

社交网络中种子用户分布对产品扩散的影响

汤胤¹, 徐永欢¹, 陈剑²

(1. 暨南大学管理学院, 广东 广州 510632;

2. 中山大学新华学院管理学院, 广东 广州 510507)

摘要: 在社交网络中同样数量的种子用户群体在社交网络中的不同分布, 往往对新产品的传播效率造成影响. 探讨在具备小世界特性的社交网络中, 种子用户分布与社交产品扩散速度的关系. 设计了基于阈值模型和基于节点相关性的仿真实验, 基于大规模实际社交网络的数据, 采用计算机仿真模拟方法来从微观角度进行观察. 研究发现, 社交网络中种子用户紧密度与社交产品扩散速度正相关, 但影响加速递减.

关键词: 创新扩散; 复杂社会网络; 种子用户分布; 网络紧密度; 扩散速度

中图分类号: TP713.36 **文献标识码:** A **文章编号:** 1000-5781(2021)02-0145-12

doi: 10.13383/j.cnki.jse.2021.02.001

Influence of seed user distribution on social product diffusion in social network

Tang Yin¹, Xu Yonghuan¹, Chen Jian²

(1. Management School, Jinan University, Guangzhou 510632, China;

2. Management School, Xinhua School, Sun Yat-sen University, Guangzhou 510507, China)

Abstract: Different seed user distribution in social network has different influences on new product spreading. This paper discusses the relationship between seed user distribution and velocity of new product spreading in online social networks with small-world characteristics. Two micro-level simulations based on threshold model and node correlations with the mass real-world social network data are conducted to test the hypothesis. The result verifies that compactness of seed users has positive correlations with the spreading speed of the product, and can accelerate the decrease effect on the diffusion rate of the product.

Key words: innovation diffusion; complex social network; seed user distribution; compactness; diffusion rate

1 引言

自 Rogers 从提出创新扩散以来, 创新的范围已经从新技术扩大到新产品, 新方法, 或者新观念、群体现象以及不同个体的行为等, 而“扩散”一词则涵盖了有计划性的传播和自发性无计划传播两层含义, 并且是一个过程, 包含创新、渠道、时间以及社会体系等四个要素^[1,2]. 实践中, 移动社交应用市场的竞争已经非常激烈, 对于很多新兴的社交应用(app 等各种形式), 鉴于其社交属性, 用户使用的基础是必须有一定的用户数量, 若不能在一定期间内集聚一批用户群体(即形成羊群效应), 产品很可能失败. 从很多初创项目中观察到,

收稿日期: 2018-04-10; 修订日期: 2019-05-06.

基金项目: 广州市产学研协同创新重大专项资助项目(201802010034); 广东省哲学社会科学“十三五”规划资助项目(GD17-CGL10); 国家自然科学基金资助项目(71771104).

社交应用本身可以满足用户需求前提下, 初始种子用户群体在整体人群中比较分散, 可能会引起传播障碍, 影响用户数增长; 而另一方面, 若初始种子用户群体非常紧密, 难免有覆盖不到的区域, 是否同样会影响传播速度? 图 1 刻画了同样数量种子用户群体在社交网络中的不同分布(x 代表每个小团体中节点的个数, y 代表彼此不直接相连的小团体个数), 到底哪种更有助于用户增长呢? 将这个问题扩大至共享经济应用(如共享单车)中, 基于物理空间的初始资源投放也提出了类似的问题。

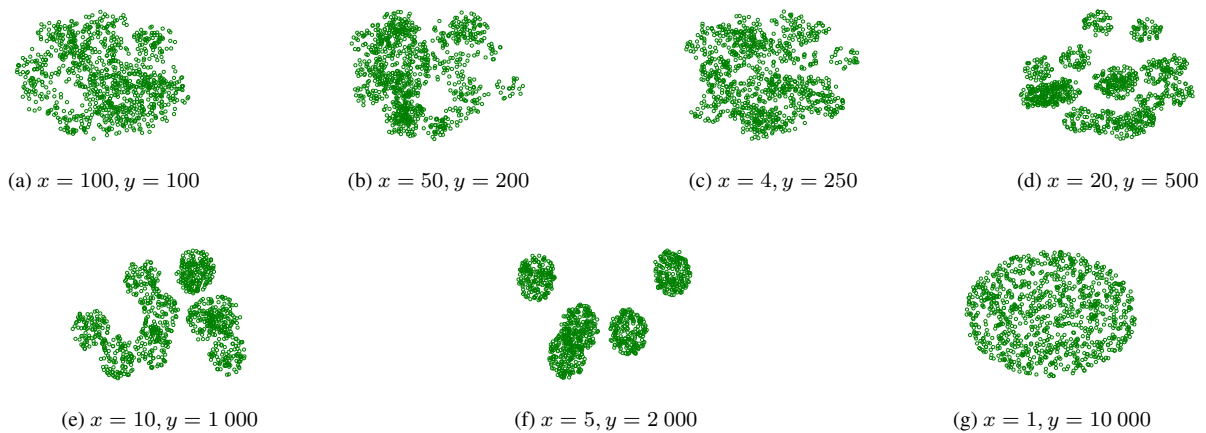


图 1 同样数量种子用户群体(初始资源)在社交网络中的不同分布

Fig 1. Different distribution of seed user of same size in social network

理论研究方面, 早期往往采用定性或者问卷调查方法, Bass 混合扩散模型最早采用仿真方式, 成功地预测了耐用消费品销量并发现影响创新扩散的重要因素^[3]. 但 Bass 模型仅反映宏观特征, 对于社会系统内单独个体交互作用后的行为的非线性的涌现缺乏对创新扩散微观机理的考察, 而且模型忽略了采纳时间和交互时间, 包括其后 Bass 模型族系均有这个问题. 而创新扩散的本质是大量潜在消费者的微观采纳决策集中涌现为宏观层面的行为模式^[4]. 大量的研究证实, 复杂社会网络与创新扩散之间存在相互影响. 一方面, 社会网络(如人际关系网络)是创新扩散的重要渠道. 社会网络中个体对某种创新(社交应用、共享单车应用等)的选择容易受到与之关联的邻居个体选择的影响, 即存在网络效应^[5-7]. 宏观的扩散模型无法考虑采纳者的异质性, 忽略了微观层面消费者个性与社会交互等特点对产品扩散的影响, 从而使其在现实应用中出现了一定的局限. 尤其是随着移动互联网的普及以及社会化媒体的出现, 使得潜在的消费个体决策行为越来越多地受到来自周边其他个体选择的影响. 忽略了这种交互作用, 将难以准确的描述和解释具有网络外部效应的创新扩散。

20 世纪 90 年代以后, 阈值模型、渗流模型、元胞自动机模型、Agent-based 模型以及多智能体模型等开始从微观角度刻画创新扩散的动力学过程, 并在创新扩散的前期预测方面体现出有效性和稳定性的特征^[8-13]. Granovetter 提出的门槛模型(阈值模型)充分考虑了信息接收者的特性和信息传播具有记忆效应, 能够更为直观地展示社会系统结构对创新扩散的影响, 特别适用于社会动态问题的仿真研究^[14,15], 如郭静等^[16]基于此模型加上影响力传播权重算法, 指出度量用户间影响力对商品营销和推广具有重要作用. 赵良杰等运用微观扩散模型仿真分析消费者交互作用对网络效应产品扩散的影响^[17].

除了考虑微观个体, 网络结构的差异也被证明可以促进或者阻碍创新在这个社会中的扩散. 真实社会网络呈现 Barabási-Albert 网络的无标度特性和 Watts-Strogatz 小世界特性, 其特征结构显著影响新产品在整个网络中的扩散^[18-20]. 高聚类网络具有很多的冗余连接拓扑结构, Centola 等发现行为在高聚类的网络上传播速度更快, 与社交产品的内秉社交属性仍然有所区别, 也未发现其增长瓶颈^[21]. Delre 等^[22]以小世界网络为背景, 发现最优的目标市场选择策略是选择多个相互分散、高度聚类的小型消费者群体. 赵保国等^[23]认为网络效应的存在不仅改变了传统扩散模式, 而且对个体采纳行为有重要影响, 进而指出直接网络效应和

间接网络效应对竞争性产品扩散具有阶段性影响。赵正龙等^[24]基于个体的差异化选择特征角度,分析了复杂社会网络上的扩散过程及其影响因素,有助于企业有效管理和控制具有差异化选择特征的扩散。Peres等^[25]对比分析了不同网络结构对创新扩散生命周期的影响效应,并指出消费者网络是诸多环境和关系(如网络外部性、消费者口碑、社会信号等)叠加作用的结果。杜宾^[26]根据社会学习的主体认知过程和消费者网络结构的线上线下差异性,探索多层次网络叠加条件下的扩散规律。孙耀吾等^[27]基于小世界网络研究企业联盟知识扩散特征,发现集聚系数增加和特征路径长度减少可以加快扩散效应。

有学者开始从网络视角出发,结合传染病模型来研究创新扩散的问题。Choi等^[28]构建了基于SIS传染病模型的平均场扩散过程,利用随机占优理论分析了网络平均度、度分布异质性以及“邻居效应”对扩散的影响。病毒传播模型为创新扩散的研究提供新的研究视角和方法,从而使得基于复杂网络的计算机仿真的研究动态演化成为可能,仿真过程也更贴近现实网络的创新扩散过程。

随着研究的深入,有部分学者开始关注种子顾客对于初期创新扩散以及创新在整个网络中的扩散的重要作用,并且更多从影响力的角度来研究意见领袖型种子顾客对创新扩散的影响^[29],或者通过构建无标度社会网络来研究病毒营销活动,探讨在消费者网络中选择初始影响者的方法^[30]。研究发现,技术扩散的早期,网络中出现种子顾客或者意见领袖会加速创新在社会网络中的扩散^[31]。黄敏学等将意见领袖所依存的网络划分为社交网络和消费咨询网络,分析了网络节点的信息创造和网络结构对于网络关系构建以及意见领袖形成的影响^[32]。Peng等^[33]在研究微博信息的最终流行度时,发现信息传播早期转发者的拓扑结构和最终流行程度有很大相关性,若用户所发信息早期的传播深度较广,则有利于信息在社会网络中的广泛传播。而典型的兴趣社交网络微博的用户转发消息的概率受其邻居中已转发此消息邻居之间形成的连通图的个数的正面影响^[34]。在实物商品在某企业员工中的扩散模拟中发现,种子用户分布越分散,创新扩散获得市场成功所需要的时间就越短^[35]。

总之,目前微观层面的研究往往对网络拓扑特征的描述以及社会网络拓朴性质的应用研究,比如说初始条件对扩散影响的相关研究,更多关注初始种子用户影响力对创新扩散的影响,忽略了扩散过程中所选取的种子用户在社交网络的网络分布特征的差异、种子用户之间的连接关系差异,也可能引起的扩散效率的不同。另外,当前关于种子用户分布对创新扩散的研究方法多为数学模型演绎或用少量节点的网络,而在线社交网络是一种典型的复杂网络,节点数量繁多、结构关系复杂,若脱离现实网络数据的复杂性,将使得其研究成果的可靠性及适用性等方面缺乏实证检验^[36]。

为尝试回答上述问题,本文利用大规模真实社交网络数据,在微观层面上设计基于阈值模型和基于节点相关性的两个实验,在种子用户的不同网络分布及扩散规则下,采用计算机仿真模拟得到结论。

2 研究假设和主要概念界定

社交应用产品是指以交友为目的,以基于各种兴趣组成的交友社区为组织形式的移动应用。社交应用产品相比其他工具类、生活实用类产品来说,具有网络外部性,社交类产品有内生的扩张需求,用户有主动传播给好友的内生动机,从而增加自身的网络价值。针对此类产品从网络外部性角度考虑,内生的传播动机远远超过其他营销因素,可以一定程度上排除其他因素的干扰。由于社交应用的网络外部性特性,之前关于其他类型新产品的推广策略的研究,可能并不适用于社交产品。

本文专指在线社交网络,原因是当前的在线社交网络已经覆盖绝大多数人口,在网络上的抽样,可以充分表达总体特征。兴趣社交网络则是基于兴趣发生用户间连接的社交网络。基于兴趣的社交网络中,用户主动进行传播扩散也是一个相对比较自然的行为,可以在一定程度上排除其他营销因素的干扰。另外就本课题的研究内容而言,社交类应用天然地在上传播,因而线下零星的口碑传播可以不做考虑。

本文中的种子用户是指产品自传播之前,社交产品开发商通过各种方式在社交网络中获得的第一批用户,也称为初始感染节点集合(感染源)。社交产品扩散则指的是用户发现,下载安装社交应用类产品并向好

友传播. 社交网络中社交产品的扩散并不完全等同于用户采纳与传播, 然而由于其社交属性和应用的特殊性, 用户传播或转发相关信息大多数情况下都是在用户下载安装使用之后, 故研究忽略转发却不采纳的情况. 基于上述细微的不同, 研究将在不同场合使用“传播”、“采纳”、“感染”和“扩散”等术语. 为简化起见, 论文认为社交网络中社交类产品的扩散过程就是类似于病毒的传播过程, 即社交网络中的节点(即用户)存在已感染(已采纳的用户)和未感染(潜在采纳用户)两种状态.

根据文献[21, 34, 35]的分析, 均发现扩散效率与网络结构或种子顾客在社会网络中分布的关系. 现实中亦可观察到, 社交应用产品本身可以满足用户需求前提下, 同样数量的种子用户, 在社交网络中的分布越集聚, 社交产品扩散的速度越快. 因此提出假设

H1: 社交网络中, 种子用户的网络紧密度与社交产品扩散速度正相关.

进一步地, 若假设 **H1** 为真, 那么种子用户分布的集聚程度对创新扩散速度的正向影响又是怎么样? 文献[13]研究了初始采纳者比例(即数量)与创新扩散之间的正相关关系, 同时也在实际传播限项中观察到初始已采纳用户比例对于创新扩散速度的影响迅速变小, 即

H2: 社交网络中, 种子用户的网络紧密度对社交产品扩散速度的正向影响加速递减.

设社交网络为 $G = (V, E)$, 其中 V 为节点集合, E 为边的集合, 初始种群所组成的网络记为 $G_0(V_0, E_0)$, 那么有 $G_0(V_0, E_0) \subset G(V, E)$, 则 G_0 的紧密度为

$$c = \frac{2L}{N(N-1)}, \quad (1)$$

其中 L 代表网络中实际存在的边的数量, N 代表网络中种子用户节点的数量.

网络紧密度 c 取值范围在 0 和 1 之间, 作为衡量种子用户分布的集聚程度的指标, 可以表征种子用户的网络紧密度越大, 网络分布越集聚. 为表述方便, 令

$$c = c_{N-y-x}, \quad (2)$$

其中 x, y 为正整数, x 代表每个小团体中节点的个数, y 代表彼此不直接相连的小团体个数, 易知 $N = xy$ (见图 1 各子图).

在实验中考察社交产品扩散速度, 主要观察从种子用户群(时钟 0)至整个目标用户群的 20%(时钟 t)所需要的时长作为扩散快慢的标志. 假设 **H1** 可量化为, 随着种子用户紧密度的增大, 创新扩散至整个目标用户群的 20% 左右所用的时间(达到创新累积采纳百分比 20%, 即创新扩散“起飞”所用的时间)越短. 假设 **H2** 量化为, 种子用户的不同网络紧密度下, 社交产品扩散至整个目标用户群的 20% 所用的时间间隔越来越短, 呈加速递减趋势.

3 仿真实验

为了检验本文假设, 使用典型社交网络为基础数据, 选取不同数量的种子用户, 并在种子用户的不同分布下进行基于阈值模型的仿真实验和基于节点相关性的仿真模拟实验. 基于阈值模型的仿真实验考虑初始种子用户的网络分布以及扩散中信息的累积效应对社交产品扩散的影响, 重复实验, 进而从数据集中观察平均值、中位数、方差和偏度等指标, 从而得到高信度的结论. 基于节点相关性的扩散仿真实验与基于阈值模型的扩散仿真实验的主要区别在于传播规则需要更进一步考虑已感染节点与未感染节点之间的相关性, 进而有倾向地选择邻居节点进行感染.

3.1 社交网络数据获取与分析

本文选取豆瓣网作为仿真实验的样本社交网络. 原因是豆瓣网运营多年, 覆盖人群非常广泛并比较稳定, 可以排除高速增长期或衰弱期平台带来的影响; 使用 Python 语言编写网页抓取及解析程序, 递归地获取

社交网络数据. 具体步骤为, 首先在网络中随机抽取一个用户作为起始用户, 获取起始用户关注的对象, 进而迭代地对关注的对象进行同样的操作, 直到获得足够的数据(代码略). 然后利用 R 语言和程序包 igraph 构建一个无向无权的社交网络图(代码略)并分析.

抓取到的豆瓣社交子网络包含 206 891 个节点, 544 748 条边, 网络紧密度为 1.27×10^5 , 各节点的度值及其关注的节点的度值表现出显著的异质性, 且各用户被其他人关注的差异较大. 采用常规社会网络分析方法可得该网络的聚类系数为 0.075, 最大强连通图的平均最短路径为 3.18, 即具有一般社会网络的短平均最短路径长度和高内聚的小世界特性, 适合作为用于社交产品扩散仿真的社交网络^[18].

3.2 基于阈值的仿真实验

为简化工作, 社交网络中的节点(即用户)只存在两种状态——已感染节点(代表已采纳的用户)和未感染节点(代表潜在采纳用户), 种子用户作为感染源, 构成初始感染节点集合; 其余用户构成未感染节点集合, 感染节点对未感染节点具有持续的传染能力, 直至该节点被感染. 此外, 由于社交类产品的社交属性所蕴含的内生传播动机, 用户转发或下载安装试用的过程中均伴随着信息再次传播, 因此可以认为未感染节点被感染后会主动向其他未感染节点传播相关信息, 从而整个感染过程会主动进行下去, 直至网络大部分节点被感染.

实验采用线性阈值模型, 感染刺激是指未感染节点接收到一次已感染节点的感染信息. 经过初步试算, 刺激阈值的大小不会影响实验的结果. 因此设定如下感染规则: 每个未感染节点累积受到来自相同或者不同易感染节点的感染刺激次数 $\delta \geq 10$, 则该节点被感染. 为控制其他因素对创新扩散过程的影响, 随机选取种子用户, 并大量重复实验过程, 从而排除异质性的干扰. 算法流程如下:

步骤 1 初始化种子用户群数量 N , 每个子团体的用户数量 x , 种子用户群子团体数量 y , 迭代次数 recursions(default = 100);

步骤 2 加载网络, $G_0 \leftarrow \text{network.seed_users}(N, x, y)$ 生成种子用户群 G_0 ;

步骤 3 当 $r < \text{recursions}$ 时, 随机获取 $v \in G_0$;

步骤 4 获取 v 的邻居节点 $n \leftarrow \text{neighbors}(v)$;

步骤 5 对每个邻居节点进行感染, 即 $n.\text{infected} \leftarrow n.\text{infected} + 1$;

步骤 6 若节点 v 的受感染超过阈值, 即 $v.\text{infected} \geq \delta$, 则将 v 加入子图 $G_0 \leftarrow G_0 + v$;

步骤 7 重复步骤 3~步骤 6.

实验假定任意一个感染节点阈值都是相同的, 先考察 $N = 100$, (y, x) 元组分别为 $(100, 1)$, $(50, 2)$, $(25, 4)$, $(20, 5)$, $(10, 10)$, $(1, 100)$, 即 $c_{100-100-1}$, $c_{100-20-5}$, $c_{100-10-10}$, $c_{100-20-5}$, $c_{100-1-100}$. 为提高实验结果的信度, 考察初始种子用户数量规模为 $N = 10$ 和 $N = 1\ 000$ 下, 种子用户的网络紧密度对初期扩散速度的影响. 鉴于 $N = 100$ 实验结果中 N 是扩散发生变化的临界值, 当 $N = 10$ 时只需分别考察 $c_{10-10-1}$, c_{10-5-2} , $c_{10-1-10}$ 三种情况, 当 $N = 1\ 000$ 时只需分别考察 $c_{1\ 000-1\ 000-1}$, $c_{1\ 000-500-2}$, $c_{1\ 000-250-4}$, $c_{1\ 000-50-20}$, $c_{1\ 000-1-1\ 000}$ 五种情况. 其他创新扩散仿真实验规则步骤和参数均与 $N = 100$ 实验相同.

1) 种子用户的节点数量为 10 个时, 上述 3 种网络紧密度情况下, 基于阈值模型的创新扩散仿真实验的 20 次迭代, 100 次重复实验结果如图 2 所示.

由于创新累积采纳数量到达目标用户群的 20% 作为初期阶段扩散取得成功的标志, 因此实验结果重点分析初期阶段创新扩散情况. 由标准差曲线可以看出, 重复实验结果的总体上离差较小. 随着种子用户的网络分布紧密度的增大, 扩散至 20% 所用的时间越短, 即扩散速度越快, 初步验证假设 H1 为真(考虑到篇幅, 仅给出较能反映数据特征的中位数和标准差指标). 由图 2 中 100 次重复试验的平均值和中位数扩散曲线可以看出, 扩散曲线 $c_{10-10-1}$ 和 c_{10-5-2} 间距较大, 而 c_{10-5-2} 和 $c_{10-1-10}$ 间距突然变小, 即当种子用户的网络分布紧密度为 0(即 $c_{10-10-1}$)的基础上小幅度的提高, 扩散速度便有了极大的提高, 随着紧密度继续增

大, 其对扩散速度增长的影响变小, 初步验证了假设 H2 为真。

2) 种子用户数量为 100 时, 其不同网络分布紧密度下的 20 次迭代, 100 次重复试验的结果如图 3 所示。由图 3 可以看出, 初始种子用户数量为 100 时, 其不同网络分布紧密度下的 100 次重复实验结果的标准差曲线和偏度曲线总体上离差较小。从整体趋势上看, 节点数量为 100 的种子用户的不同网络紧密度下, 100 次重复试验的平均值和中位数扩散曲线陡峭程度大体一致。随着种子用户的网络分布紧密度的增大, 创新扩散至 20% 所用的时间越短; 从整个扩散过程来看, 种子用户的网络紧密度越大, 创新扩散到达峰值的时间也越早, 同样初步验证了假设 H1 为真。

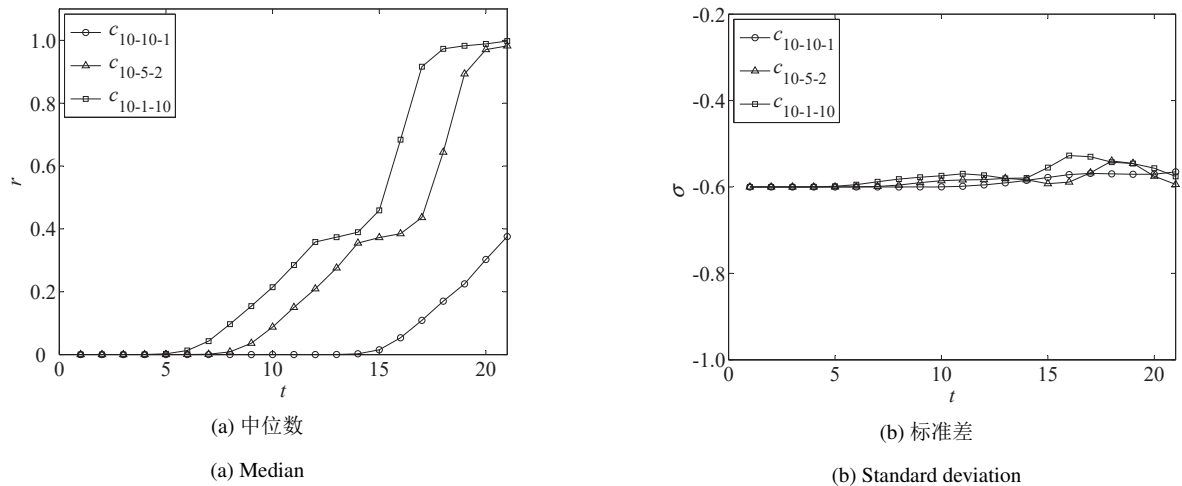


图2 种子用户的不同分布下基于阈值模型的仿真实验结果($N = 10$)

Fig. 2 The results of simulation experiments based on threshold model under different distribution of seed users ($N = 10$)

由图 3 中 100 次重复试验的中位数扩散曲线可以看出, 扩散曲线 $c_{100-100-1}$ 的启动时间远远滞后于介于 $c_{100-50-2}$ 和 $c_{100-1-100}$ 之间的扩散启动时间, 而介于 $c_{100-50-2}$ 和 $c_{100-1-100}$ 之间的扩散启动时间差异骤然变小(扩散曲线几乎重合)。可见, 当种子用户的网络分布紧密度在 0 的基础上小幅度提高, 其扩散速度便有了极大的提高, 而随着紧密度继续增大, 其对扩散速度的增长的影响加速变小, 初步验证了假设 H2 为真。

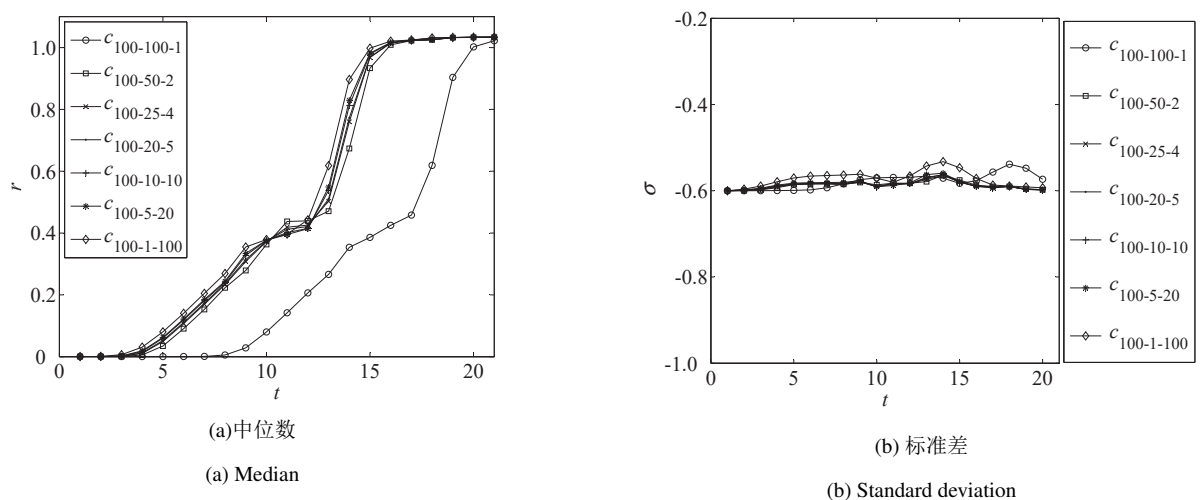


图3 种子用户的不同分布下基于阈值模型的仿真实验结果对比分析图($N = 100$)

Fig. 3 Comparison and analysis diagram of simulation experiment results based on threshold model under different distribution of seed users ($N = 100$)

3)种子用户的节点数量为1 000个时,其5种紧密度情况下的20次迭代,100次重复仿真实验结果对比分析图如图4所示。

由图4可以看出,种子用户节点数量为1 000时,其不同网络紧密度分布下,感染节点数量的增长极其迅速。从开始感染到覆盖整个网络,仅经历了15次左右的迭代过程。从扩散曲线的整体趋势上看,节点数量为1 000的种子用户网络紧密度越大,创新扩散至目标用户网络的20%所用的时间越短,或到达峰值的时间也越早。进一步验证了假设H1为真。同样可以看出,当种子用户的网络分布紧密度在0的基础上小幅度的提高,其扩散速度便有了较大的提高,当种子用户的网络分布紧密度继续增大时,其对扩散速度的增长的影响骤然变小,验证了假设H2为真。

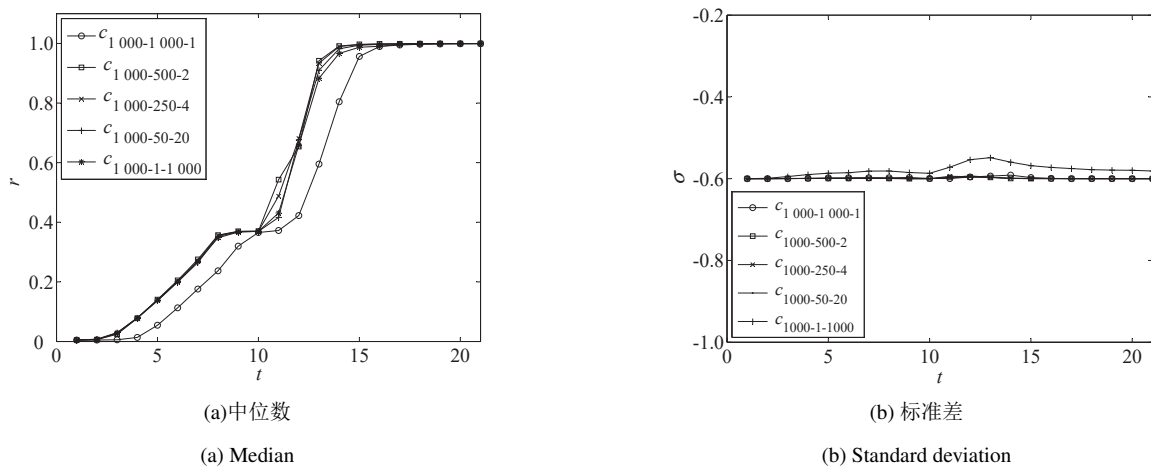


图4 种子用户的不同分布下基于阈值模型的仿真实验结果对比分析图($N = 1\,000$)

Fig. 4 Comparison and analysis diagram of simulation experiment results based on threshold model under different distribution of seed users ($N = 1\,000$)

3.3 基于阈值模型的仿真实验结果对比分析

图5为不同时刻下,节点数量为10的种子用户的创新累积采纳百分比随其网络紧密度的变化而变化的轨迹图。

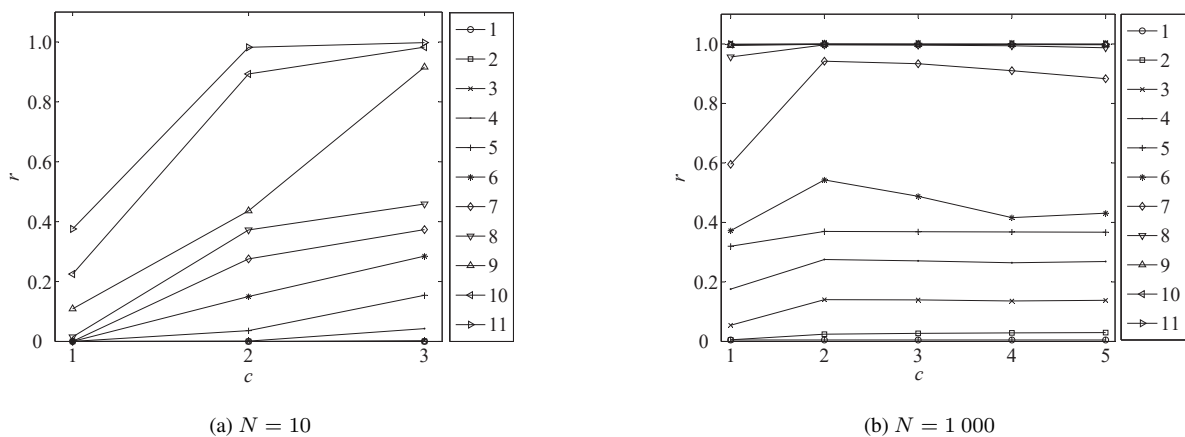


图5 不同时刻下,累积采纳百分比随种子用户紧密度变化而变化曲线($N = 10, 1\,000$)

Fig. 5 The change curve of cumulative adoption percentage with the change of seed user tightness at different times ($N = 10, N = 1\,000$)

可以看出, 初期扩散取得成功时(扩散至整个目标网络的 20% 用户), 即 $t = 18$ 时刻下, 轨迹曲线图在临界紧密度 c_{10-5-2} 前后变化存在显著差异. 在临界紧密度之前, 曲线非常陡峭, 在临界紧密度后, 曲线骤然变缓, 几乎呈水平线. 节点数量 $N = 100$ 实验中在 $t = 7$ 时刻下, 轨迹曲线图在临界紧密度 $c_{1\,000-500-2}$ 前后变化存在显著差异. 在节点数量 $N = 10$ 实验中, 类似情况出现在 $t = 13$ 时刻临界紧密度为 $c_{100-50-2}$ 时.

因此, 初始种群节点数量规模为 10, 100 和 1 000 实验的共同结论是当种子用户的网络紧密度在 0 的基础上有小幅提升时, 创新扩散的速度迅速得到提升. 然而, 紧密度提升到 1 的过程中, 种子用户的网络紧密度对创新扩散速度的正向影响呈加速式递减, 假设 H2 为真.

选取不同节点数量的种子用户(10, 100 和 1 000)进行仿真实验对比分析发现, 扩散的初始阶段, 扩散速度最慢的为 $c_{10-10-1}$, 扩散速度最快的为 $c_{1\,000-1-1\,000}$ (与 $c_{1\,000-500-2}$ 几乎重叠, 如图 6). 这一方面说明初始传播阶段, 由于初始种群的数量规模相差较大(100 倍), 种子用户群的数量越大, 扩散的速度越快, 这也比较符合现实情况.

另外一个有趣的现象是, $c_{10-1-10}$ 扩散至整个网络的 20% 所用时间比 $c_{100-100-1}$ 要短, 同样 $c_{100-1-100}$ 扩散至整个网络的 20% 所用时间比 $c_{1\,000-1\,000-1}$ 的时间短. 显然在一定节点数量范围内, 种子用户的较为集聚, 能够弥补其数量方面劣势, 种子用户的网络紧密度对扩散速度的正向影响更为显著.

3.4 基于节点相关性的仿真实验

基于阈值模型的仿真实验只考虑初始种子用户的网络分布以及扩散中信息的累积效应对社交产品扩散的影响, 没有考虑社交网络中用户之间的相关性, 即一个用户对另一个用户的创新采纳决策影响力. 在现实的社交网络中, 当社会网络结构确定时, 社交产品是通过节点(用户)间的相互连接来进行扩散的, 不同节点之间相互影响力是不同的, 而节点的影响力在一定程度上表现为节点 v_i 的度数 k_i 和与之相连的节点 v_j 各自的度 k_j 之间相关性, 相关性越强, 节点 v_i 对节点 v_j 的行为影响力越大. 因此, 在这种情况下, 潜在采用者个体(即未感染节点)形成的网络实际上是一个由节点之间的相关性形成的权重网络. 因此, 出于完备性的目的, 本文进一步设计了基于节点之间的相关性(影响力)的扩散仿真实验来考察初始种子用户的网络分布对社交产品扩散的影响.

基于节点相关性的仿真实验的种子用户选取根据上述网络紧密度模型构建规则以及网络分布模型图的具体构建思路, 以及基于阈值模型的仿真实验结果, 选取 $c_{100-100-1}$, $c_{100-50-2}$, 和 $c_{100-1-100}$ 各 3 种不同网络紧密度分布进行实验.

未感染节点在决策是否采纳某项新产品时, 会优先考虑它的邻居中那些已感染的意见领袖或者“明星节点”的建议. 但同时, 如果未感染节点的邻居非常少, 即使这些邻居只是普通节点(即度值比较小), 也会对未感染节点产生比较大的影响, 因为由于邻居数比较少增大了未感染节点看到已感染节点传播的信息的概率, 进而也会使增大该节点被感染的概率.

设节点 i 的度数为 k_i , 节点 j 的度数为 k_j , i, j 直接相连, 则节点 j 对 i 的相关性为

$$w_{j,i} = \frac{k_j}{k_j + k_i}. \quad (3)$$

基于节点相关性的扩散仿真实验的感染规则: 首先计算所有已感染节点和它的邻居未感染节点之间的相关性, 将该相关值作为刺激值, 根据其数值从大到小排列, 未感染节点受到相同或者不同节点的感染刺激阈值 $\delta \geq 10$.

仿真实验结果分析 100 次重复实验结果如图 6 所示, 种子用户的网络紧密度越大, 创新扩散曲线越陡峭; 无论是平均值还是中位数的扩散曲线都可以看出, $c_{100-50-2}$ 和 $c_{100-1-100}$ 扩散至整个网络的 20%(约 40 000 个节点)所用的时间均远小于 $c_{100-100-1}$ 时所用的时间, 即种子用户的网络紧密度越大, 创新扩散的速度越快. 其中 $c_{100-50-2}$ 和 $c_{100-1-100}$ 下速度虽然相差不大, 但是 $c_{100-50-2}$ 的扩散速度要大于 $c_{100-1-100}$ 的扩散速度, 说明基于节点之间相关性的扩散规则下, 种子用户的网络分布集聚状态下的扩散

速度均比分散状态下的要好,但是超过一定临界值后敏感性减弱,因此,在实际的应用中,需要根据节点的数量选择适当紧密度的种子用户进行社交产品的扩散.由此验证假设 H1 为真.

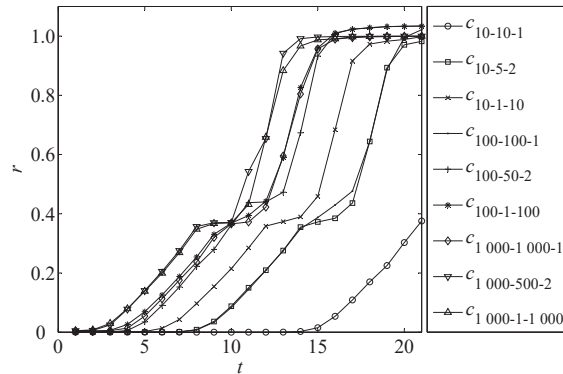
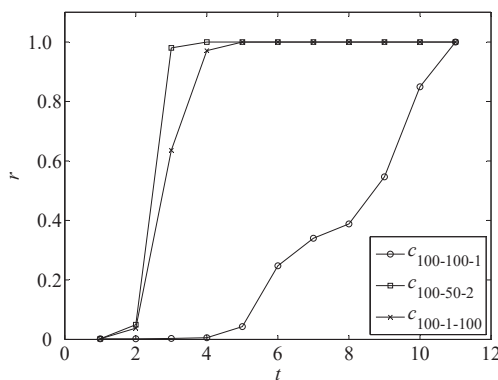


图 6 种子用户的不同分布下基于阈值模型的仿真实验结果($N = 10, 100, 1000$)

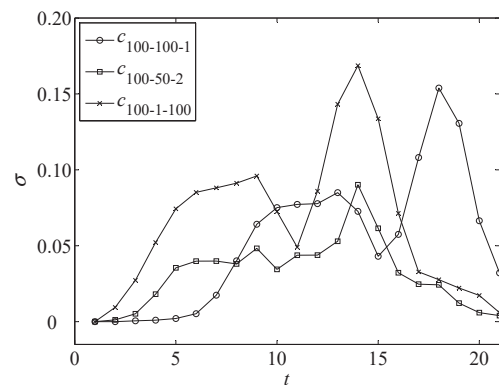
Fig. 6 Threshold model based simulation result on different seed user distribution ($N = 10, 100, 1000$)

$N = 100$ 时的种子用户的不同分布下基于节点相关性的仿真实验结果如图 7 所示. 由图 7 的不同时刻的平均值和中位数扩散曲线可以看出, 初始种群的节点数量为 100 时, 随着其紧密度的增大, 社交产品的扩散速度的增大幅度先增大大过了临界值后减慢, 即初始种群紧密度对于扩散速度正向影响是加速递减的, 验证了假设 H2 为真.



(a) 中位数

(a) Median



(b) 标准差

(b) Standard deviation

图 7 种子用户的不同分布下基于节点相关性的仿真实验结果($N = 100$)

Fig. 7 Node correlation based simulation result on different seed user distribution ($N = 100$)

图 8 为不同时刻下, 节点数量为 100 的种子用户群的创新累积采纳百分比随其网络紧密度的变化轨迹. 可以看出在 $t = 13$ 时刻下, 随着种子用户网络紧密度的增大, 创新累积采纳百分比先加速爬升, 然后骤然变缓, 临界紧密度均为 $c_{100-50-2}$, 当种子用户的网络紧密度由 $c_{100-100-1}$ 增大到 $c_{100-50-2}$ 时, 创新累积采纳百分比有大幅度提升, 而当种子用户的网络紧密度由 $c_{100-50-2}$ 继续增大到 $c_{100-1-100}$ 的过程中, 创新累积采纳百分比增大幅度差异极小. 可见, 种子用户节点数量规模为 100 情况下, 种子用户网络分布紧密度对创新扩散速度的提升作用并非都是显著的, 当种子用户的网络紧密度在 0 的基础上有小幅提升时, 创新扩散的速度迅速得到提升, 然而, 当种子用户的网络紧密度进一步提升到 1 的过程中, 种子用户紧密度对创新扩散速度的影响几乎很小, 种子用户的网络紧密度对创新扩散速度的正向影响加速递减, 进一步验证了假设 H2 为真.

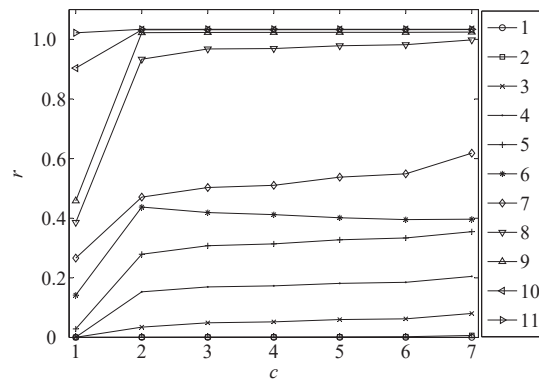


图8 基于阈值实验,不同时刻下,累积采纳百分比随种子用户紧密度变化曲线($N = 100$)

Fig. 8 Based on the threshold experiment, the change curve of cumulative adoption percentage with seed user tightness at different time points ($N = 100$)

图9为基于节点相关度实验,也有同样结论.与节点数量为100的基于阈值实验相比,不同在于时刻 $t = 9$,临界紧密度 $c_{100-5-2}$ 前后出现显著差异.显然,找感兴趣的人进行传播,效率更高,这也与人们生活实践中的思维习惯一致.

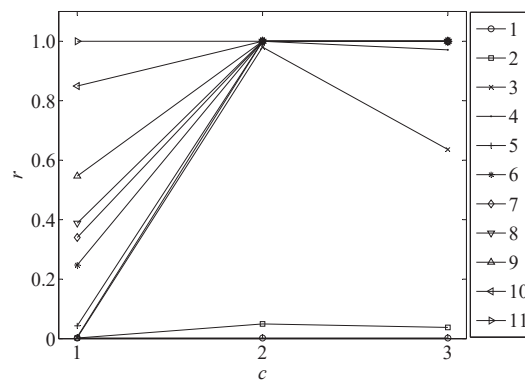


图9 基于节点相关度实验,不同时刻下,创新累积采纳百分比随种子用户紧密度变化的曲线($N = 100$)

Fig. 9 Based on the node correlation experiment, the curve of cumulative adoption percentage of innovation as a function of seed user tightness at different time points ($N = 100$)

4 结束语

本文基于大规模兴趣社交网络,选取不同数量的种子用户,利用基于不同的扩散规则结合已有的步进累积研究方法,验证了种子用户紧密程度与扩散效率的正相关关系.论文的贡献在于从微观角度出发,研究扩散过程的影响因素,同时利用大规模真实社会网络数据进行仿真,使研究在实际社交网络中得到高信度的验证.论文同时还发现了紧密度的贡献存在瓶颈,其现实解释为,紧密度是推动初期扩散的主要因素,但到了一定阈值之后作用被其他潜在因素替代,这也是未来需要进一步探讨的问题.

在营销实践中,选取的种子用户符合集聚网络的分布特征时,反而会比分散“撒网”策略更有利于新产品的扩散.论文的结论有助于企业在有限的营销预算下,选取适合种子用户网络分布策略,有效地投放营销资源,提高产品传播能力和采纳率,带来更大经济和社会效益.此外,鉴于种子用户的网络紧密度超过一定的临界值后,其紧密度对创新扩散速度的提升幅度呈加速减小趋势,因此需要根据企业的实际情况以及预期扩散效果和扩散数量,选择合适集聚程度的种子用户分布进行新产品的推广.

参考文献:

- [1] Li Q, Tao Z, Lv L Y, et al. Identifying influential spreaders by weighted leader rank. *Physica A: Statistical Mechanics and Its Applications*, 2014, 404: 47-55.

- [2] Rogers M E. Diffusion of Innovations. 5th Edition. New York: Free Press, 2003.
- [3] Bass F M. A new product growth model for consumer durables. *Management Science*, 1969, 15(5): 215–227.
- [4] 杜宾. 基于社会学习的创新扩散模型构建与实证. *系统工程理论与实践*, 2014, 34(10): 2619–2627.
Du B. Innovation diffusion modeling based on social learning. *Systems Engineering: Theory & Practice*, 2014, 34(10): 2619–2627. (in Chinese)
- [5] 庄新田, 黄玮强. 基于消费者网络的金融创新扩散研究. *管理科学学报*, 2009, 12(3): 132–141.
Zhuang X T, Huang W Q. Study on financial innovation diffusion based on consumer network. *Journal of Management Sciences in China*, 2009, 12(3): 132–141. (in Chinese)
- [6] Kapur P K, Singh V B, Anand S, et al. An innovation diffusion model incorporating change in the adoption rate. *Management Dynamics*, 2007, 16(1): 34–41.
- [7] Tucker C. Identifying formal and informal influence in technology adoption with network externalities. *Management Science*, 2008, 54(12): 2024–2038.
- [8] 何铮, 张晓军, 吴易明. 新产品扩散随机阈值模型的实证研究. *系统管理学报*, 2013, 22(1): 39–45.
He Z, Zhang X J, Wu Y M. An empirical study of stochastic threshold model on new product diffusion. *Journal of System & Management*, 2013, 22(1): 39–45. (in Chinese)
- [9] Tong H. Threshold models in time series analysis some reflections. *Journal of Econometrics*, 2015, 189(2): 485–491.
- [10] Amini M, Wakolbinger T, Racer M. Alternative supply chain production-sales policies for new product diffusion: An agent-based modeling and simulation approach. *European Journal of Operational Research*, 2012, 216(2): 301–311.
- [11] 周琦萍, 徐迪, 杨芳. 基于 Agent 的创新扩散机理研究. *软科学*, 2014, 28(1): 6–10.
Zhou Q P, Xu D, Yang F. Study of agent-based mechanism of innovation diffusion. *Soft Science*, 2014, 28(1): 6–10. (in Chinese)
- [12] 李英, 胡剑. 基于智能体的多类新能源汽车市场扩散模型. *系统管理学报*, 2014, 23(5): 711–716.
Li Y, Hu J. An agent-based diffusion modeling and simulation of multiple classes of new energy vehicles. *Journal of System & Management*, 2014, 23(5): 711–716. (in Chinese)
- [13] 黄玮强, 姚爽, 庄新田. 基于复杂社会网络的创新扩散多智能体仿真研究. *科学学研究*, 2013, 31(2): 310–320.
Huang W Q, Yao S, Zhuang X T. Research on multi-agent in innovation diffusion based on complex social network. *Studies in Science of Science*, 2013, 31(2): 310–320. (in Chinese)
- [14] Granovetter M. Threshold models of collective behavior. *American Journal of Sociology*, 1978, 83(6): 14–20.
- [15] 张晓军, 李仕明, 何铮. 社会关系网络密度对创新扩散的影响. *系统工程*, 2009, 27(1): 92–97.
Zhang X J, Li S M, He Z. Impact of social network density on innovation diffusion. *System Engineering*, 2009, 27(1): 92–97. (in Chinese)
- [16] 郭静, 曹亚男, 周川, 等. 基于线性阈值模型的影响力传播权重学习. *电子与信息学报*, 2014, 36(8): 1804–1809.
Guo J, Cao Y N, Zhou C, et al. Influence weights learning under linear threshold model in social network. *Journal of Electronics and Information Technology*, 2014, 36(8): 1804–1809. (in Chinese)
- [17] 赵良杰, 武邦涛, 段文奇, 等. 消费者交互作用对网络效应产品扩散的影响: 基于产品生命周期的视角. *系统工程理论与实践*, 2012, 32(1): 67–75.
Zhao L J, Wu B T, Duan W Q, et al. The influence of customer interactions on the diffusion of products with network effects: A view based on product life cycle. *System Engineering: Theory & Practice*, 2012, 32(1): 67–75. (in Chinese)
- [18] Watts D J, Strogatz S H. Collective dynamics of small-world networks. *Nature*, 1998, 393(6684): 440–442.
- [19] 郭海霞. 基于小世界网络的微博信息影响力研究. *情报科学*, 2014, 32(9): 60–63.
Guo H X. Influential of weibo based on small world. *Science of Intelligence*, 2014, 32(9): 60–63. (in Chinese)
- [20] 赵健宇, 袭希, 苏屹. 知识流动网络演化中小世界效应的仿真研究. *管理评论*, 2015, 27(5): 70–81.
Zhao J Y, Xi X, Su Y. A simulation research on small world effect in knowledge flow's network evolution. *Management Review*, 2015, 27(5): 70–81. (in Chinese)
- [21] Centola D. The spread of behavior in an online social network experiment. *Science*, 2010, 329(5996): 1194–1197.
- [22] Delre S A, Jager W, Janssen M A. Diffusion dynamics in small-world networks with heterogeneous consumers. *Computational and Mathematical Organization Theory*, 2007(13): 185–202.
- [23] 赵保国, 余宙婷. 基于网络效应的竞争性产品微观扩散研究. *管理科学学报*, 2013, 16(9): 33–43.
Zhao B G, Yu Z T. Competitive product diffusion at the level of individuals based on network effect. *Journal of Management Sciences in China*, 2013, 16(9): 33–43. (in Chinese)
- [24] 赵正龙, 陈忠, 孙武军, 等. 具有差异化选择特征的复杂社会网络扩散研究. *管理科学学报*, 2010, 13(3): 38–49.
Zhao Z L, Chen Z, Sun W J, et al. Diffusion with property of differential choice in complex social network. *Journal of Management Sciences in China*, 2010, 13(3): 38–49. (in Chinese)

- [25] Peres R, Muller E, Mahajan V. Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, 2010, 27(2): 91–106.
- [26] 杜 宾. 基于多维学习的服务扩散模型构建与实证. *系统工程理论与实践*, 2014, 34(S1): 210–215.
Du B. Modeling and empirical study of service diffusion based on multidimensional learning. *System Engineering: Theory & Practice*, 2014, 34(S1): 210–215.
- [27] 孙耀吾, 卫英平. 高技术企业联盟知识扩散研究: 基于小世界网络的视角. *管理科学学报*, 2011, 14(12): 17–26.
Sun Y W, Wei Y P. Study on knowledge diffusion of high tech enterprise alliance from the small-world network perspective. *Journal of Management Sciences in China*, 2011, 14(12): 17–26. (in Chinese)
- [28] Choi H, Kim S, Lee J. Role of network structure and network effects in diffusion of innovations. *Industrial Marketing Management*, 2010, 39(1): 170–177.
- [29] Eck P S V, Jager W, Leeftang P S H. Opinion leaders' role in innovation diffusion: A simulation study. *Journal of Product Innovation Management*, 2011, 28(2): 187–203.
- [30] Kiss C, Bichler M. Identification of influencers-measuring influence in customer networks. *Decision Support Systems*, 2008, 46(1): 233–253.
- [31] Valente T W, Davis R L. Accelerating the diffusion of innovations using opinion leaders. *The Annals of the American Academy of the Political and Social Sciences*, 2000, 566(1): 55–67.
- [32] 黄敏学, 王琦缘, 肖邦明, 等. 消费咨询网络中意见领袖的演化机制研究: 预期线索与网络结构. *管理世界*, 2015(7): 109–121.
Huang M X, Wang Q Y, Xiao B M, et al. A study on the evolutionary mechanism of the idea leader in the consumer advice network of expectancy clue and the network structure. *Management World*, 2015(7): 109–12. (in Chinese)
- [33] Peng T Q, Zhang L. Breadth, depth and speed: Diffusion of advertising messages on microblogging sites. *Internet Research*, 2015, 25(3): 453–464.
- [34] Zhang J, Tang J, Li J, et al. Who influenced you: Predicting retweet via social influence locality. *ACM Transactions on Knowledge Discovery from Data*, 2015, 9(3): 1–26.
- [35] 陈 锬. 种子顾客的网络分布对创新扩散的影响. *管理科学*, 2010, 23(1): 38–43.
Chen K. The impact of seed customers' network distribution on innovation diffusion. *Journal of Management Science*, 2010, 23(1): 38–43. (in Chinese)
- [36] Lin M, Li N. Scale-free network provides an optimal pattern for knowledge transfer. *Physica A: Statistical Mechanics and Its Applications*, 2010, 389(3): 473–480.

作者简介:

汤 胤(1975—), 男, 福建周宁人, 博士, 教授, 研究方向: 大数据与商务智能, Email: ytang@jnu.edu.cn;

徐永欢(1990—), 女, 河南驻马店人, 硕士生, 研究方向: 电子商务, Email: xuyonghuan08@163.com;

陈 剑(1981—), 女, 湖南永州人, 硕士, 副教授, 研究方向: 电子商务, Email: cici919@126.com.