

文章编号: 1003-0077(2016)02-0074-09

## 利用社交网络的影响力骨架探索信息传播

黄俊铭, 沈华伟, 程学旗

(中国科学院 计算技术研究所, 中国科学院网络数据科学与技术重点实验室, 北京 100190)

**摘要:** 理解社交网络上的信息传播机制, 通常包括对拓扑结构的分析和对用户行为的分析。由于社交网络上连边的强度具有异质性, 只有一部分连边对于信息传播有实质作用, 构成隐藏在社交网络中的影响力骨架。对影响力骨架的拓扑研究可帮助我们获得比直接研究社交网络拓扑结构更深刻的认识。我们从连边正负性和个体节点角色分化入手, 探讨了微观层面连边和节点在信息传播中的作用, 进而从宏观层面分析信息传播所依赖的影响力骨架的连通性和扩散效率, 发现信息传播具有一定程度的脆弱性, 且其传播效率低于对社交网络本身研究的预期。

**关键词:** 信息传播; 社交网络; 影响力骨架

**中图分类号:** TP391      **文献标识码:** A

## Understanding Information Propagations via Influence Backbone Analysis on Social Networks

HUANG Junming, SHEN Huawei, CHENG Xueqi

(CAS Key Lab of Network Data Science and Technology, Institute of Computing Technology,  
Chinese Academy Sciences, Beijing 100190, China)

**Abstract:** Understanding intrinsic mechanism of information propagations on social networks has attracted growing attention, including social network topology analysis and user behavior analysis. Due to the heterogeneity of links in social networks, only a portion of links significantly contribute to information propagations. The influence backbone of a social network, consisting of those links, might provide deeper insight to information propagations. Focused on the influence backbone, we analyze the signs of links with social structural balance theory, and the roles of nodes with heterogeneous distributions of out-degrees, so as to find the roles played by links and nodes in information propagations in a microscopic. Furthermore, we investigate the network connectivity and information spread efficiency of the influence backbone, finding that information propagations are more fragile and less effective.

**Key words:** information propagations; social networks; influence backbone

### 1 引言

近年来社交网络的飞速发展, 海量的信息通过社交网络分发、共享和传播, 对大规模人群的线上和线下生活产生不可忽视的影响。例如, 人们在 Twitter 和新浪微博上传播身边的新闻, 或是在 Facebook 和豆瓣上交换对于音乐的看法。随着社交网络上的信息传播逐渐衍生为蕴含丰富价值的大数据来源, 利用信息传播数据预测政治<sup>[1]</sup>、分析金融<sup>[2]</sup>、指导市场营销<sup>[3]</sup>、预警疾病传播<sup>[4]</sup>等数据挖掘

应用逐渐兴起, 使得理解信息传播的机制成为具有兼具研究价值和现实意义的重要问题<sup>[5]</sup>。

对信息传播的分析通常包括基于拓扑结构的分析和基于行为历史的分析。一些研究者分析社交网络的拓扑结构对于信息传播的速度、范围、广度等指标的影响, 希望寻找对信息传播具有指导性意义的网络拓扑机制, 例如, 探讨复杂网络的小世界性质对信息传播速度的影响<sup>[6]</sup>, 利用信息传播随机跳转的平稳分布寻找关键节点<sup>[7]</sup>, 讨论结构洞节点对于信息传播的影响<sup>[8]</sup>, 揭示局部拓扑结构的多样性对于个体信息传播的决定性作用<sup>[9]</sup>, 分析信息传播早期

收稿日期: 2013-09-15    定稿日期: 2014-04-15

基金项目: 国家基础研究发展计划(973)(2012CB316303, 2013CB329602); 国家自然科学基金(61232010, 61202215); 北京市自然科学基金(4122077)

人群局部拓扑性质<sup>[10]</sup>等。另一些研究者对社交网络中信息传播参与者的行为历史进行大规模分析以寻找信息传播的统计规律,例如,通过信息传播历史估计信息传播概率<sup>[11-14]</sup>,讨论用户评价对于信息传播的影响<sup>[15]</sup>,分析不同性别、年龄等人口学因素对信息传播的差异化作用<sup>[16]</sup>,拟合新信息随时间衰减的趋势<sup>[17]</sup>,评估信息传播中蕴含的信任程度<sup>[18]</sup>等。对拓扑结构和行为历史的分析,从不同侧面丰富了对信息传播机制的定性认识,能够更准确地指导针对信息传播的定量应用。

然而,作为两种主流分析方法之一的拓扑结构分析,目前几乎所有的工作都集中于分析社交网络本身的拓扑结构。事实上,社交网络中连接人与人的连边并不是全同的,而是具有非常显著的异质性,不同的连边有不同程度的信任和影响力<sup>[19]</sup>。事实上,只有一小部分边表达了显著的信任和影响(称为影响力边),能够在实质上推动信息传播。我们相信,由影响力边所构成的社交网络的子图(称为影响力骨架)的拓扑结构提供了更精确了解信息传播机制的拓扑途径。相比于对社交网络所有连边不加区分地合并分析,集中讨论影响力骨架或许能够提供更深入的认识。因此,本文我们试图从真实的社交网络中提取影响力骨架,以之作为实证分析信息传播的拓扑结构的载体。

我们首先从微观角度入手,利用社会平衡理论与度分化现象分析连边和节点在信息传播中的角色,探讨了影响力边的正负性和个体节点的主导者/追随者两种角色的分化现象。在微观讨论的基础上,我们从宏观层面分析影响力骨架的连通能力与扩散效率,发现信息传播对于关键节点的依赖程度与信息传播的扩散效率都与此前对社交网络研究的结果存在显著差异,被低估的脆弱性和被高估的传播效率使得我们意识到需要重新审视基于信息传播的各种量化应用。

## 2 提取影响力骨架

### 2.1 数据准备

我们采集了结合在线社交网络和在线分享的网站豆瓣网(www.Douban.com)和 Goodreads(www.Goodreads.com)的数据作为观测对象。在这两个网站上,用户可以对电影、音乐和书籍评分,并关注其他用户的行为。用户之间的关注关系构成

了社交网络,而对电影、音乐和书籍(统称为对象)的评分意见作为信息在这一社交网络上传播(数据抓取于2010年8月)。我们从网站上若干个热门讨论组开始,抓取讨论组成员的行为数据,并扩展到关注这些成员以及被他们关注的用户,迭代直到不再能扩展到新的未抓取用户。我们一共抓取了1 014 890位豆瓣用户和912 973位 Goodreads 用户,并移除了关注者少于五个或评分少于五次的不活跃用户,因为这些用户的可用数据太少,无法提供稳定的统计结果。清洗整理后的数据如表1所示。

表1 采集数据说明

	豆瓣	Goodreads
用户数	552 827	308 789
关注关系数	22 541 688	4 999 871
对象数	1 089 914	1 648 546
评分数	111 001 250	26 900 929

### 2.2 影响力骨架提取算法

为了提取影响力骨架,我们采用 Huang 等人提出的基于后验体验的影响力算法<sup>[20]</sup>来量化计算每条边上的影响力。具体来说,在用户  $v$  关注用户  $u$  所构成的有向边  $(u, v)$  上,用户  $u$  对用户  $v$  的影响力  $f_{u,v}$  定义为式(1)。

$$f_{u,v} = \frac{\sum_{i \in I_{u,v}} r_{v,i} - \hat{r}_{v,i}}{|I_{u,v}|} \quad (1)$$

其中  $I_{u,v}$  表示用户  $u$  向用户  $v$  推荐的对象集合,  $r_{v,i}$  表示用户  $v$  在接收到推荐后对于对象  $i$  的评分,  $\hat{r}_{v,i}$  是利用矩阵分解算法<sup>[21]</sup>估算出假如用户  $v$  未收到推荐的情况下对于对象  $i$  的评分。为避免小样本导致的不稳定统计,只考虑  $|I_{u,v}| \geq 5$  的边。将一条  $f_{u,v} \geq 0.1$  的有向边  $(u, v)$  记为影响力边,表明这条边上存在显著的影响力可以导致信息传播。在这条影响力边上,称用户  $u$  为影响者,称用户  $v$  为被影响者。注意影响者和被影响者都定义在单条边上而非全局,一条影响力边的影响者可能是另一条影响力边的被影响者(事实上这种情况非常常见,构成了信息传播的网络基础)。在两个数据集中分别有 0.095% 和 2.19% 的有向边被标记为影响力边,分别覆盖了 0.078% 和 7.23% 的用户。由这些影响力边所构成的子图标识了原社交网络中可以明确传播信息的骨架结构,称为影响力骨架,如表2所示。

表2 影响力骨架的统计

	豆瓣 (原社交网络)	豆瓣 (影响力骨架)	Goodreads (原社交网络)	Goodreads (影响力骨架)
用户数	552 827	18 243	308 789	146 494
连边数	22 541 688	21 941	4 999 871	217 494
连边密度	1.4e-4	1.3e-4	1.0e-4	2.0e-5
互惠系数	0.686	0.006	0.980	0.109

### 3 利用影响力骨架分析信息传播

#### 3.1 影响力边的正负性

我们首先从符号网络的角度分析影响力边的正负性。符号网络理论将网络中的连边分为表达积极关系的正边和表达消极关系的负边,许多复杂网络中都存在正负边对立的现象。理解影响力边的正负性对于预测信息传播有重要的理论意义。若影响力是正关系,则信息传播的效果是让群体意见趋向一致。若影响力是负关系,则信息传播的效果是让群体分裂成多个意见集团。

我们计算了影响力骨架的连边互惠系数,较高的互惠系数是正边的标识之一。虽然两个数据集上的影响力骨架的互惠系数相比于原网络有大幅度降低,但仍然高于随机抽取相同大小的随机子图的互惠系数,如表2所示。进一步地,我们利用社会平衡理论<sup>[22]</sup>检验影响力骨架中的三角形统计。社会平衡理论认为,如果三个个体之间的三条连边全为正关系,则这个三角形处于平衡态,可以稳定存在;反之若全为负关系则难以稳定存在。在影响力骨架中,任意三个用户之间的影响力若能构成三角形,无外乎传递( $a \rightarrow b, b \rightarrow c, a \rightarrow c$ )和循环( $a \rightarrow b, b \rightarrow c, c \rightarrow a$ )两种情况(若存在双向边,视为构成两个不同的三角形)。我们将影响力边全部随机重连,计算100次随机实验中这两种三角形在随机状态下的期望个数,与实际影响力骨架中两种三角形的个数相比。若一种三角形的实际个数远大于期望个数,表明这种三角形能稳定存在,反之则表明不能稳定存在。

如表3所示,两种三角的实际个数都远远大于期望个数。以z-score度量的显著水平表明这一现象并非由于随机误差导致,而是由于影响力骨架中的三角形处于平衡态,从而能大量稳定存在。这表明影响力边表征了用户之间的正关系,信息传播的

结果是使得相邻个体的态度趋向一致。

表3 社会平衡理论检验影响力边的正负性

	三角形(传递)	三角形(循环)
若全为正关系的平衡状态	平衡	平衡
若全为负关系的平衡状态	不平衡	不平衡
实际个数(豆瓣)	1 180	25
随机重连期望个数(豆瓣)	510	1.2
z-score 显著水平(豆瓣)	10	19
实际个数(Goodreads)	33 454	4 194
随机重连期望个数(Goodreads)	256	26
z-score 显著水平(Goodreads)	2 283	760

#### 3.2 个体角色分化

信息传播中的个体作为社交网络上的节点,度中心度对于理解个体在信息传播中的地位有较好的指示作用。在原社交网络中,由于互惠系数较高,大部分节点的出度入度几乎相等。而在互惠系数大幅降低的影响力骨架中,节点的出度和入度产生了明显的分化,这种分化表明了个体在信息传播中处于不同的地位。

图1为每个节点的出度-入度分布图。在豆瓣和Goodreads两个社交网络中可以看到一个明显的趋势,大部分样本点集中在 $y=x$ 直线附近,其出度与入度表现出强烈的相关性,表明大部分节点倾向于保持相等的出度和入度。在影响力骨架中几乎不存在这一趋势,很多样本点分布在远离 $y=x$ 直线的位置,表明影响力骨架中相当一部分节点的出度入度是严重不平衡的。为了更清晰地展示这一趋势,我们对每个节点计算其出度入度之比的对数。如图2所示,在两个社交网络中大部分节点集中在对数为0的位置,表明其出度入度基本一致。而在两个影响力骨架中,虽然仍有不少节点集中在对数为0的位置,但两侧有明显的多个分布峰,表明有相当一部分节点的出度远大于入度,亦有相当一部分节点的入度远大于出度。这一普遍的不平衡性现象背后是节点角色的分化。出度远大于入度的节点,对应的用户能够对一定数量的人群产生实质影响,同时自己受到他人的影响较小,在社交网络上的信息传播中常常扮演主导者的角色。与之对称的,入度远大于出度的节点,对应的用户会受到较多人的影响,而自己影响他人的能力有限,在信息传播中常

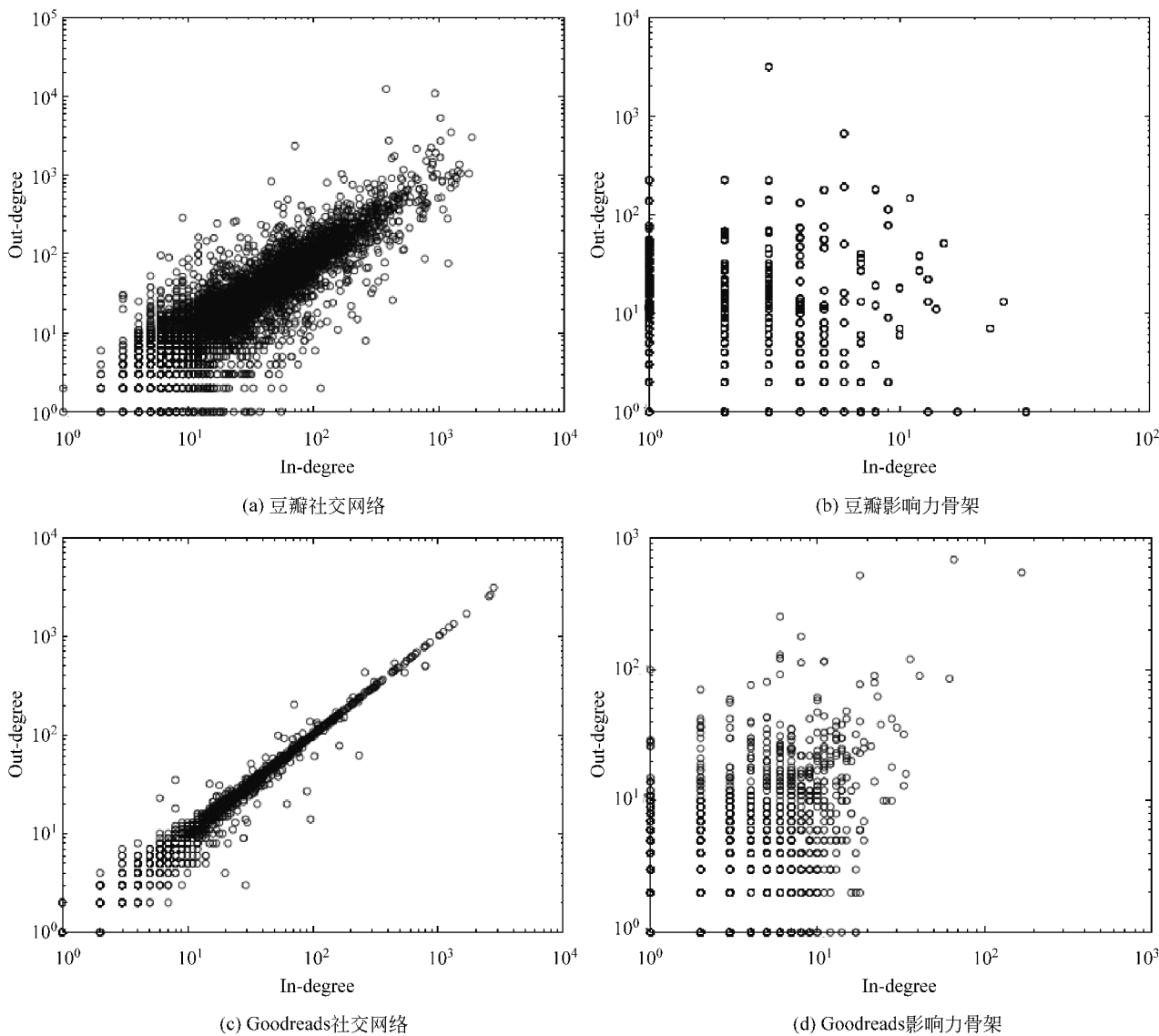


图 1 原社交网络和影响力骨架的出度—入度分布图

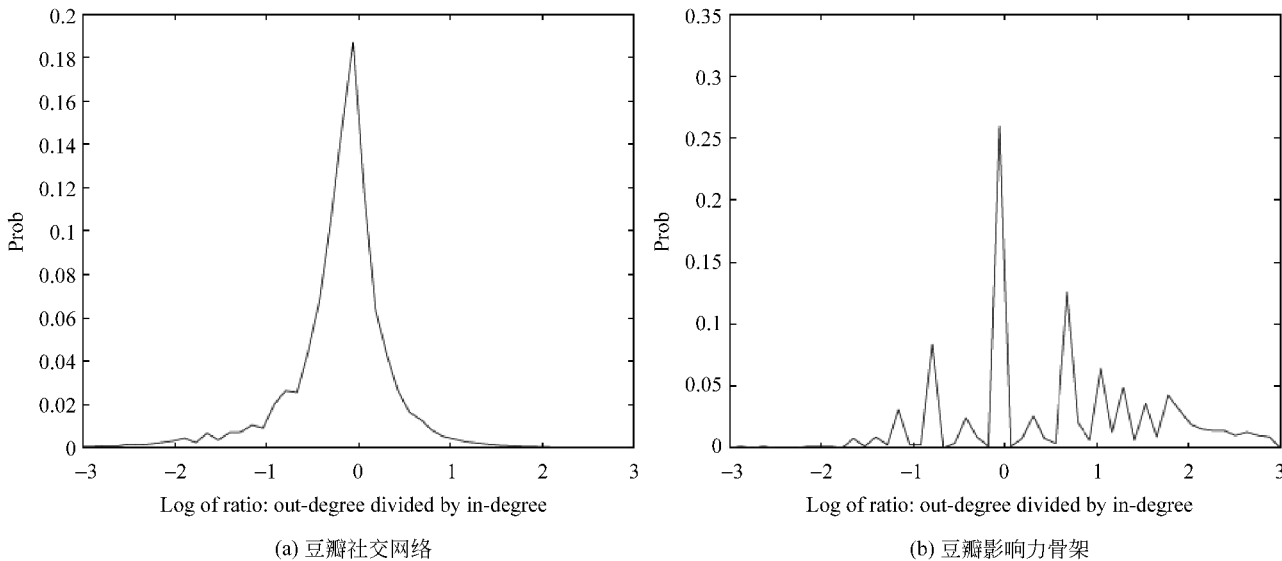


图 2 节点出度入度之比的对数分布

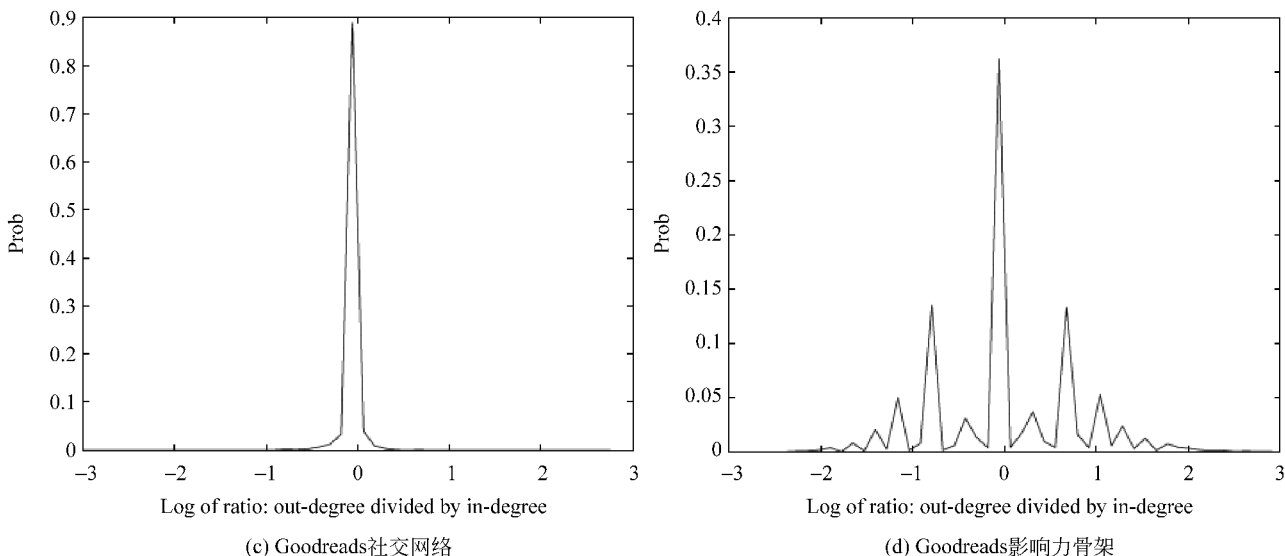


图 2(续)

常扮演追随者的角色。这两类人群的分化,使得信息能够有效地从主导者传播到追随者,构成了信息扩散的微观机制。

为了进一步分析这种不平衡性,我们选择出度为某一特定值的节点计算其平均入度,画出平均入度关于出度分布的散点图,以分析给定影响力能力的人群受到影响的程度。毫无意外地,两个社交网络的分布图集中在  $y=x$  一线,即平均入度严格随出度递增,不再赘述。图 3 左侧两图给出了两个影响力骨架上的分布情况。可以看到,随着出度的增加,平均入度并没有表现出相应地增加,而是始终稳定在一定水平线以下。豆瓣数据中,任何出度的节点的平均入度都在 10 以下。Goodreads 数据中,任何出度的节点的平均入度稳定在 10 左右(虽然较大出度的节点由于样本数量少产生波动)。这表明,对具有任何程度影响力的人来说,能影响他的人数始终是有限的。对于影响力较大的人,虽然其关注的人相应地较多(原社交网络中入度随出度递增),但是他所关注的人中能够实质影响他的人并不多。我们认为这是因为个体的信息处理能力是有限的,从而只能维持一个较小规模的信任对象群体,而不可能呈现无上界的线性增长。

相应地,我们也选择入度为某一特定值的节点计算其平均出度,画出平均出度关于入度分布的散点图,以分析给定受影响倾向的人群所具有的影响力。图 3 右侧两图给出了两个影响力骨架上的分布情况。平均出度关于入度的分布在两个数据集上表现出不同的趋势。豆瓣表现出较弱的负相关性,入

度越大的节点的平均出度越低,表明容易受影响的用户的影響力较弱。与之相反,Goodreads 表现出一定程度的正相关性,入度越大的节点的平均出度越大,表明 Goodreads 社区中有较多的枢纽节点。这些枢纽节点具有较强的信息转发能力,既能受到多个来源的影响,又能利用自己的高影响力将信息扩散出去。

### 3.3 传播拓扑

在个体角色的微观分析基础上,我们进一步考察网络的拓扑连通性,以便在宏观层面讨论信息传播。我们在社交网络和影响力骨架上分别随机选择一个节点,将其五阶以内的邻居节点构成的子图作为局部切片,在一定程度上提供对网络连通性的直观认识。在豆瓣和 Goodreads 上重复多次实验,所得切片结构类似,篇幅所限仅在图 4 中列出 Goodreads 的两个切片。

从图 4 中可以看到,Goodreads 社交网络的局部切片呈现出明显的复连通性质,任何两个节点之间都有多条路径可达。这是社交网络作为复杂网络的典型特征,路径的冗余程度保证了连通能力的鲁棒性,随机移除少量节点并不会严重破坏网络的连通性。同时,节点之间的相对地位是模糊的,缺乏清晰的拓扑结构。与之形成鲜明对比的是,影响力骨架的局部切片表现出清晰的层级关系,接近于单连通的树状结构,绝大多数节点只有一个父节点,任意两个节点之间通常只有一条路径。可以很容易把这一切片分成若干个小团体,每个小团体是一个典型

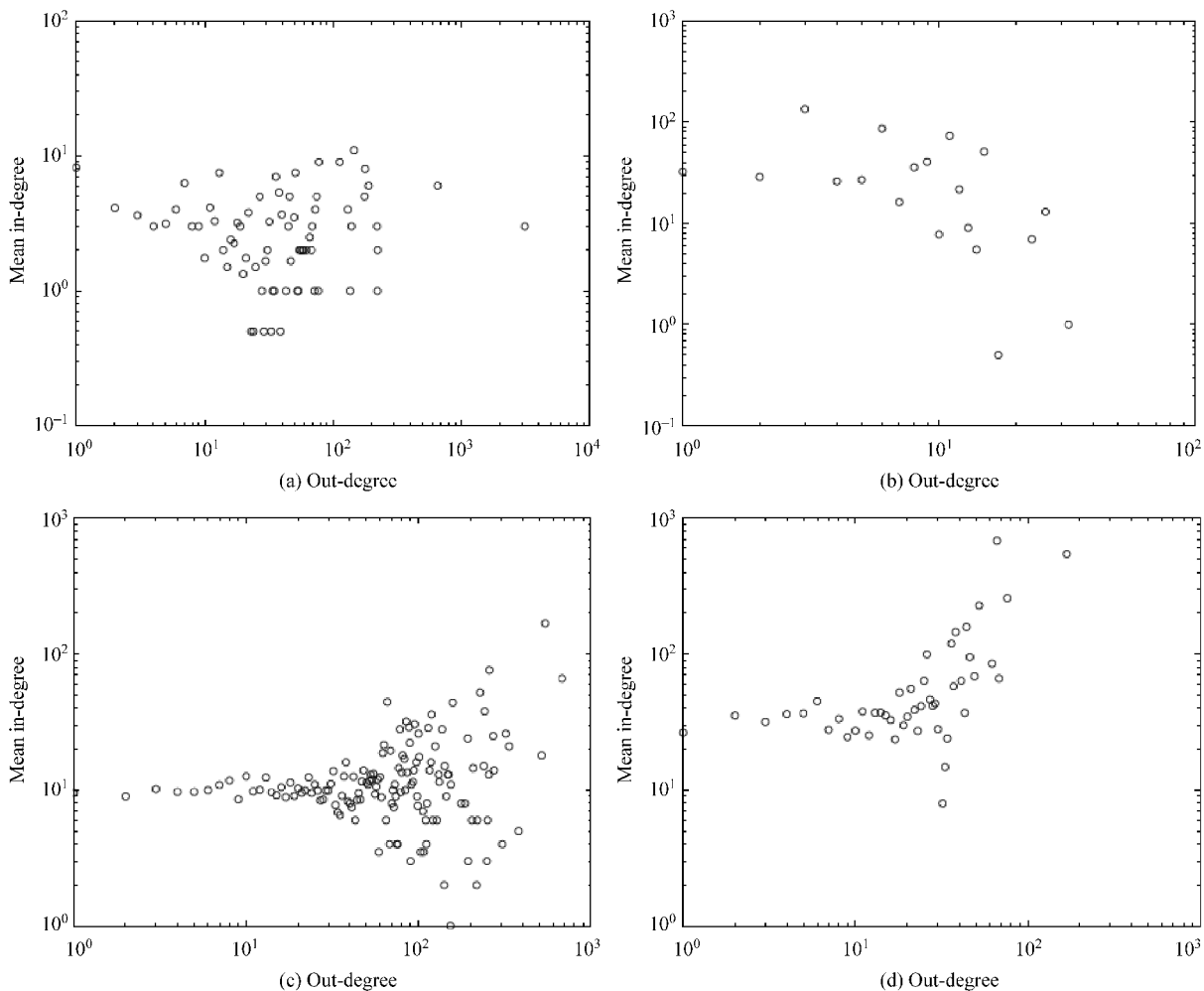


图 3 影响力骨架中,固定出度/入度的节点的平均入度/出度值  
(a)和(b);豆瓣;(c)和(d);Goodreads

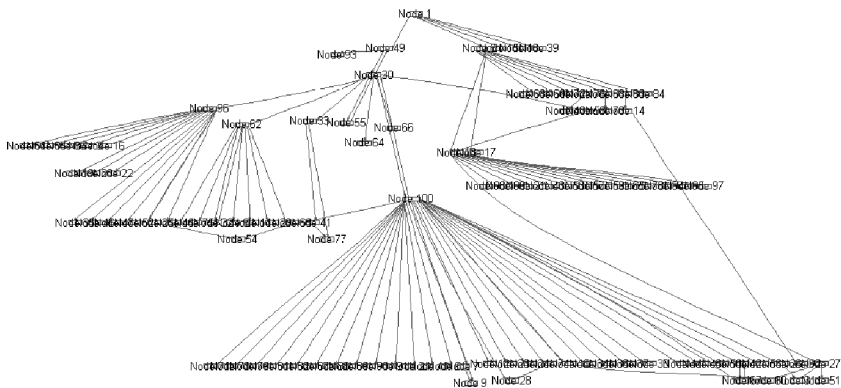
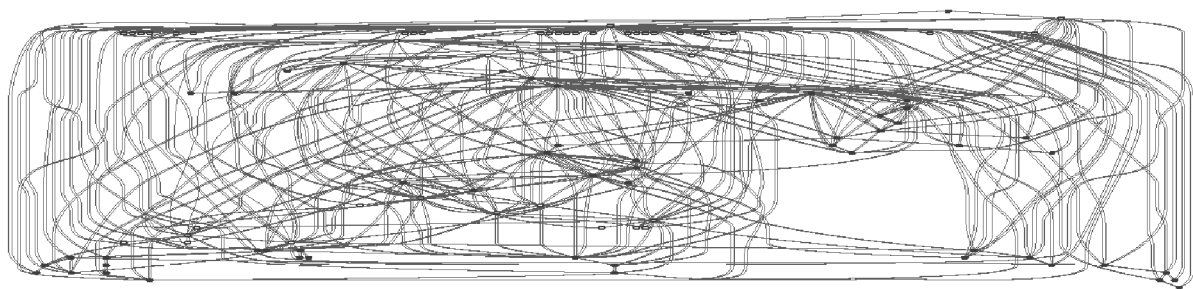


图 4 Goodreads 社交网络(上图)和影响力骨架(下图)的局部切片

的星型结构,由一个中心节点和一组边缘节点构成。中心节点在小团体中的重要地位远远超过边缘节点,一方面它是连接外部的主要(常常是唯一)通道,承担整个小团体的信息外源角色,另一方面它是连接内部各个节点的桥节点,承担小团体的内部交互功能。

这两个切片的显著对比引发了我们对社交网络的信息传播能力的思考。从表面上看,社交网络上的信息传播的途径是冗余而多样化的,具有很强的鲁棒性。但由于很多连边并不具有实质影响的能力,真实的信息传播是沿着树状结构的层级进行的。信息在每一个小团体之间传递,很多情况下信息必须到达每个小团体的中心节点才能接触到围绕在其周围的边缘节点。这些中心节点作为结构洞大量存在,导致信息传播严重依赖于这类节点。这提醒我们,真实信息传播的抗干扰能力并不像直接分析社

交网络所表现出来的那么强,事实上只需要识别出具有结构洞地位的局部中心节点就能有针对性地针对信息传播发起有效攻击、监测和干预。

### 3.4 传播效率

信息传播的效率是信息传播分析的重要尺度。能否从少量节点出发让信息覆盖大部分节点,体现了信息传播的效率,构成了病毒式营销等各种应用的基础。在社交网络中,节点的影响力通常被定义为节点作为信息源能够把信息扩散到的范围大小。虽然社交网络和影响力骨架都具有相同的度分布(如图5所示,社交网络和影响力骨架的节点出度分布都服从幂律分布且幂指数接近),但这并不保证它们具有相似的传播效率。我们以大度节点的覆盖能力作为网络扩散能力的度量。将社交网络中的节点按出度排序,选择出度最大的一部分节点,观察这些

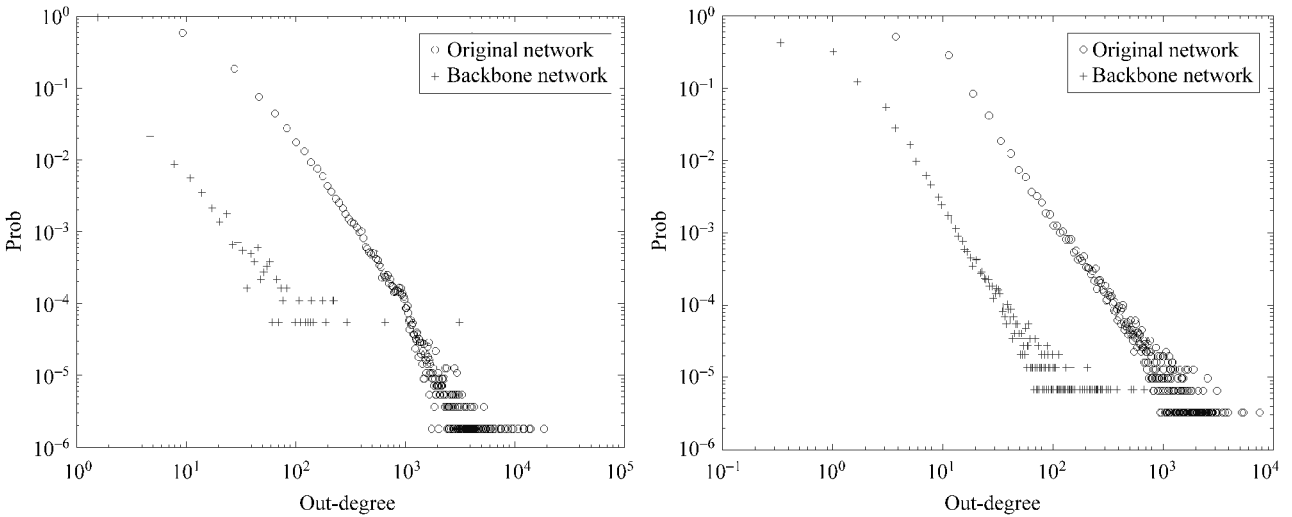


图5 豆瓣(左)和 Goodreads(右)的社交网络(圆圈)和影响力骨架(十字)的节点出度分布

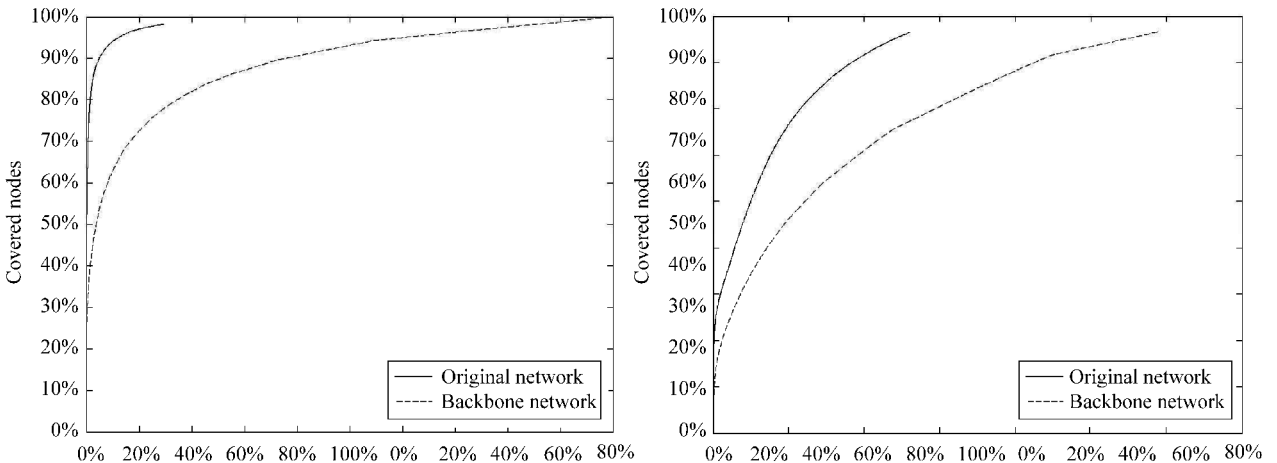


图6 豆瓣(左)和 Goodreads(右)的社交网络(实线)和影响力骨架(虚线)的节点覆盖率,即出度排序靠前的一部分节点(横轴)的出边所能覆盖的节点数(纵轴)

节点的出边所能覆盖的节点占全网的比例,类似于按出度选择一组种子节点以激发病毒式营销。我们看到,社交网络上出度最大的 20% 的节点的出边能够覆盖全网 80% 的节点,表明社交网络是一个直径较小且大多数节点围绕在大度节点周围的网络。这样的网络具有高效的传播能力,符合 20-80 帕累托定律,只需借助 20% 的高影响力者就可以将信息扩散到达 80% 的人群。然而,真实的信息传播效率并没有这么高。对影响力骨架进行同样的分析,我们发现影响力骨架中出度最大的 20% 的节点只能覆盖到全网 50% 的节点。要覆盖全网 80% 的节点,我们需要动员出度排名靠前的 60% 的节点。影响力骨架较低的连边密度和接近单连通的拓扑性质是其主要因素。也就是说,要让 80% 的人群“知道”某一观点,只需发动少量高影响力者即可,但要让 80% 的人群“同意”某一观点,需要动员过半人群才能实现。这昂贵的动员成本表明,无论是病毒营销还是舆论干预,其真实难度都高于我们此前对社交网络分析所得的认识。

## 5 小结

社交网络表征了信息传播的基础结构,而通过量化用户之间影响力提取的影响力骨架则表征了信息传播的功能。这两个网络在各方面存在一定的差异,表明了结构与功能的不一致性。我们认为影响力骨架体现了信息传播的真实路径和范围,对其拓扑结构分析可以帮助我们更准确地认识社交网络上的信息传播的规律。我们首先从符号网络的角度探讨了信息借以传播的影响力边的正负性,利用社会平衡理论验证了影响力边是正关系,指出信息传播的结果是意见同步,即促使相邻个体的态度趋向一致。在微观层面,我们讨论了个体在信息传播中的角色分化,虽然社交网络的高互惠系数导致绝大部分节点的出度入度平衡,难以划分出清晰的角色。但在真实信息传播中个体明显分化为两类不同的节点,即出度远大于入度的主导者和入度远大于出度的追随者,信息在这两类个体之间的传播构成了信息的微观传播机制。同时我们发现,对于任何个体,能真实影响到他的影响力来源总是有限的,这一现象可能归因于个体接受信息的能力是有限的。在宏观层面,我们讨论了网络的拓扑性质对于信息传播的影响。在冗余的复连通社交网络上,真实信息传播路径更接近于单连通的树状结构,信息在大小不

一的星型结构小团体之间跳转,存在相当数量的结构洞可以对信息传播产生关键性影响。因此信息传播的抗干扰能力是比较脆弱的,只需识别并控制结构洞即可实现高效的攻击和干预。最后我们讨论了信息传播效率。旨在“了解”的信息传播符合传统的 20-80 帕累托定律,只需动员 20% 的高影响力人群即可帮助信息到达 80% 的目标人群,但旨在“说服”的信息传播的效率较低,需要动员过半的高影响力人群才能让 80% 的目标人群接受一条信息,高昂的动员成本可能导致病毒营销等信息传播应用的难度高于预期。

## 参考文献

- [1] Robert M Bond, et al. A 61-million-person experiment in social influence and political mobilization [J]. *Nature*, 2012, 489(7415): 295-298.
- [2] Johan Bollen, Huina Mao. Twitter Mood as a Stock Market Predictor [J]. *Computer*, 2011, 44(10): 91-94.
- [3] David Kempe, Jon Kleinberg, Eva Tardos. Maximizing the spread of influence through a social network [C]//Proceedings of the 9th ACM SIGKDD international conference on knowledge discovery and data mining. Washington, D. C. 2003.
- [4] Jeremy Ginsberg, et al. Detecting influenza epidemics using search engine query data [J]. *Nature*, 2009, 457(7232): 1012-1014.
- [5] Jimeng Sun, Jie Tang. A Survey of Models and Algorithms for Social Influence Analysis [M]. *Social Network Data Analytics*. C. C. Aggarwal, Springer US, 2011: 177-214.
- [6] Lü Linyuan, Chen Duan-Bing, Zhou Tao. The small world yields the most effective information spreading [J]. *New Journal of Physics*, 2011, 13(12): 123005.
- [7] Lü Linyuan, et al. Leaders in Social Networks, the Delicious Case [J]. *PLoS ONE*, 2011, 6(6): e21202.
- [8] Jie Tang, Tiancheng Lou, Jon Kleinberg. Inferring social ties across heterogenous networks [C]//Proceedings of the 5th ACM international conference on Web search and data mining. Seattle, Washington, USA, ACM, 2012: 743-752.
- [9] Johan Ugander, et al. Structural diversity in social contagion [J]. *Proceedings of the National Academy of Sciences*, 2012, 109(16): 5962-5966.
- [10] Peng Bao, et al. Popularity prediction in microblogging network: a case study on sina weibo [C]//Proceedings of the 22nd international conference on World Wide Web companion, Rio de Janeiro, Brazil,



2013

- [11] Kazumi Saito, Ryohei Nakano, Masahiro Kimura. Prediction of Information Diffusion Probabilities for Independent Cascade Model [C]//Proceedings of the Knowledge-Based Intelligent Information and Engineering Systems. Springer Berlin Heidelberg. 2008: 67-75.
- [12] Amit Goyal, Francesco Bonchi, Laks V S Lakshmanan. Learning influence probabilities in social networks [C]//Proceedings of the 3rd ACM international conference on Web search and data mining, New York, USA, ACM, 2010.
- [13] Manuel Gomez-Rodriguez, Jure Leskovec, Andreas Krause. Inferring Networks of Diffusion and Influence [J]. ACM Trans. Knowl. Discov. Data, 2012, 5(4): 1-37.
- [14] Manuel Gomez-Rodriguez, et al. Structure and dynamics of information pathways in online media [C]//Proceedings of the 6th ACM international conference on Web search and data mining, Rome, Italy, ACM, 2013.
- [15] Paula Fitzgerald Bone. Word-of-mouth effects on short-term and long-term product judgments [J]. Journal of Business Research, 1995, 32(3): 213-223.
- [16] Sinan Aral, Dylan Walker. Identifying influential and susceptible members of social networks [J]. Science, 2012, 337(6092): 337-341.
- [17] Fang Wu, Bernardo A. Huberman. Novelty and collective attention [J]. Proceedings of the National Academy of Sciences, 2007, 104(45): 17599-17601.
- [18] Meredith Ringel Morris, et al. Tweeting is believing? understanding microblog credibility perceptions [C]//Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work. Seattle, Washington, USA, ACM, 2012.
- [19] Mark SGranovetter. The strength of weak ties [J]. American journal of sociology, 1973, 78(6): 1360-1380.
- [20] Junming Huang, et al. Exploring social influence via posterior effect of word-of-mouth recommendations [C]//Proceedings of the 5th ACM international conference on Web search and data mining. Seattle, Washington, USA, ACM, 2012.
- [21] Ruslan Salakhutdinov, Andriy Mnih. Probabilistic Matrix Factorization [J]. Advances in Neural Information Processing Systems 20. Cambridge, MA, MIT Press, 2008: 1257-1264.
- [22] Heider F. Attitudes and cognitive organization [J]. Journal of Psychology, 1946, 21(1): 107-112.



黄俊铭(1984—), 博士, 主要研究领域为信息传播、社交网络分析。

E-mail: mail@junminghuang.com



沈华伟(1982—), 博士, 副研究员、硕士生导师, 主要研究领域为网络科学、社交网络分析、数据挖掘。

E-mail: shenhuawei@ict.ac.cn



程学旗(1971—), 研究员、博士生导师, 主要研究领域为网络科学、网络与信息安全、互联网搜索与服务。

E-mail: cxq@ict.ac.cn