪等也可以在社交网络上传播。文献[1-4]的研究发现失眠、吸烟、吸毒等行为都能在基于人际关系的社交网络上传播,发现如果一名学生有一个吸毒的朋友,那么他接触毒品的几率会增加42%,失眠和使用毒品的传染性甚至能够传播到4层人际关系。因此,在研究复杂网络上的传播动力学时,有必要把研究视觉扩大到更大的范围,研

究诸如观念、行为、情绪等更多的传播主体在网络上的传播现象,无疑具有重要的现实意义。

在过去的研究中,人们更多地研究了疾病、信息和舆论等的传播现象,并基于平均场方法建立了各种网络传播模型^[5-11],但对于行为在社交网络上的传播现象研究甚少。2010年,文献[12]发表了关于行为传播的研究成果,引起人们对行为传播研究的关注。通过对在线社会网络上的健康行为传播实验研究发现,行为与疾病传播具有不同规律,由于行为传播具有社会强化效应,导致行为在拥有更多冗余连接的规则网络上比小世界网络上传播更快更广。最近,文献[13]提出一个信息传播模型,模型考虑了信息传播的记忆性、社会强化性和传播链路使用一

收稿日期: 2014-06-22; 修回日期: 2015-02-07

基金项目: 国家自然科学基金(61164020)

作者简介: 陈玟宇(1988-),女,硕士,主要从事复杂网络方面的研究。

次性原则,建立了信息传播演化机制,并且分别在规则网络、小世界网络和随机网络上模拟了信息传播过程。研究发现,当个体接收到一次信息就确认的概率较小的时候,信息在规则网络上比随机网络上传播更快更广,并且在小世界上能够达到最好的传播效果,该研究部分支持了文献[12]中发表的结果。文献[14]在文献[12]提出的“社会加强”概念的基础上建立了一个行为传播模型,着重考察了社会加强效果对行为传播的影响,并通过仿真实验验证了文献[12]中关于行为传播实验的结果。文献[15]进一步研究了社会强化效应对信息传播的影响,验证了文献[13]的一些研究结果。然而,所有这些研究都没有考虑到信息与行为传播的交互作用,没有把二者结合起来研究。事实上,信息与行为之间是相互联系、相互影响的,人们的许多行为都是受到相关信息的刺激而产生的。如在股市中,当人们听到某上市公司利好或不好的消息时,会产生购买或抛售该公司股票的行为;当一个人看到周围的朋友中有很多人吸烟或酗酒时,他也会受到感染而学着去吸烟或酗酒;当顾客听到或看到人们对某种商品的好评时,也可能产生购买该物品的冲动等。这些行为的产生,体现了信息与行为的交互作用和协同传播的特点。一方面,个体的行为会受到相关信息(对某商品的好评)的刺激和影响,当这种刺激累积到一定程度就可能激发该个体产生相应行为(如购买该物品),这种信息的累积作用就是一种社会强化效应。另一方面,个体的行为本身也会给周围的个体一种示范信息,从而刺激他人产生效仿行为,这样,信息与行为在网络上交互影响并传播开来。鉴于此,本文提出一个新的行为传播模型。与以往不同的是,该模型反映了信息与行为的交互作用与协同传播的特点,并且以一个真实Facebook网络为例,重点研究在线社交网络上的信息与行为传播规律。

1 模型描述

通过在线社交网络(如Facebook、Twitter、腾讯QQ、微信等),人们可以在个人空间或朋友圈上发布与某种行为相关的信息,如发布某时装的图片、对某文章评论、对某商品的赞誉等,他的朋友可以收到并转播此信息,当网络中的个体受到这些信息刺激后,就可能产生相应行为(如穿着此时装、阅读此文章、购买此商品等),并且接收到的信息累积次数越多,产生此行为的概率越大。同时,这种行为本身对周围的人也起到示范作用,也会传递相应的

信息,从而刺激他人产生效仿行为。基于这种信息与行为交互作用和协同传播的特点,本文提出一个社交网络上基于信息驱动的行为传播模型。先给出模型中的几个重要概念:社会强化效应是指个体对接收的信息具有记忆性,接收信息的次数对驱动个体产生相应行为具有累加作用,个体接收信息次数越多,产生相应行为的概率越大。转播率是指个体接收到一条信息并把它转播出去的概率,它的大小反映网络群体(平均水平而言)对该信息或行为感兴趣的程度。信息时效是指信息对个体产生的刺激作用是有限度的,当个体接收信息的次数达到某个上限而不能产生相应行为,那么就认为再多的信息刺激对他也不起作用了,表明该个体对此行为具有了免疫力,这时称个体转化成了疲惫状态。本文把个体所能接收信息次数的上限称为信息时效,它的大小反映网络群体(平均水平而言)对信息和行为的兴趣持续的时长。

1.1 模型假设

考虑一个社交网络,节点代表社交个体,边代表个体之间的朋友或关注关系。

假设 1 信息 Q 与行为 B 相对应。信息 Q 可以刺激节点产生行为 B , 个体接收信息 Q 的次数越多, 产生行为 B 的概率越大, 且行为者(具有行为 B 的节点)本身也传播信息 Q 。

假设 2 网络中所有节点均以相同的概率 λ (称为转播率)转播信息 Q , 且每条边只转播信息一次。

假设 3 信息 Q 具有有限时效性。如果节点接收信息 Q 的次数累积达到 L 次(称 L 为信息时效)还没有产生行为 B , 那么就认为信息 Q 对该节点已失效, 该节点不再参与信息 Q 的转播, 也不再产生行为 B 。

假设 4 在每一时间步, 网络中的节点都处于下列4种状态之一。

1) 未知信息态(记为 U 态): 节点不知道或未见到信息 Q , 类似于SIR模型中的易感态。 U 态节点只能接收信息而不具有传播信息的能力。

2) 已知信息态(记为 K 态): 节点已知信息并且转播信息 Q 。 K 态节点可以多次接收信息 Q 并把它转播出去, 但没有产生相应行为 B 。

3) 行为态(记为 B 态): 节点产生了行为 B , 同时也可以传播信息 Q 。

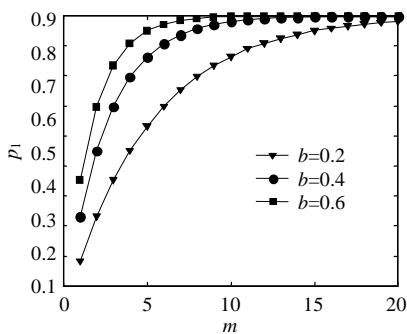
4) 疲惫态(记为 E 态): 信息刺激对此类节点失效, 它不再参与信息转播, 也不会产生行为 B , 相

当于SIR模型中的免疫态。

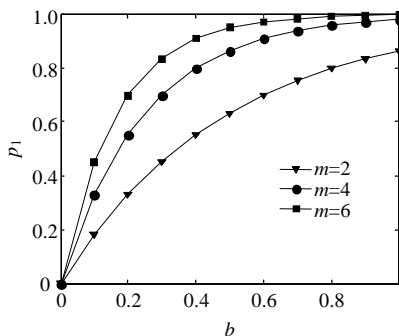
1.2 传播过程

初始时刻, 随机地从网络中选取一个节点 v_0 作为源节点, 携带信息 Q , 而其余节点均处于 U 态。源节点 v_0 在下一时间步以概率 λ 把信息 Q 转播给它的邻居节点。在 t 时刻, K 态或 B 态节点 v_i 以概率 λ 把信息 Q 转播给它的 U 态或 K 态邻居节点 v_j , 节点 v_j 每接收一次信息, 它接收信息的次数就累加一次, 这体现个体对信息具有记忆性。

在每一时间步, 所有节点的状态都同步更新一次: U 态节点或以概率 p_1 (见式(1)) 转化为 B 态, 或以概率 p_2 (见式(2)) 转化为 K 态; K 态节点或以概率 p_1 转化为 B 态, 或以概率 p_2 停留在 K 态, 或以概率 p_3 (见式(3)) 转化为 E 态。重复上述步骤, 直到网络中 B 态节点的密度达到稳定状态, 即网络中不再产生新的 B 态节点为止。



a. 状态转移概率 p_1 随接收次数 m 的变化关系



b. 状态转移概率 p_1 随接收次数 b 的变化关系

图1 状态转移概率 p_1 随参数 b 和 m 的变化节点的状态转移概率算法如下:

设 U 态或 K 态节点 v_i 在 t 时刻接收到信息的累积次数为 $m(t)$, 当 $1 \leq m(t) \leq L$ 时, v_i 转化成 B 态的状态转移概率为:

$$p_1 = 1 - e^{-bm(t)} \tag{1}$$

因此, 当 $1 \leq m(t) < L$ 时, v_i 仍处于 K 态的概率为:

$$p_2(t) = e^{-bm(t)} \tag{2}$$

当 $m(t) = L$ 时, v_i 转化为 E 态的概率为:

$$p_3 = e^{-bL} \tag{3}$$

式中, 参数 b 称为社会加强因子。概率 p_1 随着参数 $m(t)$ 和 b 的增大而增大, 如图1所示, 它反映了信息的记忆性和社会强化效应的对行为产生的激励作用。

2 仿真与分析

本文采用真实的在线社交网络Facebook上的用户数据信息^[16], 把注册用户作为节点, 用户之间的关注关系作为连边, 这样就构成了一个Facebook网络。该网络的总节点数 $N = 4\ 039$, 共有84 243条连边, 平均度为41.7, 其度分布如图2所示, 近似幂律分布^[17]。

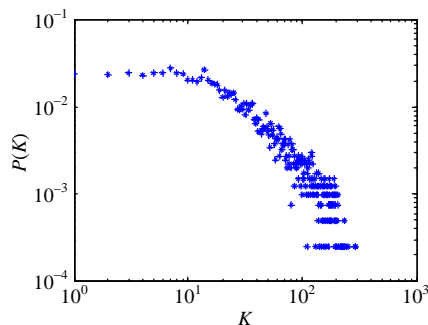


图2 Facebook网络的度分布图

下面将在这个Facebook网络上, 按照模型描述的传播过程进行仿真实验, 分别考察各种模型参量和网络参量对行为传播的影响。网络中 B 态节点在 t 时刻的密度记为 ρ_t , 稳态下的密度记为 ρ_∞ , 通过考察 ρ_t 到 ρ_∞ 的演化来获得网络上的行为传播规律。为消除随机因素的干扰, 所有实验结果均取10次独立重复实验的平均值。

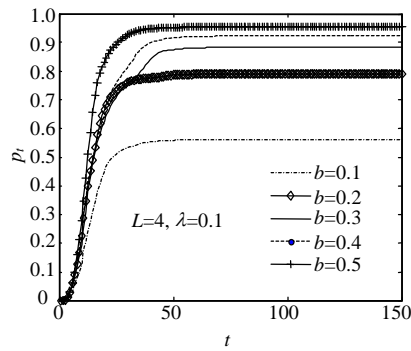
2.1 模型参量对传播的影响

首先, 考察社会加强因子 b 和信息转播率 λ 对行为传播的影响。

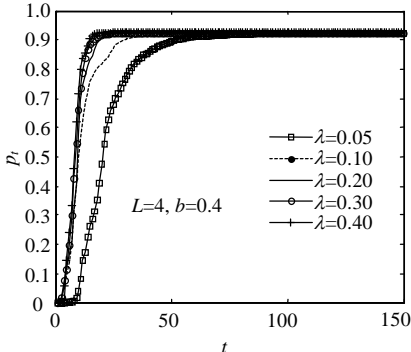
图3a和3b分别展示了参数 b 、 λ 在不同取值下 B 态节点密度 ρ_t 随着时间的演化过程。由图3a可见, 在参数 b 的不同取值下, ρ_t 趋于稳定状态的时间相差不大, 但其终态值 ρ_∞ 随着 b 的增大而增大, 表明社会加强因子对行为传播的范围有显著影响, 但对传播的速度影响不大。由图3b可见, 随着转播率 λ 的增大, ρ_t 趋于稳定状态所需的时间减少, 即行为在网络中的传播速度加快, 但对终态值 ρ_∞ 的影响不大。这表明在当前参量取值下转播率 λ 只影响行为传播的速度, 而不影响传播的范围。

从图3c可以进一步考察参量 b 和 λ 对传播的影响。观察图3c中 ρ_∞ 的变化, 可以把参数空间 $R(\lambda, b)$

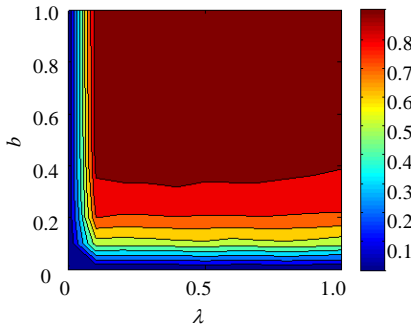
划分为3个显著差异的区域: 区域 I = $\{(b, \lambda): 0.38 < b \leq 1, 0.1 < \lambda \leq 1\}$, 区域 II = $\{(b, \lambda): 0 < b \leq 1, 0 < \lambda \leq 0.1\}$, 区域 III = $\{(b, \lambda): 0 < b \leq 0.38, 0.1 < \lambda \leq 1\}$ 。在区域I上, b 和 λ 对传播范围的影响极小, 表明当社会加强因子 b 和信息转播率 λ 较大时, 行为几乎扩散到整个网络, ρ_∞ 不再随 b 和 λ 的变化而变化, 这与SIR模型中疾病传播规律一致, 即当感染概率较大时, 传播都会大范围蔓延开来^[18]。在区域II上, 可见 λ 对传播有显著影响, 但 b 的影响不显著, 表明在此区域上信息转播率 λ 是影响传播范围的主要因素; 相反, 在区域III上, ρ_∞ 随着 b 的增大而增大, 而 λ 对传播的影响不大, 表明在此区域上社会加强效果 b 是传播的主要影响因素。



a. 不同 b 下, ρ_i 随时间的演化关系



b. 不同 λ 下, ρ_i 随时间的演化关系



c. 参数空间 $R(\lambda, b)$ 上 ρ_∞ 的变化色彩图

图3 参数 b 和 λ 对行为传播的影响

综上, 当社会加强因子 b 和信息转播率 λ 其中一个较小时才能体现出它们对传播的影响, 且较小

者为传播的主要影响因素; 但随着二者取值的增大, 其影响力逐渐减弱; 当二者的取值同时增大到一定值时, 行为迅速在网络上传播开来, 并达到饱和状态, 所以它们不再影响传播的范围, 但行为传播的速度随着二者的增大而加快。

其次, 考察信息时效 L 对行为传播的影响。

信息时效 L 是节点所能接收信息的最大次数, 也是节点转化为 E 态的转折点, 即只要节点接收信息次数到达 L 而没有转化为 B 态, 那么该节点就转化为 E 态, 并不再参与信息与行为的传播。图4展示了不同的信息时效对行为传播的影响, 可见, ρ_i 的增长速度变化不大, ρ_∞ 随着 L 的增大而增大的, 这表明信息时效只影响传播范围而不影响传播速度, 信息时效越长, 行为传播的范围越大, 这体现了病毒或信息传播的一般规律。显然, 人们对信息或行为的兴趣持续时间越长, 传播的范围自然就越大。

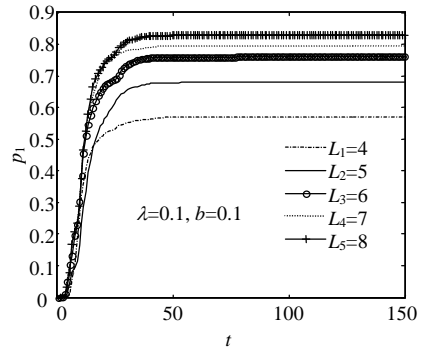


图4 不同的信息时效下 ρ_i 的演化图

2.2 网络参量对传播的影响

首先考察源节点的度与核数对行为传播的影响。这里, 先给节点的核数的定义^[19]。

定义 一个网络中所有度值不小于 k 的节点组成的连通片称为 k -核 (k -core); 若节点存在于 k -核中, 而不存在于 $(k+1)$ -核中, 则称该节点核数为 k 。

在社会网络分析中, 节点的核数往往代表节点在网络中的重要程度。节点的核数越大, 则该节点在网络中的地位越重要^[20]。

图5展示了不同传播参量下, 源节点分别在同度数但不同核数、同核数但不同度数下, B 态节点密度 ρ_i 随着时间的演化状态。比较图5a~5d可见, 源节点的度数或核数越大, ρ_i 到达稳态的时间越短, 表明行为扩散的速度越快, 但最终扩散的范围 ρ_∞ 与源节点在网络中的度数和核数大小无关, 即源节点在网络中的地位只影响传播的速度而不影响传播的范围。这一结果与疾病在网络上的传播规律相差甚远^[18,20], 这是由于信息和行为传播与疾病传播在传播机制上的根本差异所造成的, 也反映了在线社交

网络上的信息与行为传播的特别之处。从一个具有大量粉丝的明星发出的信息或倡导的行为，其传播速度相比普通人发出的要更快些，但随着时间的增长，传播的最终范围却与倡导者的地位无关，即无论是明星还是草根发出的信息或行为，都同样可以在网络上传播开来，其传播范围的广度应该更多的依赖于网络群体对信息或行为的兴趣程度和持续时间，这点体现在转播率和信息时效性对传播的影响上。

其次考察被传播节点的度和位置与行为传播的相关性。

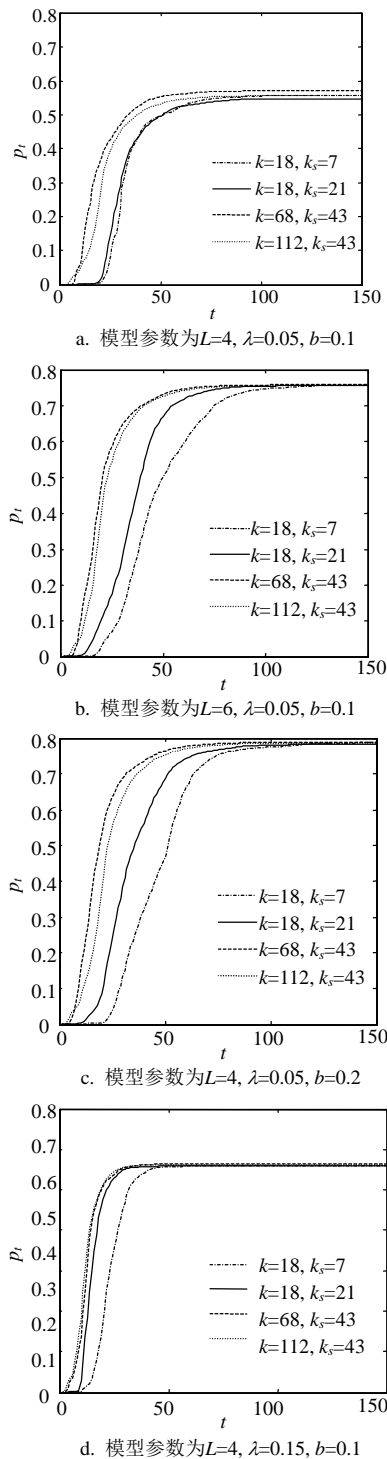
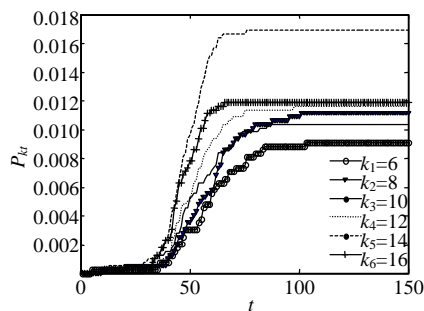
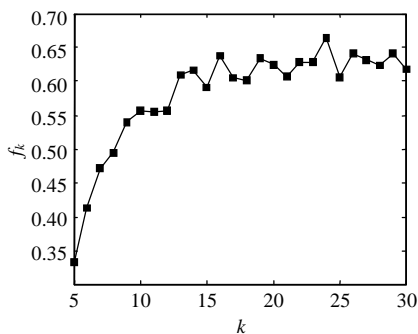


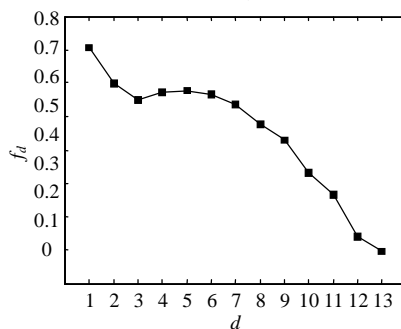
图5 源节点的不同度和核数下 ρ_t 的演化



a. 不同度的 B 态节点的相对密度 ρ_{kt} 随时间变化关系



b. 相对转化率 f_k 与度 k 的关系



c. 相对转化率 f_d 与距离 d 的关系

图6 被传播节点的度与源节点的距离对传播的影响 (参数取值为 $L=4, b=0.1, \lambda=0.05$)

图6a展示了不同度的 B 态节点的相对密度 ρ_{kt} 的演化状态。可见， ρ_{kt} 达到稳定状态所需的时间随着 k 的增大而减小， ρ_{kt} 的终值大致随着 k 的增大而增大。表明网络中度大的节点更趋向于转化为行为者，且转化的速度更快。图6b展示了不同度的节点中最终相对转化率 f_k 与度 k 的相关关系，即度为 k 的节点中行为态所占比例随度值的变化关系。可见，当节点的度 $k < 15$ 时， f_k 随着 k 的增大而快速增大，即 f_k 与 k 呈正相关关系，这与图6a的结果相互印证。但当度数 $k > 15$ 以后，度 k 对 f_k 影响变化相对较小。这表明，在社交网络上，当你的好朋友个数越多时，就越容易受到感染而成为行为者，但当朋友个数达到某个值(这里约为15)以后，受影响的程度将趋于稳

定。图6c展示了B态节点的相对转化率 f_d (距源节点距离为 d 的节点中行为态所占比例)与距离 d 的关系。可见,二者呈负相关关系,即与源节点距离越近节点的相对转化率越高。显然,距源节点越近的节点受影响越大,正体现了所谓的“近朱者赤,近墨者黑”的常理。此外,从图6c曲线的下降趋势还可看出,在距源节点三层关系以内,行为传播的强度最大,四到六层关系传播程度相对稳定,以后关系层次越远传播程度衰减越快。表明源节点的行为能较大可能的传播到三层关系,某种程度上印证了James Fowler的实证数据分析结果^[1-4],也说明可以通过控制与源节点三层关系内的接触率来控制行为的广泛流行。

3 结论

本文提出了一个基于信息驱动的行为传播模型,综合反映了社交网络上信息与行为的交互作用和不同于疾病的传播机制。以一个在线社交网络(Facebook网络)为实例的仿真研究表明:信息时效性、社会加强效果、信息的传播率等模型参量在一定范围内对行为传播具有显著影响,但他们在不同参数环境下的对行为传播的影响是有差异。一方面,当社会加强因子与信息的传播率都取较大值时,只影响传播的速度,而对传播范围的影响由于行为者密度迅速达饱和状态被淹没;但当二者中有一个取较小值时,较小值者成为行为传播范围的主要影响因素。另一方面,源节点的度和核数越大时,其行为在网络中传播快,但不影响传播的最终范围;网络中节点相对转化率与它的度都呈正相关关系,与它距源节点的距离呈负相关关系,即网络中度越大的节点和越靠近源节点的节点受感染而转化成行为者的可能性越大,而且行为在源节点的三层关系内的传染性最强,所以,减少网络个体与源节点的三层关系内的接触是减少行为流行的有效方法。此外,研究中还发现,本文模型中的传播率的临界值 λ_c 接近于0,推测这是由于网络的度分布近似幂律分布的缘故。行为与疾病在无标度网络上的传播临界值是否也有类似的结论^[19],这点还有待于进一步探索研究。

在线社交网络上的行为传播既有不同于疾病传播的独特个性,也有疾病传播的某些一般规律。使人们对行为的传播机制和演化规律有了更深入的探索,这也有助于更深刻理解、预测和控制社交网络上流行行为的形成与演化方向,同时,对于某些不

良行为(如吸毒、吸烟等)在社会上扩散的控制和某些良好行为(如运动、阅读等)的形成与倡导都具有重要意义。

参 考 文 献

- [1] CHRISTAKIS N A, FOWLER J H. The collective dynamics of smoking in a large social network[J]. *New England journal of medicine*, 2008, 358(21): 2249-2258.
- [2] ROSENQUIST J N, MURABITO J, FOWLER J H, et al. The spread of alcohol consumption behavior in a large social network[J]. *Annals of Internal Medicine*, 2010, 152(7): 426-433.
- [3] MEDNICK S C, CHRISTAKIS N A, FOWLER J H. The spread of sleep loss influences drug use in adolescent social networks[J]. *PloS one*, 2010, 5(3): e9775.
- [4] ROSENQUIST J N, FOWLER J H, CHRISTAKIS N A. Social network determinants of depression[J]. *Molecular psychiatry*, 2010, 16(3): 273-281.
- [5] PASTOR-SATORRAS R, VESPIGNANI A. Epidemic dynamics and endemic states in complex networks[J]. *Physical Review E*, 2001, 63(2): 066117.
- [6] PASTOR-SATORRAS R, VESPIGNANI A. Epidemic spreading in scale-free network[J]. *Phys Rev Lett*, 2001, 86(4): 3200-3203.
- [7] PASTOR-SATORRAS R, VESPLGNANI A. Epldemies and immunization in scale-free networks[M]//BORNHOLDT S SEHUSTER H G. *Handbook of Graph and Networks, From the Genome to the Internet*. Berlin: WILEY-VCH, 2003: 111-130.
- [8] CASTELLANO C, PASTOR-SATORRAS R. Thresholds for epidemic spreading in networks[J]. *Physical review letters*, 2010, 105(21): 218701.
- [9] GARAS A, ARGYRAKIS P, ROZENBLAT C. Worldwide spreading of economic crisis[J]. *New Journal of Physics*, 2010, 12(11):113043.
- [10] ZHOU J, LIU Z H. Epidemic spreading in communities with mobile agents[J]. *Physical A*, 2009, 388(7): 1228-1236.
- [11] 张海峰, 张文耀, 孙桂全, 等. 信息的滞后性诱导传染病的周期爆发[J]. *中国科学: 物理学 力学天 文学*, 2012, 42(6): 631-638.
ZHANG Hai-feng, ZHANG Wen-yao, SUN Gui-quan, et al. Time-delayed information can induce the periodic outbreaks of infectious diseases[J]. *Sci Sin-Phys Mech Astron*, 2012, 42(6): 631-638.
- [12] DAMON. The spread of behavior in an online social network experiment[J]. *Science*, 2010, 329(5996): 1194-1197.
- [13] LÜ L, CHEN D B, ZHOU T. The small world yields the most effective information spreading[J]. *New Journal of Physics*, 2011, 13(12): 123005.
- [14] ZHENG M, LÜ L, ZHAO M. Spreading in online social networks: The role of social reinforcement[J]. *Physical Review E*, 2013, 88(1): 012818.