

节点中心性对复杂网络传播模式的影响分析

苏臻 高超 李向华

Analysis of the effect of node centrality on diffusion mode in complex networks

Su Zhen Gao Chao Li Xiang-Hua

引用信息 Citation: *Acta Physica Sinica*, 66, 120201 (2017) DOI: 10.7498/aps.66.120201

在线阅读 View online: <http://dx.doi.org/10.7498/aps.66.120201>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2017/V66/I12>

---

您可能感兴趣的其他文章

Articles you may be interested in

基于多重影响力矩阵的有向加权网络节点重要性评估方法

Evaluation method of node importance in directed-weighted complex network based on multiple influence matrix

物理学报.2017, 66(5): 050201 <http://dx.doi.org/10.7498/aps.66.050201>

基于最小刚性图代数特性的无线网络拓扑优化算法

Topology optimization algorithm for wireless networks based on the algebraic properties of minimum rigid graph

物理学报.2016, 65(24): 240201 <http://dx.doi.org/10.7498/aps.65.240201>

基于复杂网络理论的多元混合空管技术保障系统网络特征分析

Analysis on network properties of multivariate mixed air traffic management technical support system based on complex network theory

物理学报.2016, 65(14): 140203 <http://dx.doi.org/10.7498/aps.65.140203>

交织型层级复杂网

Interlacing layered complex networks

物理学报.2014, 63(19): 190201 <http://dx.doi.org/10.7498/aps.63.190201>

基于有限资源提升网络可生存性的拓扑重构方法

A topological reconfiguration method for enhancing networks survivability with limited resources

物理学报.2014, 63(17): 170201 <http://dx.doi.org/10.7498/aps.63.170201>

## 节点中心性对复杂网络传播模式的影响分析\*

苏臻 高超† 李向华

(西南大学计算机与信息科学学院, 重庆 400715)

(2017年1月8日收到; 2017年3月11日收到修改稿)

在众多的重要节点评估方法研究中, 具有较高中心性的节点一直是关注的焦点, 许多传播行为的研究也主要围绕高中心性节点展开, 因此在一定程度上忽略了低中心性节点对传播行为的影响. 本文从传播异构性角度, 通过初始感染最大中心性节点和最小中心性节点揭示网络结构异构性对信息传播的影响. 实验结果表明, 传播过程中存在“链型”和“扇型”两种传播模式, 在初始感染比例不断提升的情况下, 两种传播模式的相互转换引发传播速率的变化, 进一步促使非线性传播规模交叉现象的产生. 这一现象说明, 在宏观的信息传播过程中, 最小中心性节点的影响力不容忽视, 尤其在初始感染比例升高时, 最小中心性节点比最大中心性节点更具传播优势.

**关键词:** 复杂网络, 结构异构性, 中心性, 传播速率

**PACS:** 02.10.Ox, 89.75.Fb, 89.75.-k

**DOI:** 10.7498/aps.66.120201

## 1 引言

复杂网络中的重要节点挖掘作为网络科学的重要研究内容之一, 已经被应用到诸多领域, 尤其在各类传播行为(如流行病爆发<sup>[1]</sup>、信息扩散<sup>[2]</sup>、病毒传播<sup>[3]</sup>)的研究中体现得尤为重要. 所谓重要节点是指相比于网络中的其他节点, 能够对网络的整体结构(如网络连通性、平均距离等)以及功能(如网络的鲁棒性、同步等)产生重要影响力的一类特殊节点, 其影响力主要体现在当这部分占比极少的节点受到影响时, 会加速传播过程, 扩大传播规模<sup>[4,5]</sup>. 由此可见, 对于重要节点的挖掘非常重要.

目前, 已有很多指标用于评估节点中心性(或者说影响力、重要性)<sup>[6,7]</sup>, 比如度数<sup>[8]</sup>、核数<sup>[9]</sup>、介数<sup>[10]</sup>以及特征向量中心性<sup>[11]</sup>等. 尽管已有方法对传播行为控制<sup>[12]</sup>、传播源定位<sup>[13,14]</sup>以及传播网络推演<sup>[15]</sup>等起到了重要作用, 但是这些研究都将研究的重点集中在具有较高中心性的节点上, 在一定程度上忽视了具有较低中心性的节点对传播行为的影响. 因此, 本着探究复杂网络中传播行为异

构性的思路, 首先发现了两种不同的传播模式(“链型”和“扇型”), 然后通过实验展现最大、最小中心性节点对传播过程的影响.

本文基于四种常用的中心性评估指标(度数、核数、介数和特征向量), 通过大量的传播仿真实验, 揭示了传播过程中的两种传播模式(“链型”和“扇型”)以及传播规模交叉现象, 即随着初始感染比例的升高, 初始感染最小中心性节点情况下的病毒传播规模以及传播速率会超越初始感染最大中心性节点的情况. 与此同时, 伴随着初始感染比例的提升, 两种传播过程下的传播规模交叉时间点不断提前. 这一现象一方面说明节点中心性的高低对信息传播起着重要作用, 另一方面也反映复杂网络中传播行为的复杂性.

## 2 结构异构性

## 2.1 网络结构异构性分析

从网络的全局结构特征考虑, 常见的规则网络模型(如全局耦合网络、最近邻耦合网络以及星形

\* 国家自然科学基金(批准号: 61402379, 61403315)、中央高校基本科研业务费专项资金(批准号: XDJK2016A008, XDJK2016B029)和重庆市科技研发基地建设计划(国际科技合作)项目(批注号: cstc2015gjh40002)资助的课题.

† 通信作者. E-mail: cgao@swu.edu.cn

耦合网络)在整体结构上表现出网络平均距离小、链接度比较高的特点<sup>[16]</sup>,但节点间的差异性不明显,无法对实际网络结构进行很好的刻画.与规则网络相反的随机网络,虽然与实际网络存在相似的稀疏性、最大块特性以及小世界特性,但在聚类特性以及度分布上仍然存在明显的差异.考虑到规则网络和随机网络分别在刻画实际网络小世界特征与聚类特性时的不足,小世界网络模型作为过渡模型被提出<sup>[17]</sup>.由于Erdos-Renyi随机图和Watts-Strogatz小世界模型中网络的度分布可利用泊松分布来近似刻画,因此这类网络也被称为均匀网络或指数网络<sup>[18]</sup>.与此同时,Barabási和Albert<sup>[19]</sup>于1999年提出的Barabási-Albert(BA)无标度网络模型作为揭示复杂网络无标度特性的经典模型,使得无标度网络成为当今网络科学研究的重要课题之一<sup>[20]</sup>.

从网络传播的角度考虑,对于均匀网络,存在一个有限的正传播临界值 $\lambda_c = \frac{1}{\langle k \rangle}$ ( $\langle k \rangle$ 表示网络的平均度),当且仅当有效传播率 $\lambda$ 大于 $\lambda_c$ 时,传播行为才能得以保证并最终达到某一平衡状态.在无标度网络中,传播临界值 $\lambda_c = \frac{\langle k \rangle}{\langle k^2 \rangle}$ ,对于幂指数为 $2 < \gamma \leq 3$ 的无标度网络,当网络规模无限增大时,该临界值趋于0,即只要存在传播源,传播行为就会在整个网络中蔓延<sup>[21,22]</sup>.但这些分析有个假设,即利用平均场理论进行分析,使用平均度来度量网络异构性.然而,鉴于节点位置差异导致网络异构性(如中心性大小),尚需进行仿真分析.

## 2.2 节点结构异构性分析

从前面的分析可知,在规则网络中节点间的差异性(如度差异性)不明显,而在指数网络以及无标度网络中这种差异性表现突出,正是源于这种差异性造成了各节点中心性的差异.

本文从众多评估方法中选择四种常用指标(节点度、核数、介数和特征向量),讨论结构异构性对传播过程的影响.度中心性是刻画节点重要性最简单的指标,节点的度数越大则认为其影响力越高,因此度中心性主要是对节点的局部影响力进行刻画<sup>[8,23]</sup>. $k$ -core分解与 $k$ -shell分解相似<sup>[9]</sup>,都是对网络节点重要性的一种粗粒化划分,同样地, $k$ -core值越大则认为节点的影响力越高,利用 $k$ -core划分在某些情况下能排除某些度中心性较高的节点从而找出具有更高影响力的节点.介数中心性刻画了节点对于网络中沿最短路径传输的信息流的控制力<sup>[10]</sup>,即经过某一节点的最短路径越多则认为节点的重要性越高.介数中心性在控制网络传播<sup>[24]</sup>、对度中心性相同的节点进行更精确的划分等方面能起到重要作用<sup>[25]</sup>.特征向量中心性在考虑节点度的同时也考虑了邻居节点的重要性<sup>[11]</sup>,所以特征向量中心性能够对节点的长期影响力进行很好的刻画.

## 2.3 两种传播模式

基于已有的研究,从直观的传播效果来看,如果节点中心性越高,则认为其影响力越大,传播能力越强,因此这类节点被作为传播源的研究也相对广泛.但是,在真实的传播网络中,并不是所有的传播行为都是从高中心性节点开始,从低中心性节点开始的传播行为在现实情况下也普遍存在.

本文通过前期实验发现,在信息传播过程中,由于网络结构的影响,存在两种可相互转换的传播模式:“扇型”和“链型”,如图1所示.

当传播过程从高中心性节点开始时,如图1(a)所示,红色传播源均是高度数的节点,在短时间内,呈现出“扇型”的扩散方式,所以信息传播的速率相对较快.但是,随着传播的继续,一方面,越来越多的高链接度节点成为感染节点,无法为后续传播过

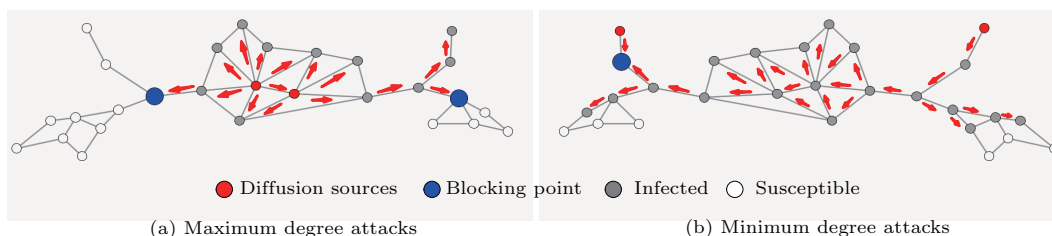


图1 (网刊彩色) 传播过程中的两种传播模式 (a) “扇型”向“链型”转变; (b) “链型”向“扇型”转变

Fig. 1. (color online) Two kinds of propagation processes: (a) The change of diffusion process from “fan-shaped” type to “single-strand” type if the highest-degree nodes are infected sources; (b) the change of diffusion process from “single-strand” type to “fan-shaped” type if the lowest-degree nodes are infected sources.

程提供持续快速的传播速率;另一方面,传播过程中存在“扇型”模式转化为“链型”的情况,从而降低了传播速率;再者,如果在链式结构中出现由于免疫等因素引起的传播“中断”,如图1(a)中蓝色的“中断”点,后续节点则无法接收到传播信息,从而进一步降低了传播速率.

从低中心性节点初始的传播过程则刚好相反,如图1(b)所示,红色传播源均是度最小的节点,在传播开始的短时间内,传播模式主要呈现“链型”,所以传播速率相对较慢.随着传播的继续,一方面传播模式会逐渐由“链型”转化为“扇型”,从而提升传播速率;另一方面,即使在“链型”传播过程中出现“中断”,只要任意一条传播链得以继续传播,传播速率都会得到增强,如图1(b)中左侧传播路径由于蓝色“中断”点的影响无法继续,但是右侧传播路径保证了传播过程的继续.由于两种传播模式的相互转化,在不同初始感染条件下的传播过程中,可能存在传播规模交叉的现象,即两种初始条件下的传播过程在某时刻出现交叉点.

### 3 仿真分析

#### 3.1 传播模型

为了消除传播模型对结果的影响从而验证实验结果的普适性,本文使用两种不同信息传播模型:交互式电子邮件模型<sup>[12]</sup>和传统SI模型<sup>[26]</sup>.

交互式电子邮件模型可以被看作是基于Agent的模型<sup>[12]</sup>,在该模型中,传播过程需要借助用户的行为激发才能进行.具体来说,当网络中的一个用户被感染后,该用户会试图感染周围用户(即向邮件列表中的朋友发送邮件),而邻居用户只有在查看并点击该用户发送的邮件之后才会被感染.因此,传播过程由两种用户行为决定:用户查看邮件的时间和用户是否点击可疑邮件.定义 $N$ 表示网络中用户的总数, $i$ 表示网络中的用户( $i = 1, 2, 3, \dots, N$ ), $T_i$ 表示用户 $i$ 查看电子邮件的时间(即用户 $i$ 检查电子邮件的时间间隔为 $T_i$ ), $P_i$ 表示用户 $i$ 点击可疑邮件的概率.通过分析可知用户 $i$ 查看电子邮件的时间 $T_i$ 主要取决于用户的个人习惯,而用户 $i$ 点击可疑邮件的概率 $P_i$ 主要取决于用户对电子邮件病毒的了解程度以及系统自身的防护功能.由于网络中用户各自的行为相互独立,所以该模型假设不同用户间查看邮箱的时间间隔 $T$ 与点击可疑邮件概率 $P$ 服从高斯分

布,即 $T \sim (\mu_t, \delta_t^2)$ 和 $P \sim (\mu_p, \delta_p^2)$ ,对于某一具体用户 $i$ ,在不同时刻查看邮箱的频率分布服从以 $T_i$ 为参数的分布.依据文献<sup>[12]</sup>,本文用户查看邮件的概率服从指数分布,其幂指数服从正态分布 $N(40, 20^2)$ ,而用户点击可疑邮件的概率服从正态分布 $N(0.5, 0.3^2)$ .

SI模型,作为典型的传染病模型之一<sup>[26]</sup>,在研究许多传播行为时都具有一定的适用性.在该模型中,只存在两种状态的个体,易染状态(S:个体在被感染之前处于易染状态,处于该状态的个体可能被邻居个体感染)和感染状态(I:个体感染了病毒将处于感染状态,并且处于该状态的个体会尝试感染其邻居个体).在任意时刻,感染个体都会以概率 $\beta$ 感染与之接触的易染个体,直到传播系统处于稳定状态.在传播网络中表现为感染节点会以概率 $\beta$ 尝试感染与之直接链接的所有节点.为避免传播速率过快的的问题,本文设定易染节点被感染的概率 $\beta$ 为0.2.

#### 3.2 数据集

仿真实验中采用的信息传播网络( $G_1$ — $G_{12}$ )均是具有无标度特性的网络,在无标度网络下的仿真结论能更好地付诸实际应用.人工网络( $G_1$ — $G_7$ )中, $G_1$ — $G_3$ 、 $G_4$ — $G_7$ 分别是基于BA模型<sup>[19]</sup>和GLP (generalized linear preference)算法<sup>[27]</sup>生成的,其中 $G_7$ 包含多个社区结构.五个标准网络( $G_8$ — $G_{12}$ )包含安然邮件网络( $G_8$ )<sup>[28]</sup>、大学邮件网络( $G_9$ )<sup>[29]</sup>、自治层Internet网络( $G_{10}$ )<sup>[30]</sup>以及科学协作网络( $G_{11}$ ,  $G_{12}$ )<sup>[31,32]</sup>.所有网络的结构特征信息如表1所列.

#### 3.3 传播实验

本次传播实验基于不同中心性评估方法(度数、核数、介数、特征向量),重点分析了最大中心性节点以及最小中心性节点对传播速率以及传播规模的影响.其中,初始感染最大中心性节点的情景包括最大度、最大节点介数、最大核值以及最大特征向量中心性初始感染,而初始感染最小中心性节点的情景包括最小度、最小节点介数以及最小特征向量中心性初始感染.在不同的初始感染策略下,重复传播100次,每次传播过程的仿真迭代次数 $t$ 为2000,保证最终传播系统的稳定性的同时降低实验结果数据的波动性,最后统计100次运行结果的均值作为实验结果.

表1 网络结构特征

Table 1. Networks with different structures used in our experiments.

Notation	Type	Nodes	Edges	$\langle k \rangle$	$\alpha$	$D$	$C$	$L$
G <sub>1</sub>	BA 无标度网络	1000	2995	5.990	2.78	6	0.034	3.470
G <sub>2</sub>	BA 无标度网络	1000	2003	4.006	2.64	8	0.039	3.997
G <sub>3</sub>	BA 无标度网络	1000	5980	11.960	2.54	4	0.049	2.826
G <sub>4</sub>	人工网络	1000	4228	7.490	2.42	8	0.124	3.283
G <sub>5</sub>	人工网络	1000	4226	8.452	2.64	8	0.043	3.462
G <sub>6</sub>	人工网络	1000	4232	8.464	3.00	8	0.040	3.513
G <sub>7</sub>	人工社团网络	4000	16733	8.366	2.39	11	0.079	5.463
G <sub>8</sub>	安然邮件网络	1238	2106	3.402	2.28	12	0	4.670
G <sub>9</sub>	大学邮件网络	1133	5451	9.622	2.42	8	0.254	3.606
G <sub>10</sub>	自治层 Internet 网络	22963	48436	4.219	2.42	11	0.350	3.842
G <sub>11</sub>	科学协作网络	4158	13428	5.526	2.06	17	0.687	6.049
G <sub>12</sub>	科学协作网络	8638	24827	5.259	3.50	18	0.600	5.945

注:  $\langle k \rangle$ ,  $D$ ,  $C$ ,  $L$  分别表示平均度数、网络直径、平均聚类系数、平均路径长度.

实验结果包含两部分: 传播交叉图和交叉点非线性变化图. 其中, 传播交叉图是对传播规模与传播速率交叉现象的直观刻画, 包含传播感染图(即  $t$  时刻被感染节点总数  $I(t)$ ) 和传播增量图(即  $t$  时刻新增感染数目  $\Delta I(t)$ ),  $I(t)$  和  $\Delta I(t)$  满足:  $\Delta I(t) = I(t+1) - I(t)$ . 而交叉点非线性变化图则进一步展现传播规模交叉的时间点与初始感染比例间的变化关系.

### 3.3.1 传播交叉现象

利用表 1 中的网络数据, 以电子邮件模型以及 SI 模型为传播实验平台, 在不同中心性评估方法下分别以 5%, 10%, 20% 的初始感染比例进行仿真实验. 以 BA 无标度网络 ( $G_1$ ) 作为仿真结果示例, 得到如图 2 和图 3 中的传播交叉图, 其他数据集下的仿真证明见补充材料 (online).

#### 3.3.1.1 传播感染图

传播感染图是对传播过程中传播规模交叉的直观呈现, 在不同传播初始比例下的四个传播图分别代表四种中心性评估方法下的传播实验, 图中的黑色标记点代表传播规模发生交叉的时间点. 图 2 展现了不同传播模型下 BA 无标度网络 ( $G_1$ ) 中的仿真结果, 其中图 2(a1)—(a3) 来自电子邮件模型, 图 2(b1)—(b3) 来自 SI 模型, 其他数据集下的仿真结果见补充材料的图 S1—图 S6 (online). 图 2 中不同传播模型下的仿真结果显示以下两点现象:

1) 本次实验设计的初始感染比例下, 在无标度网络中, 存在传播规模交叉的现象, 即初始感染最

小中心性节点情况下的病毒传播规模超越初始感染最大中心性节点的情况;

2) 随着初始感染比例的提升, 交叉时间点呈现出不断提前的趋势, 传播规模交叉的现象也愈加明显.

#### 3.3.1.2 传播增量图

进一步分析传播过程中传播速率的变化, 统计每次迭代过程中新增感染节点数目. 传播增量图从传播速率的角度对传播过程进行刻画. 图 3 与图 2 中的传播实验一一对应, 图 3(a1)—(a3) 和图 3(b1)—(b3) 分别展示了 BA 无标度网络 ( $G_1$ ) 下基于电子邮件模型和 SI 模型的传播增量图, 其他数据集下的传播增量图见补充材料的图 S7—图 S12 (online). 图 3 中不同传播模型下的传播增量图, 显示以下两点特性.

1) 在初始感染最大中心性节点的情况下, 初期的传播速率相对较快, 但是初始感染比例的提升对这种传播优势有着明显的抑制作用. 相对地, 在初始感染最小中心性节点的情况下传播速率几乎不受负面影响.

在电子邮件病毒传播模型下, 随着初始感染比例的提升, 初始感染最小中心性节点的情况下, 传播速率保持相对稳定. 对比图 3(a1)—(a3) 中的  $m_1$ ,  $m_2$  和  $m_3$  (即  $m_1 > m_2 > m_3$ ) 可以发现, 初始感染最大中心性节点情况下的传播过程受到明显的影响: 达到最大传播速率前的传播加速度逐渐减缓; 传播过程中的最大传播速率逐渐降低; 相同传播时刻下的传播速率逐渐减小.

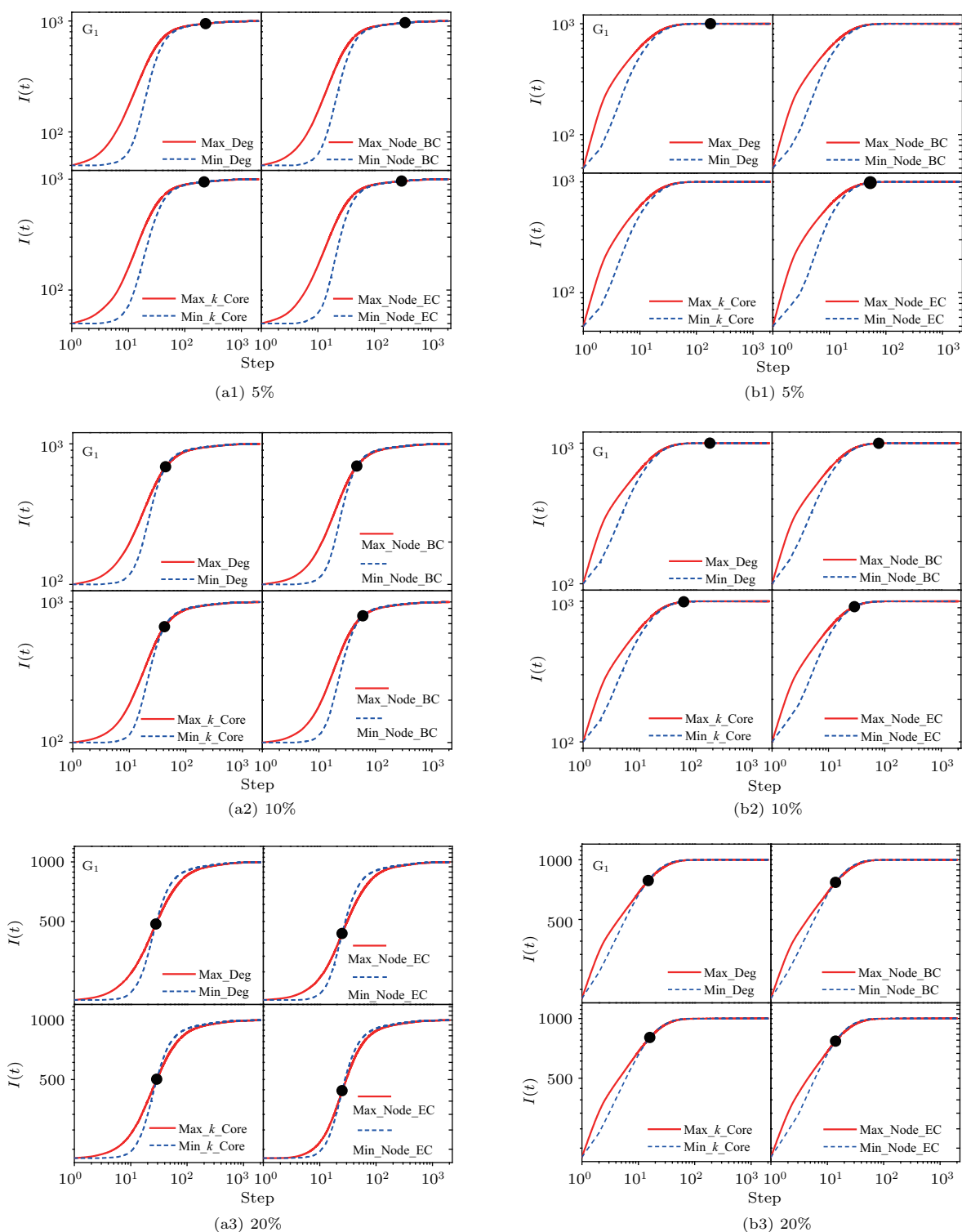


图2 (网刊彩色) 基于 (a1)—(a3) 邮件病毒传播模型和 (b1)—(b3) SI 模型, 初始感染比例为 5%, 10% 和 20% 时病毒在 BA 无标度网络 ( $G_1$ ) 中的传播示意图

Fig. 2. (color online) The illustration of dynamic propagation process in  $G_1$  based on the (a1)—(a3) email propagation model and (b1)—(b3) SI model, in which 5%, 10% and 20% nodes are selected as the initial infected nodes according to four different centralities (i.e., degree, betweenness,  $k$ -core and eigenvector).

在 SI 模型中, 通过对比图 3(b1)—(b3) 中的  $n_1$ ,  $n_2$  和  $n_3$  (即  $n_1 > n_2 > n_3$ ) 可以发现, 随着初始感染比例的提升, 初期传播加速率在数值上呈

现出升高的趋势, 进一步引起相同传播时刻下的传播速率逐渐减小. 而在初始感染最小中心性节点的情况下, 前期的传播速率呈现增强的趋势.

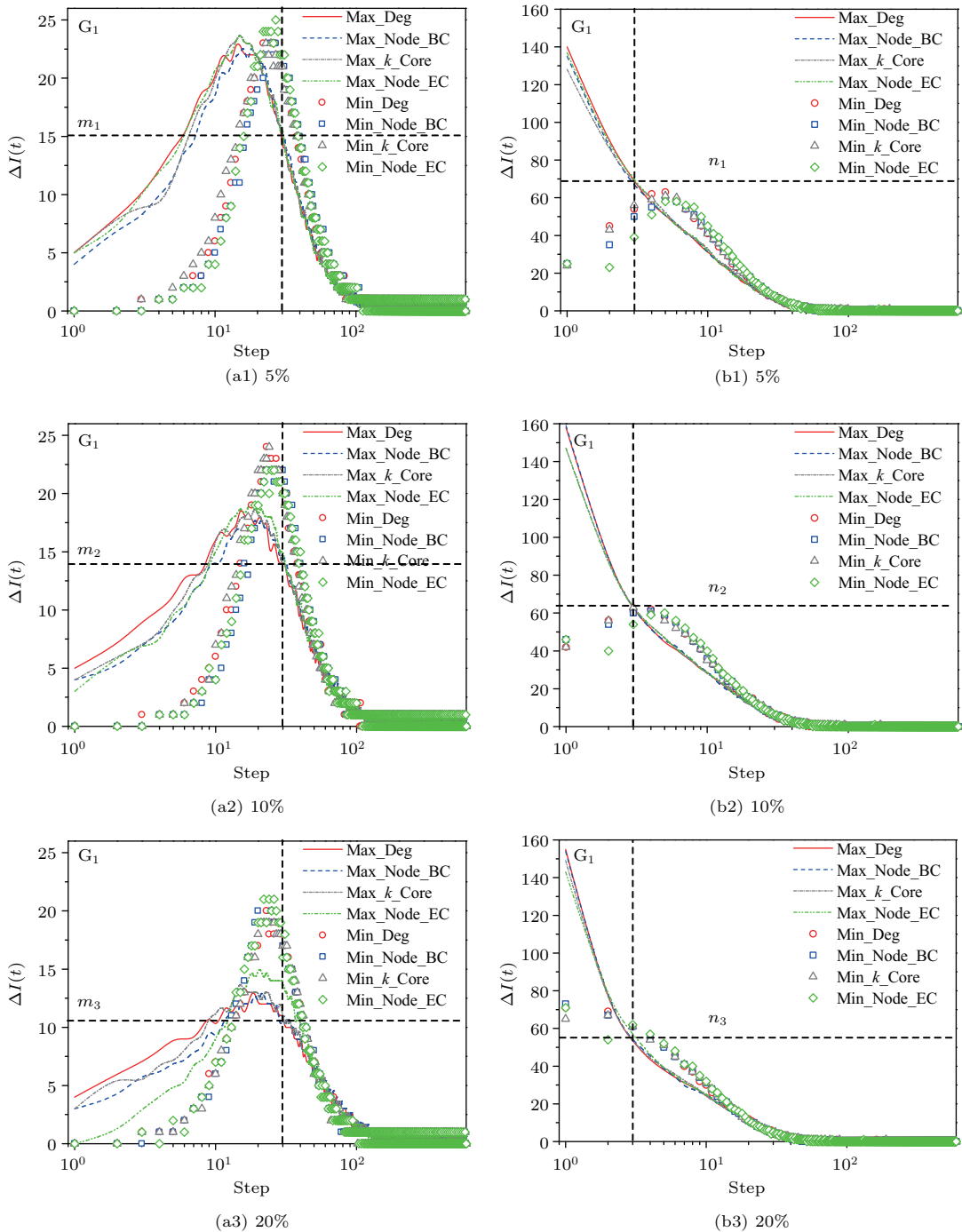


图3 (网刊彩色) 基于 (a1)—(a3) 邮件病毒传播模型和 (b1)—(b3) SI 模型, 初始感染比例为 5%, 10% 和 20% 时病毒在 BA 无标度网络 ( $G_1$ ) 中的传播增量示意图

Fig. 3. (color online) The illustration of dynamic changes of the incremental infected nodes in  $G_1$  based on the (a1)—(a3) email propagation model and (b1)—(b3) SI model, in which 5%, 10% and 20% nodes are selected as the initial infected nodes according to four different centralities (i.e., degree, betweenness,  $k$ -core and eigenvector).

2) 随着初始感染比例的提升, 初次传播速率交叉点呈现出逐渐提前的趋势.

结合图 2 中的传播感染图, 得到以下与传播速率和传播规模交叉相关的传播特性:

1) 在无标度网络下, 初始感染比例的不提高, 会抑制初始感染最大中心性节点情况下的传播

速率, 进一步影响传播规模的扩散; 相对地, 初始感染最小中心性节点情况下的传播速率几乎不受负面影响;

2) 两种传播过程受到的不同影响, 促使其中的初次传播速率交叉点呈现不断提前的趋势, 进一步可能引起传播规模交叉点的产生以及不断提前;

3) 不同结构下的无标度网络中, 传播规模与传播速率交叉点存在一定差异.

### 3.3.1.3 传播模式分析

为了验证传播过程中存在的两种传播模式(即“扇型”传播和“链型”传播), 以BA无标度网络

$G_1$ 、人工网络( $G_4$ )以及标准网络( $G_9$ )为例, 统计每次迭代过程中新增感染节点的平均度数 $\Delta D(t)$ . 图4给出了BA无标度网络 $G_1$ 下的统计结果, 其他数据集下的传播模式验证见补充材料的图S13和图S14 (online).

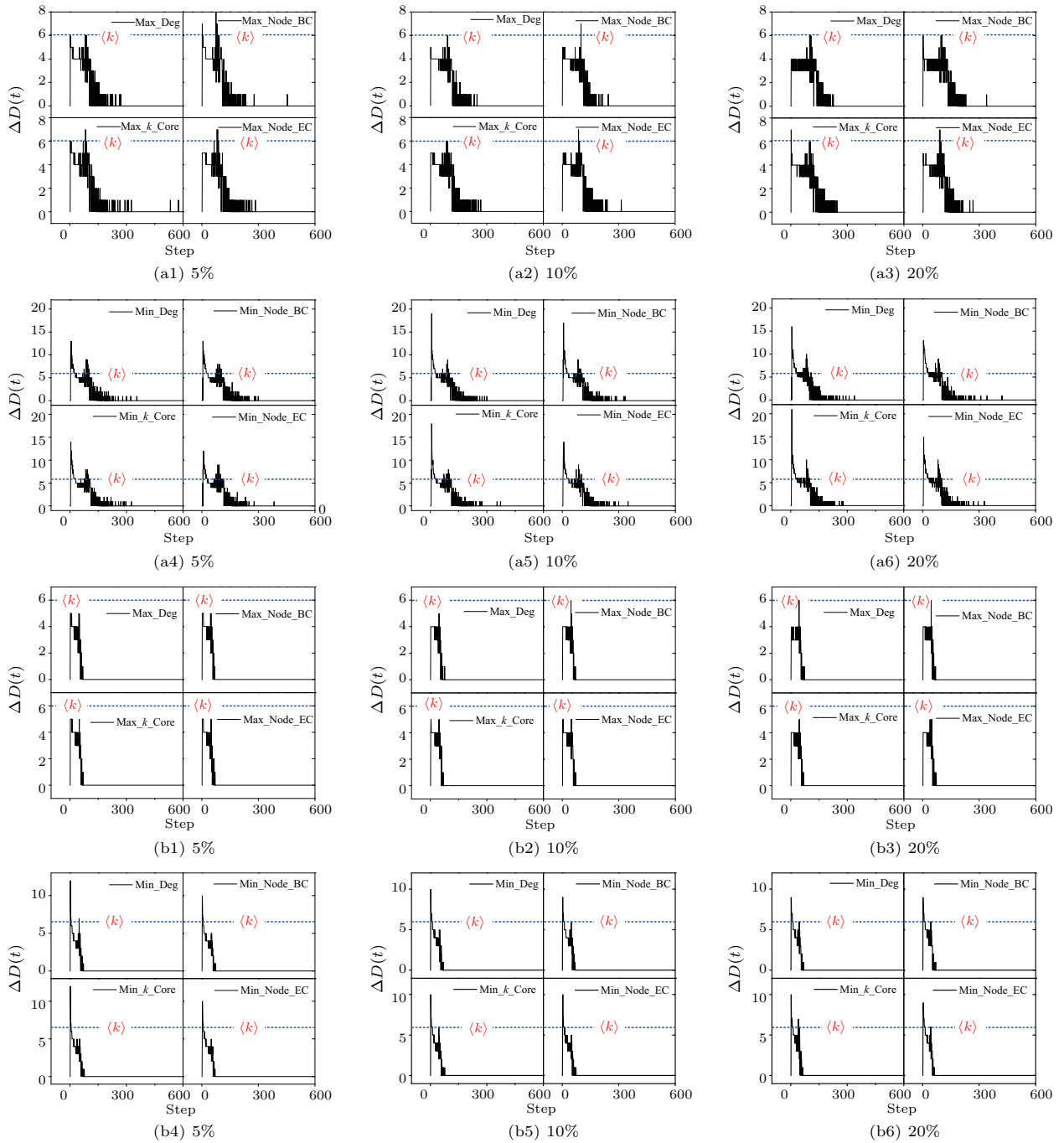


图4 (网刊彩色) 基于(a1)—(a6)邮件病毒传播模型和(b1)—(b6) SI传播模型, 初始感染比例为5%, 10%和20%时病毒在BA无标度网络( $G_1$ )中的平均感染度示意图

Fig. 4. (color online) The illustration of average degree of the incremental infected nodes in  $G_1$  based on the (a1)–(a6) email propagation model and (b1)–(b6) SI model respectively, in which 5%, 10% and 20% nodes are selected as the initial infected nodes according to four different centralities (i.e., degree, betweenness,  $k$ -core and eigenvector).



图4(a1)—(a6)和(b1)—(b6)分别来自电子邮件模型和SI模型,其中,图4(a1)—(a3)和(b1)—(b3)、图4(a4)—(a6)和(b4)—(b6)分别是初始感染最大、最小中心性节点情况下的统计结果,图中 $\langle k \rangle$ 值表示网络 $G_1$ 的平均度.图4揭示出无标度网络下,传播过程中存在以下3点特性.

1) 传播过程中的平均感染度变化整体呈现下降趋势直至传播系统稳定,但是即使在初始感染最小中心性节点的情况下,传播行为趋向于优先感染高链接度的节点.

2) 基于邮件病毒的传播过程中,平均感染度具有两处明显的峰值变化.在相同初始感染比例与中心性评价指标下,对比两种传播过程的平均感染

度变化图(如图4(a1)和(a4), (a2)和(a5), (a3)和(a6))可以发现,随着初始感染比例的升高,初始感染最小中心性节点的情况下,处于两个峰值间的平均感染度明显高于网络平均度 $\langle k \rangle$ ,而初始感染最大中心性节点的情况下,整个传播过程中的平均感染度明显低于网络平均度 $\langle k \rangle$ .与此同时,对比图4(a4)—(a6)可以发现,初始感染最小中心性节点下的峰值间的平均感染度呈现增强的趋势.

在SI模型下,对比图4(b1)和(b4), (b2)和(b5), (b3)和(b6)可以发现,初始感染最小中心性节点情况下的峰值依然高于初始感染最大中心性节点的情况.这一现象为传播速率以及传播规模交叉现象的产生提供了必要条件.

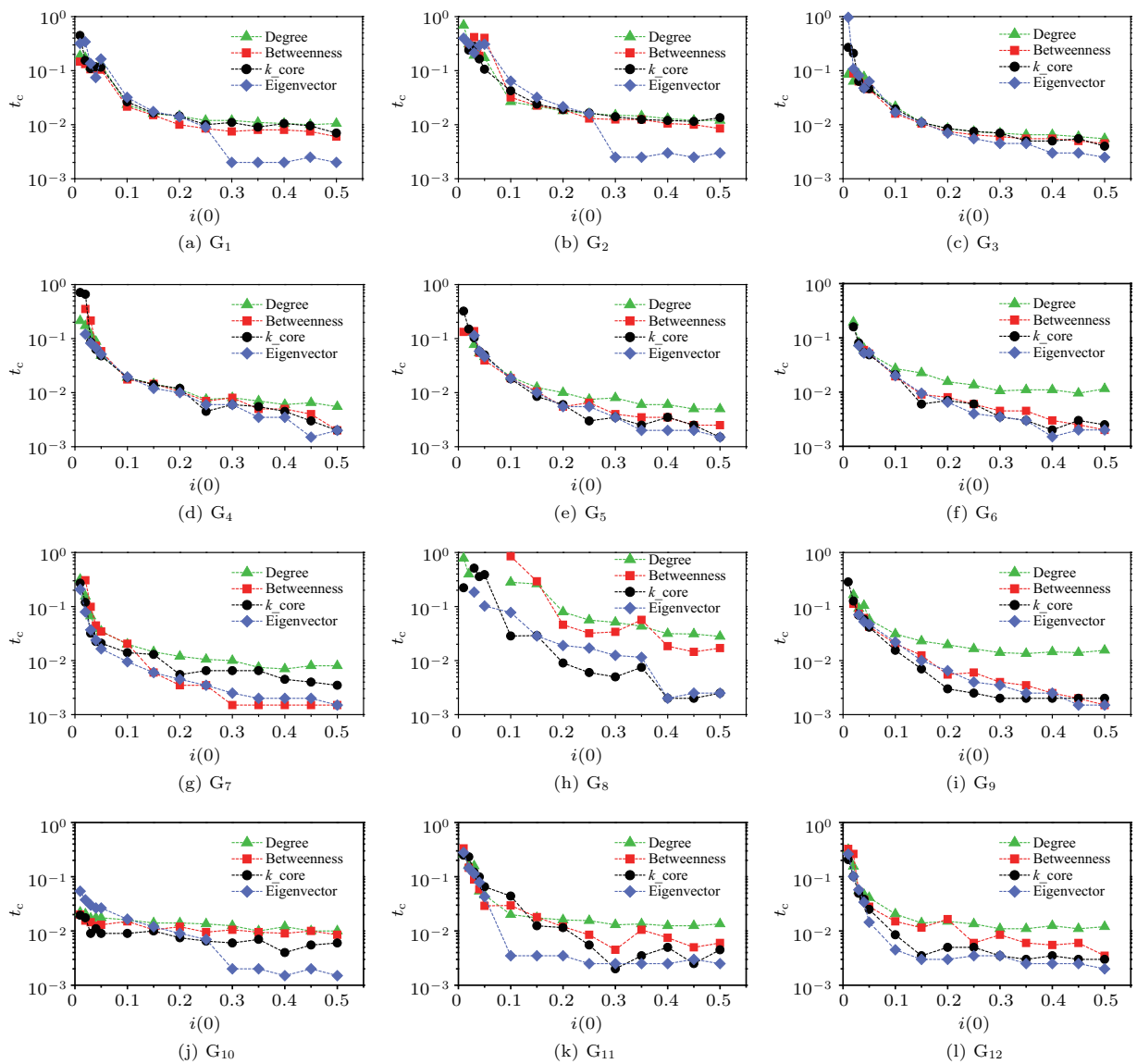


图5 (网刊彩色) 电子邮件病毒传播模型下的交叉点非线性变化图

Fig. 5. (color online) The relationship between  $t_c$  and the  $i(0)$  based on the email propagation model.

3) 基于邮件病毒的传播过程中, 两处明显的峰值现象以及传播过程中的波动现象, 说明传播过程中的确存在两种传播模式. 峰值间的传播过程对应于传播增量图中  $\Delta I(t)$  大于 0 的阶段, 在这一阶段, 平均感染度呈现不低的水平, 说明传播过程表现为“扇型”模式. 当传播增量图中的  $\Delta I(t)$  逐渐趋于 0 时, 平均感染度从第二峰值处呈现降低趋势, 说明传播过程表现为从“扇型”模式向“链型”模式的逐渐转化.

### 3.3.2 初始感染比例对交叉点的影响

3.3.1 小节中的仿真实验揭示出, 传播规模以及传播速率的交叉时间点会随着初始感染比例的提升而提前. 基于相同的传播模型以及中心性评估方法, 下面进一步探究交叉时间点与初始感染比例间的关系.

实验中, 初始感染比例  $i(0)$  从 1% 开始, 分为两

段: 1%—5%, 按照 1% 的比例增长; 5%—50%, 按照 5% 的比例增长. 不同初始感染比例下的交叉时间点  $t_c$  与 3.3.1 小节中的交叉时间点对应,  $t_c$  满足:  $I^{\max}(t) < I^{\min}(t)$ , 其中  $I^{\max}(t)$ ,  $I^{\min}(t)$  分别代表初始感染最大、最小中心性节点情况下,  $t$  时刻被感染节点总数. 为了归一化所有交叉时间点, 令  $t_c = t/2000$ ,  $t_c$  越小, 说明交叉时间点越靠前.

不同的中心性评估方法、传播模型以及网络结构下, 图 5 和图 6 的交叉点非线性变化图揭示出以下 3 点共性:

1) 初始感染比例较低时, 存在传播规模交叉的现象, 但是, 由于传播过程具有一定的随机性, 当初始感染比例提升至 10%—20% 才能保证这一现象的稳定性;

2) 交叉时间点随着初始感染比例的提升逐渐提前, 最后趋于稳定;

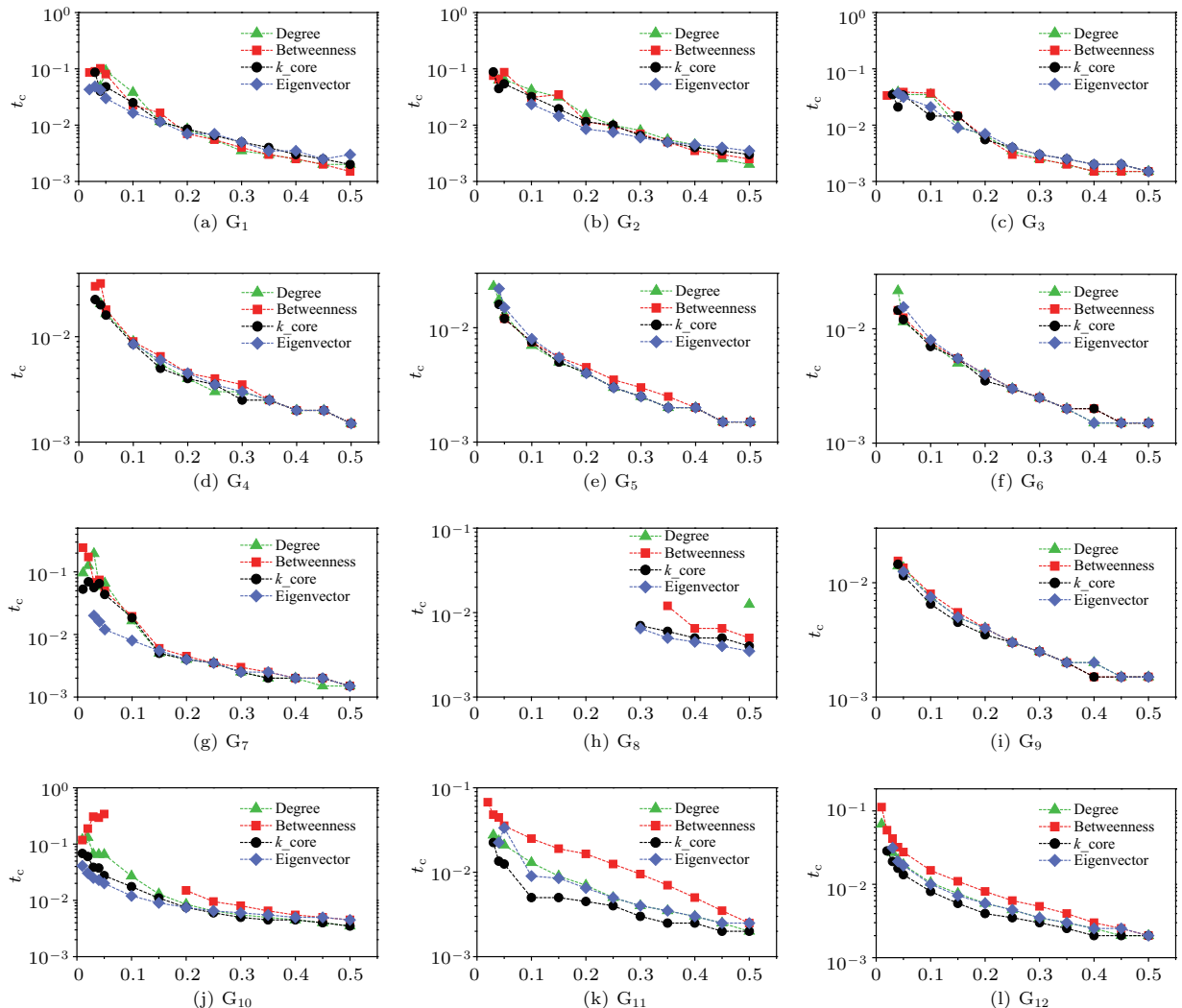


图 6 (网刊彩色) SI 模型下的交叉点非线性变化图

Fig. 6. (color online) The relationship between  $t_c$  and the  $i(0)$  based on SI model.

3) 传播过程受到网络结构的影响, 进一步影响传播规模交叉点的变化; 在电子邮件病毒传播模型下,  $G_8$  网络的交叉点表现出明显的波动性, 在 SI 模型下, 初始感染比例提升至 30% 才产生传播规模交叉的现象.

### 4 讨论

在第 3 部分的实验中, 利用不同的传播模型以及无标度网络证实了传播过程中传播速率与传播规模交叉现象的存在. 其中, 3.3.1 小节给出了这一现象的直观验证, 3.3.2 小节则通过进一步的实验分析验证了初始感染比例与传播交叉现象间的关系. 下面主要针对这一现象做进一步的分析.

基于微分方程的传播模型存在均匀混合的假设, 在一定程度上高估了传播过程中的传播速率 [33]. 本文关注两种极端情况下传播过程中的传播规模交叉现象, 因此, 采用离散化的方法对该现象进行分析.

定义如下:  $N$ ,  $P(k)$  和  $\langle k \rangle$  分别表示网络中的节点总数, 度为  $k$  的节点的比例和网络的平均度;  $t$  时刻度为  $k$  的节点被感染的概率  $i_k(t)$ ,  $t$  时刻感染节点的密度  $i(t)$ ;  $t$  时刻感染节点数  $I(t)$ ,  $t-1$  时刻到  $t$  时刻新增感染节点数  $\Delta I(t)$ , 与第 3 部分中的统计量一致; 平均传播率  $\bar{\lambda}$ , 在 SI 模型下, 该传播率与节点被传染的概率  $\beta$  一致, 而在邮件病毒模型中, 可将人类行为因素等效为异构化的传播率, 在分析过程中使用平均的传播率. 两种传播过程下, 各变量分别采用上标进行区分.

仿真实验中只考虑节点从健康转变为感染状态, 根据文献 [33, 34], 传播过程中  $t-1$  到  $t$  时刻:

$$i_k(t) = i_k(t-1) + \bar{\lambda}k(1 - i_k(t-1))\Theta(t-1),$$

$$\Theta(t) = \frac{\sum_k (k-1)P(k)i_k(t)}{\langle k \rangle}, \quad (1)$$

其中,  $\Theta(t)$  表示度为  $k$  的节点的感染邻居的密度. 进一步得到  $t$  时刻传播速率与  $t$  时刻传播规模的离散化表示:

$$\Delta I(t) = \sum_k \bar{\lambda}k(1 - i_k(t))\Theta(t), \quad (2)$$

$$I(t) = \sum_k i_k(t)P(k)N. \quad (3)$$

根据 (1) 式可知, 初始感染状态  $i(0)$  对整个传播过程影响重大. 同时, (3) 式表明, 当传播系统中感染节点比例较大的情况下, 此时传播规模交叉点的存

在性取决于  $i(t)$ , 因此, 需要对初始感染状态  $i(0)$  以及  $i(t)$  进行分析.

初始状态下满足  $I^{\max}(0) = I^{\min}(0)$ , 由于无标度网络的度分布可以用幂律分布 (即  $P(k) \sim k^{-\gamma}$ ) 近似刻画, 因此, 在初始状态下, 存在感染密度的差异  $\sum_k i_k^{\max}(0) > \sum_k i_k^{\min}(0)$ , 这必然会引起后续传播行为的不一致性.

当系统中的感染节点比例足够大 (即感染节点中包含不同度值的节点), 对 (1) 式从初始状态  $i_k(0)$  推导:

$$i_k(t) = i_k(t-1) + \bar{\lambda}k(1 - i_k(t-1))\Theta(t-1)$$

$$= i_k(0) + \bar{\lambda}k \left( \sum_{t'=0}^{t-1} \Theta(t') - \sum_{t'=0}^{t-1} \Theta(t')i_k(t') \right)$$

$$\approx i_k(0) + \bar{\lambda}k(f_a(t)X - f_b(t)Y),$$

$$X \sim \Theta(0), Y \sim (\Theta^2(0)), t \in [1, +\infty), \quad (4)$$

其中,  $f_a(t)$  和  $f_b(t)$  是关于  $t$  的多项式,

$$f_a(t) = 2^t - 1,$$

$$f_b(t) = \frac{1}{3}[(4^t - 1) + \sum_{i=0}^{t-1} \sum_{t'=0}^i (4^{t'} - 1)],$$

$$t \in [1, +\infty).$$

(4) 式第三步推导忽略了关于  $\Theta(0)$  的高次项, 第一项  $i_k(0)$  描述初始状态, 第二项是对传播增量密度的刻画. 传播系统达到稳定状态前, 构建辅助函数  $\Phi(t, \Theta(0)) = I^{\min}(t) - I^{\max}(t)$  且  $I^{\max}(0) = I^{\min}(0)$ , 则

$$\Phi(t, \Theta(0)) = \sum_k (i_k^{\min}(t) - i_k^{\max}(t))P(k)N$$

$$= \bar{\lambda}N \sum_k (\Theta^{\min}(0) - \Theta^{\max}(0))$$

$$\times [f_a(t) - f_b(t)(\Theta^{\min}(0) + \Theta^{\max}(0))]$$

$$\times kP(k). \quad (5)$$

当某时刻  $t$  出现传播规模交叉的现象, 根据  $\Phi(t, \Theta(0)) \geq 0$  进一步得到

$$\frac{f_a(t)}{f_b(t)} \leq (\Theta^{\min}(0) + \Theta^{\max}(0)), \quad (6)$$

随着传播行为的进行, 存在

$$\lim_{t \rightarrow +\infty} \frac{f_a(t)}{f_b(t)} = 0. \quad (7)$$

因此, 为了避免传播系统短时间内达到稳定状态, 当网络规模趋于无限时, 必然存在时刻  $t_c$  使

得(6)式成立. 对于有限的网络规模, 通过提升初始感染密度  $i(0)$  同样可以满足(6)式, 从而促使两种传播过程下传播规模交叉现象的产生, 并且随着初始感染密度的提升, 传播规模交叉点的产生时间会得到提前. 这一现象表明, 初始感染最大中心性节点情况下, 初始感染比例的提升会对传播速率产生相对较强的负面影响.

结合第3部分的仿真实验, 无标度网络下, 由于网络结构的影响, 传播过程中存在两种传播模式(“链型”和“扇型”), 即使在初始感染最小中心性节点的情况下, 传播行为仍趋向于优先感染高链接度的节点, 从而为传播规模交叉点的产生创造了有利条件. 再者, 随着初始感染比例的提升, 初始感染最大中心性节点的情况下, 高链接度的节点在初始状态被感染, 后续传播速率受到相对抑制, 进一步增大这一现象产生的可能性.

## 5 结 论

本文从网络结构异构性的角度, 基于不同的节点中心性, 分析了结构异构性对传播过程的影响. 通过对传播过程的分析, 详细阐述了传播过程中存在的“链型”和“扇型”两种传播模式以及两种传播模式之间可能的转换过程. 以此为出发点, 利用仿真实验模拟病毒传播过程, 实验结果表明, 由于结构异构性的影响, 传播过程中的确存在两种传播模式(“链型”和“扇型”), 并且在初始感染比例相对较高的情况下, 这两种传播模式的相互转换引起两种极端传播过程下的非线性传播交叉现象. 该现象揭示出初始感染比例的提高可以改变最小中心性节点的影响力, 同时也提醒我们在今后的研究工作中有必要将这一因素纳入考虑范围.

## 参考文献

- [1] Zhang H F, Zhang J, Zhou C S, Small M, Wang B H 2010 *New J. Phys.* **12** 023015
- [2] Saito K, Kimura M, Ohara K, Motoda H 2016 *Inform. Sci.* **329** 985
- [3] Fu C, Min L, Yang J, Xu D L, Liu X Y, Han L S 2015 *Proceedings of IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing* Liverpool, United Kingdom, October 26–28, 2015 p1725
- [4] Ren X L, Lü L Y 2014 *Chin. Sci. Bull.* **59** 1175 (in Chinese) [任晓龙, 吕琳媛 2014 科学通报 **59** 1175]

- [5] Zhao J, Yu L, Li J R, Zhou P 2015 *Chin. Phys. B* **24** 058904
- [6] Song B, Jiang G P, Song Y R, Xia L L 2015 *Chin. Phys. B* **24** 100101
- [7] Liu J G, Ren Z M, Guo Q, Wang B H 2013 *Acta Phys. Sin.* **62** 178901 (in Chinese) [刘建国, 任卓明, 郭强, 汪秉宏 2013 物理学报 **62** 178901]
- [8] Freeman L C 1978 *Soc. Networks* **1** 215
- [9] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E, Makse H A 2010 *Nat. Phys.* **6** 888
- [10] Liu Y Y, Slotine J J, Barabási A L 2011 *Nature* **473** 167
- [11] Borgatti S P 2005 *Soc. Networks* **27** 55
- [12] Gao C, Liu J M, Zhong N 2011 *Knowl. Inf. Syst.* **27** 253
- [13] Jiang J J, Wen S, Yu S, Xiang Y, Zhou W L 2016 *IEEE Trans. Depend. Secure.* pp 1
- [14] Zhang X Z, Zhang Y B, Lü T Y, Yin Y 2014 *Physica A* **442** 100
- [15] Han X, Shen Z S, Wang W X, Di Z R 2015 *Phys. Rev. Lett.* **114** 028701
- [16] Wang X F, Chen G R 2003 *IEEE Circ. Syst. Mag.* **3** 6
- [17] Strogatz S H 2001 *Nature* **410** 268
- [18] Watts D J, Strogatz S H 1998 *Nature* **393** 440
- [19] Barabási A L, Albert R 1999 *Science* **286** 509
- [20] Wang X F, Li X, Chen G R 2012 *Network Science: An Introduction* (Beijing: Higher Education Press) pp270–275 (in Chinese) [汪小帆, 李翔, 陈关荣 2012 网络科学导论 (北京: 高等教育出版社) 第270—275页]
- [21] Pastor-Satorras R, Vespignani A 2001 *Phys. Rev. Lett.* **86** 3200
- [22] Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A 2004 *Phys. Rev. Lett.* **92** 178701
- [23] Albert R, Barabási A L 2002 *Rev. Mod. Phys.* **74** 47
- [24] Goh K I, Kahng B, Kim D 2001 *Phys. Rev. Lett.* **87** 278701
- [25] Estrada E, Rodríguez-Velázquez J A 2006 *Physica A* **364** 581
- [26] Pastor-Satorras R, Vespignani A 2001 *Phys. Rev. E* **63** 066117
- [27] Bu T, Towsley D 2002 *Proceedings of the 21st Annual Joint Conference of the IEEE Computer and Communications Societies* New York, USA, June 23–27, 2002 p638
- [28] UC Berkeley Enron Email Analysis Project, MIT <http://bailando.sims.berkeley.edu/enron/enron.sql.gz> [2017-4-18]
- [29] <http://deim.urv.cat/?aarenas/data/welcome.htm> [2017-4-18]
- [30] University of Oregon Route Views Archive Project, David Meyer <http://routeviews.org/> [2017-4-18]
- [31] Stanford Network Analysis Project, Leskovec J <https://snap.stanford.edu/data/ca-GrQc.html> [2017-4-18]
- [32] Stanford Network Analysis Project, Leskovec J <https://snap.stanford.edu/data/ca-HepTh.html> [2017-4-18]
- [33] Zou C C, Towsley D, Gong W 2007 *IEEE Trans. Depend. Secure.* **4** 105
- [34] Boguñá M, Pastor-Satorras R, Vespignani A 2003 *Statistical Mechanics of Complex Networks* (Berlin: Springer-Verlag) p127

# Analysis of the effect of node centrality on diffusion mode in complex networks\*

Su Zhen Gao Chao<sup>†</sup> Li Xiang-Hua

(College of Computer and Information Science, Southwest University, Chongqing 400715, China)

( Received 8 January 2017; revised manuscript received 11 March 2017 )

## Abstract

The centrality reflects the importance of a node in a complex network, which plays an important role in the propagation dynamics. Many researches in the field of node ranking estimation have revealed the characteristics of higher centrality in the structural dynamics and propagation dynamics. However, there are few reports about the effect of nodes with a relatively lower centrality on propagation process. In this paper, we focus on the effect of heterogeneous structural characteristics on propagation dynamics. First, we select four centrality measurements (i.e., degree, coreness, betweenness, and eigenvector) and initialize source nodes with the maximum and minimum centralities respectively. Then, based on the email propagation model and the SI model, the massive numbers of elaborate simulations are implemented in twelve scale-free networks. These networks include three networks generated by the Barabási-Albert model, four synthetic networks compiled by the GLP (generalized linear preference) algorithm, and five benchmark networks. The simulation results contain two parts: one is the crossover phenomenon of two propagation processes, and the other is the correlation between the crossover point and the proportion of the initial source nodes. We present the crossover of two propagations by calculating the total infected nodes, the incremental infected nodes, and the average degree of the incremental infected nodes. The average degrees of the incremental infected nodes in both synthetic networks and benchmark networks show that there exist two kinds of diffusion modes (i.e., “fan-shaped” type and “single-strand” type). With the increase of the initial source nodes, the interaction between two modes results in the different dynamic changes of two propagations with respect to propagation speed, which may lead to the crossover of two propagations in terms of propagation scale in the propagation process. Specifically, the increase of the initial source nodes would suppress the propagation process in which nodes with the maximum centralities are portrayed as propagating sources. However, such an effect is not observed in the propagation process in which nodes with the minimum centralities are portrayed as propagating sources. Our further simulation indicates that the crossover points appear earlier as the proportion of the initial source nodes increases. And by employing the discrete-time method, we find that such a phenomenon can be triggered exactly by increasing the initial source nodes. This work reveals that the influence of the nodes with the minimum centralities should be taken into consideration because the initial infected nodes with a lower centrality will lead to a larger propagation scale if the initial proportion is high.

**Keywords:** complex networks, heterogeneous structure, centrality measures, propagation speed

**PACS:** 02.10.Ox, 89.75.Fb, 89.75.-k

**DOI:** 10.7498/aps.66.120201

\* Project supported by the National Natural Science Foundation of China (Grant Nos. 61402379, 61403315), the Fundamental Research Funds for the Central Universities of Ministry of Education of China (Grant Nos. XDJK2016A008, XDJK2016B029), and the Chongqing Science and Technology R & D Base Construction (International Science and Technology Cooperation) Project, China (Grant No. cstc2015ghz40002).

<sup>†</sup> Corresponding author. E-mail: [cgao@swu.edu.cn](mailto:cgao@swu.edu.cn)