



尖叫效应传播机制及仿真研究

姜政旭¹, 姚嘉祺¹, 孟令钧², 王 建², 曹永昌²

(1. 中国海洋大学 海德学院, 山东 青岛 266071; 2. 中国海洋大学 数学科学学院, 山东 青岛 266071)

摘要: 基于社交网络模型, 建立了边权偏重因子、主动影响力函数、被动影响度函数和兴趣阈值等一系列网络指标; 借鉴 SEIR 模型, 引入脉冲函数族研究“尖叫效应”的形成过程, 刻画了信息传播中各类人群动态变化过程及变化趋势; 提出了提高兴趣阈值、降低脉冲强度和屏蔽特定用户等控制或消除“尖叫效应”的策略; 基于真实的社交网络结构数据, 通过仿真研究模拟了“尖叫效应”的形成机制, 验证了本文提出的控制或消除“尖叫效应”策略的有效性。

关键词: 社交网络; 尖叫效应; 脉冲函数; SEIR 模型; 谣言阻断

中图分类号: O29

文献标志码: A

文章编号: 2095-3070(2023)04-0064-09

DOI: 10.19943/j.2095-3070.jmmia.2023.04.09

0 引言

随着网络信息技术的发展, 在网络信息领域, 一些网络平台利用大数据和人工智能技术, 智能分析用户们的浏览记录和兴趣爱好, 并大量推送恶搞、色情等低俗内容, 以快速博取公众眼球并从中获得巨大的流量与点击率, 这便是“尖叫效应”^[1]. 这种哗众取宠的公共行为容易使公众失去对其他不同信息的接触机会和辨识能力, 导致“信息茧房”的形成. 因此, 探究“尖叫效应”的形成机制, 具有现实意义.

1 社交网络结构

1.1 用户节点与用户联系

社交网络是一个具有小世界、集群分布、幂律度分布等特性的复杂网络^[2-3], 如图 1 所示.

给定一个有向赋权图 $G=(V, E)$, 用点 v_i 表示社交网络中的用户 i , 所有点的集合记为点集 V ; 社交网络中从用户 i 指向用户 j 的有向联系记为有向边 e_{ij} , 所有有向边的集合记为边集 E ; $w_{ij} \in [0, 1]$ 表示用户 i 对用户 j 的影响力, 其值与传播的信息内容相关; 全体社交网络中的用户总数记为 n .

d_i^+ 称为节点 v_i 的出度, 是以 v_i 为起点的边数; 类似地, d_i^- 称为节点 v_i 的入度, 是以 v_i 为终点的边数. 以 v_i 为终点

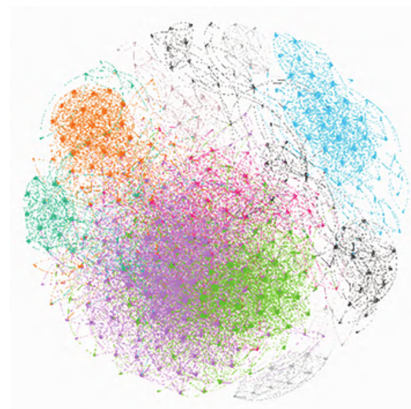


图 1 社交网络图^[4]

收稿日期: 2023-05-07

基金项目: 山东省本科教学改革研究项目(M2021089); 山东省本科教学改革研究重点项目(Z2022044); 中国海洋大学本科教学研究项目(2023JY126)

通讯作者: 曹永昌, E-mail: caoyongchang@ouc.edu.cn

引用格式: 姜政旭, 姚嘉祺, 孟令钧, 等. 尖叫效应传播机制及仿真研究[J]. 数学建模及其应用, 2023, 12(4): 64-72+106.

JIANG ZH X, YAO J Q, MENG L J, et al. Research of the propagation mechanism of scream effect and its simulation(in Chinese)[J]. Mathematical Modeling and Its Applications, 2023, 12(4): 64-72+106.

的有向边的起点的集合称为 v_i 的上游节点集, 记为 $N^-(v_i)$; 以 v_i 为起点的有向边的终点的集合称为 v_i 的下游节点集, 记为 $N^+(v_i)$.

1.2 影响力指标

为了描述信息传播过程, 给出以下影响力指标.

定义 1 (边权偏重因子) 网络节点 v_i 会受到其上游节点的影响, 上游节点 v_{k_0} 对 v_i 的影响力在上游节点集的总影响力中的占比称为 v_{k_0} 对 v_i 的边权偏重因子, 记为

$$r_{k_0 i} = (w_{k_0 i}) / \left(\sum_{v_k \in N^-(v_i)} w_{ki} \right). \quad (1)$$

显然, $r_{k_0 i} \in [0, 1]$.

定义 2 (主动影响力函数) 不同网络节点自身的影响力是不同的, 其大小与该网络节点的出度和时间都有关系, 称为主动影响力函数, 记为

$$w_i^+(t) = f(v_i) \cdot g(t). \quad (2)$$

现实中, 一般情况下, 拥有更大出度的网络节点能够将信息传递给更多的网络节点, 其对外影响力也更强.

在实际选取 $f(v_i)$ 时, 可以利用不同节点的出度大小对网络节点影响力进行排序. 目前常见的排序方案大多基于局部属性、社团关系、全局属性和随机游走 4 个视角来展开^[5].

在实际选取 $g(t)$ 时, 考虑到目前的社交网络环境, 节点的对外影响力随时间的变化关系往往为非光滑函数, 如脉冲函数.

主动影响力较大的用户通常可以将信息、观点散播给同在社交网络中的大量用户, 引发用户们的关注, 并带动、引导舆论走向, 这类节点被称为“意见领袖”^[6]. 当网络节点接收到了来自上游节点的信息, 是否会对该信息产生兴趣, 与上游节点的主动影响力以及两点的边权偏重因子关系密切.

定义 3 (被动影响度函数) 网络节点 v_i 受到的上游节点的影响程度称为节点 v_i 的被动影响度, 记为 $w_i^-(t)$, 其值取决于上游节点的主动影响力和由上游节点指向 v_i 的边权偏重因子, 即:

$$w_i^-(t) = \sum_{v_k \in N^-(v_i)} w_k^+(t-1) \cdot r_{ki}. \quad (3)$$

定义 4 (兴趣阈值) 网络节点 v_i 的被动影响度不同时, 其反应也会不同, 反应变化的阈值称为节点 v_i 的兴趣阈值, 记为 $\overline{w_i^-}$. 当节点的被动影响度超过自身的兴趣阈值时, 用户会传播该信息, 否则不会产生任何信息传播行为. 用 $\chi_i(t)$ 来衡量用户是否传播该信息, 则:

$$\chi_i(t) = \begin{cases} 1, & w_i^-(t) > \overline{w_i^-}, \\ 0, & w_i^-(t) \leq \overline{w_i^-}. \end{cases} \quad (4)$$

由于不同用户的兴趣与认知范围不同, 他们的兴趣阈值是存在差异的, 对于同一个信息, 用户会根据自己的兴趣和认知情况决定是否传播. 对于部分用户, 对该信息比较喜欢, 会倾向于传播该信息, 即其对该信息的兴趣阈值较低; 而对于另一部分用户, 对该信息不感兴趣甚至持抵制、厌恶的态度, 则不会产生传播行为, 即其对该信息的兴趣阈值较高.

2 信息传播过程

2.1 SEIR 模型

在信息传播领域, 由于宏观的信息传播行为与传染病扩散存在相似性, 即用户对网络信息的传播、知晓行为, 相似于流行病扩散过程中感染者传染、易感人群感染^[7], 因此传统的流行病仓室模型常被类比用于研究信息传播过程.

当某一网络节点接收到来自上游节点传来的信息时, 用户可根据对信息产生的兴趣强弱, 选择对信息是否浏览. 用户既可能不会对信息进行浏览并知晓信息具体内容, 也可能会进行浏览并知晓具体内容. 用户在浏览该信息后, 既可以选择立刻向其他网络用户传播信息, 也会选择暂不转发.

借助流行病学模型中的一些较为成熟的研究成果^[8], 将社交网络的全体用户节点分为 4 种状态.

1) 易感节点 S(Susceptible). 这类节点在信息传播初期大量存在于社交网络中, 且目前尚未浏览某一特定信息具体内容, 但在将来的某一时刻可能会选择浏览, 也可能直到信息被淹没也未浏览过, 这取决于用户的兴趣强弱; 或浏览过某特定信息, 但是无向外传播信息的潜在可能的网络用户。

2) 潜伏节点 E(Exposed). 这类节点已浏览过某特定信息, 且有潜在的向外传播信息的可能, 但目前并未将信息传播出去, 在将来的某个时间, 会将信息传播出去。

3) 感染节点 I(Infected). 这类节点已浏览过某特定信息或者是信息爆发的原点, 并正在向其下游网络节点传播该信息。

4) 免疫节点 R(Recovered). 这类节点的下游节点全部都成为 E 节点或 I 节点, 即已完成对外的信息传播, 因此退出了后面的信息传播过程; 或由于其他原因不会接收到该信息。

社交网络中的所有用户节点, 在任意时刻 t , 都处于 4 种状态中的某一种。

2.2 尖叫效应机理分析

尖叫效应指的是突然性输出大量引人注目的内容, 从而实现吸引他人注意力并将特定信息传播出去的目的。这类效应在目前的自媒体平台等十分常见, 例如: 某些信息在传播初期, 借助大 V 的刻意宣传, 再辅以引人眼球的标题与内容, 达到引发舆论、快速实现大范围传播的目的^[1]。

尖叫效应的实现离不开两个关键性因素:

1) 突然性, 即在某时刻, 某些信息忽然出现在社交网络中。这些信息是全新的, 这也意味着, 尖叫效应往往是某类特定信息在某个网络节点爆发, 并开始向周围扩散;

2) 高频率性, 即在短时间内系统大量迸发某特定信息。只有在这种情况下, 单位时间内的信息强度和密度才会足够强大, 使得信息传播的影响力足够大。

在这两种因素的互相促进下, 尖叫效应便在社交网络中爆发。

综合社交网络和传染病模型两种视角, 出现尖叫效应的网络节点即为社交网络中关于某种信息的“超级传播者”, 其拥有极强的信息传播能力, 使得接受其信息的网络用户快速而大量转化为新的信息传播者, 进而开启众多新的传播途径。

因此, 尖叫效应发挥作用时的突出特征是: 向外传播信息的网络节点大量增多, 其主动影响力在某段时间内大大增强。为了描绘这种特征, 定义脉冲函数^[2]如下。

定义 5 形如

$$\delta_i(t-t_0) = \begin{cases} \alpha_i / (2\epsilon_i), & t_i - \epsilon_i < t < t_i + \epsilon_i, \\ 0, & \text{其他} \end{cases} \quad (5)$$

的与时间相关的函数称为脉冲函数, 其中: t_0 是某特定信息在整个社交网络中爆发的时间原点; t_i 是信息开始传播后的某一时刻; 脉冲函数取值代表了脉冲强度, 对于某支特定脉冲函数而言为定值, α_i 是脉冲函数 $\delta_i(t-t_0)$ 的强度系数; ϵ_i 为时间区间半径, 即一支脉冲函数的作用时间长度的一半, 其值为大于零的常数。

利用定义 2, 将脉冲函数 $\delta_i(t-t_0)$ 加入到处于 I 状态的节点的主动影响力中。此外, 考虑不同网络用户在社交网络中的不同位置, 结合各类排序算法的优缺点, 考虑社交网络的节点数量多, 社群网络结构复杂等特点, 使用 LeaderRank 算法, 基于全网络节点的出度进行影响力排序^[9], 得到从高到低的排序集, 记为 $\text{rank}(V)$, 即:

$$\text{rank}(V) = \{x_1, x_2, \dots, x_n\}, \quad (6)$$

其中, $x_i (i=1, 2, \dots, n)$ 是节点 v_i 在社交网络 $G(V, E)$ 中的影响力排名。再依据排名从高到低对节点 v_i 赋值, 即是 LeaderRank 值, 即:

$$\text{LR}(v_i) = \mathcal{L}(x_i), x_i \in \text{rank}(V), \text{LR}(v_i) \in [0, 1], \quad (7)$$

其中, 映射 $\mathcal{L}: \text{rank}(V) \rightarrow [0, 1]$ 是单射。

由公式(2), 在单次尖叫效应下, 正在传播信息的网络节点的主动影响力为:

$$w_i^+(t) = \text{LR}(v_i) \cdot (\delta(t-t_0) + 1). \quad (8)$$

在实际的信息传播过程中, 可能存在多次尖叫效应的叠加与组合。例如: 当社交网络忽然爆出一

个重磅信息,引发了广大用户的关注,掀起一定范围的舆论浪潮.过了一段时间,当流量的热潮即将褪去时,某些大 V 在社交媒体上再度发布关于该事件的反转性信息,与先前形成鲜明对比,此时往往会再度引发广大用户热议.这种现象,可以被称为“ m 次尖叫效应”.

“ m 次尖叫效应”实质上是 m 个尖叫效应的组合,为了描述这种现象,引入脉冲函数族:

$$\delta(t-t_0)=\begin{cases} \alpha_1/(2\varepsilon_1), & t_1-\varepsilon_1<t<t_1+\varepsilon_1, \\ \alpha_2/(2\varepsilon_2), & t_2-\varepsilon_2<t<t_2+\varepsilon_2, \\ \vdots \\ \alpha_l/(2\varepsilon_l), & t_l-\varepsilon_l<t<t_l+\varepsilon_l, \quad l=1,2,\cdots,m, \\ \vdots \\ \alpha_m/(2\varepsilon_m), & t_m-\varepsilon_m<t<t_m+\varepsilon_m, \\ 0, & \text{其他} \end{cases} \quad (9)$$

其中,整个脉冲函数族是由 m 支子脉冲函数在不同的时间区间 $(t_l-\varepsilon_l, t_l+\varepsilon_l)$ 组合而成; $\alpha_1, \alpha_2, \cdots, \alpha_m$ 表示不同子脉冲函数的强度,不同时间区间的中心 t_l 与时间区间半径 ε_l 存在关系:

$$t_l+\varepsilon_l \leq t_{l+1}-\varepsilon_{l+1}, \quad l=1,2,\cdots,m-1, \quad (10)$$

即整个脉冲函数族内各子脉冲函数的时间区间是互不相交的.

公式(9)也可以写成如下形式:

$$\delta(t-t_0)=\sum_{l=1}^m \delta_l(t-t_0). \quad (11)$$

由公式(2),在多次尖叫效应下,正在传播信息的网络节点的主动影响力为:

$$w_i^+(t)=\text{LR}(v_i) \cdot (\delta(t-t_0)+1)=\text{LR}(v_i) \cdot \left(\sum_{l=1}^m \delta_l(t-t_0)+1\right). \quad (12)$$

在社交网络中,某个特定信息开始出现并传播的起始网络节点被称为信息爆发原点,其他节点被称为非信息爆发原点.

值得注意的是,节点的主动影响力并不完全等同于其实际对外影响力.接收上游节点集传来的信息但并未达到兴趣阈值的非信息爆发原点的用户,不会主动参与信息传播过程,实际对外影响力为 0.

而已达到兴趣阈值的非信息爆发原点的用户,其信息传播强度还与被影响程度正相关,即:若其被动影响度越高,则其主动影响力越大.记衡量这种效应的指标为传播因子 $\mu_i(t)$,即:

$$\mu_i(t)=w_i^-(t)/\overline{w_i^-}, \quad (13)$$

则公式(4)可被重新表示为:

$$\chi_i(t)=\begin{cases} 1, & \mu_i(t)>1, \\ 0, & 0<\mu_i(t)\leq 1. \end{cases} \quad (14)$$

综上,用户节点在进行信息传播过程中的实际主动影响力函数 $\widehat{w_i^+}(t)$ 如下:

①信息爆发原点的实际主动影响力函数

$$\widehat{w_i^+}(t)=w_i^+(t)=\text{LR}(v_i) \cdot \left(\sum_{l=1}^m \delta_l(t-t_0)+1\right); \quad (15)$$

②非信息爆发原点的实际主动影响力函数 $\widehat{w_i^+}(t)=\chi_i(t)\mu_i(t)w_i^+(t)$, 即

$$\widehat{w_i^+}(t)=\chi_i(t)\mu_i(t)\text{LR}(v_i) \cdot \left(\sum_{l=1}^m \delta_l(t-t_0)+1\right). \quad (16)$$

2.3 信息传播过程

自媒体时代的社交网络,某一重磅信息的爆发往往是从某一个具有高度对外影响力的官媒、自媒体博主等开始.基于现实情景,描述信息传播过程中的 3 个关键环节.

1)信息爆发

针对某一特定信息,全社交网络都处于 S 节点状态.

假定某信息自 $t=t_0$ 时刻从一个网络节点开始传播, 该节点为信息爆发原点, 其节点状态转变为 I, 并开始向下游节点传播信息, 其实际主动影响力函数如公式(15)所示.

2) 接触传播

下游节点在上游节点发出信息后的下一时刻接收到该信息, 并会在一定程度上受到该信息的影响, 受到影响的程度用被动影响度函数进行衡量, 即:

$$w_i^-(t) = \sum_{v_k \in N^-(v_i)} \widehat{w_k^+(t-1)} \cdot r_{ki}. \quad (17)$$

大量网络节点对该信息产生兴趣, 在浏览后产生了向下游网络节点传播信息的倾向, 即 $\chi_i(t) = 1$, 这些网络节点的状态转化为 E.

事实上, 即使尖叫效应极易引发公众舆论和用户关注, 但并不意味着所有的用户都会浏览该信息, 或在浏览后传播该信息. 这通常是因为: 信息本身并非处于用户的兴趣领域, 或传播信息的用户并没有太大的影响力, 不易使人信服. 对于这些用户, $\chi_i(t) = 0$, 它们仍处于 S 节点状态.

随着舆论的不断演化, 该信息可能会出现反转或多次爆发, 某些用户即使一开始在信息传播时保持缄默, 但在多次尖叫效应下, 兴趣和关注度逐渐累积, 并在不久后成为信息传播过程的参与者, 即:

$$w_i^-(t) = \sum_{v_k \in N^-(v_i)} \widehat{w_k^+(t-1)} \cdot r_{ki} + e^{-\lambda} w_i^-(t-1), \quad (18)$$

其中, λ 是衰减系数. 公式(18)说明被动影响度是可以被动态累加的, 即用户再次接收到该信息时, 会再次受到该信息的影响, 并与之前受到的影响共同累积, 且多次累加的被动影响度函数值有可能达到用户的兴趣阈值. 但由于之前受到的影响会随着时间的推移而减弱, 因此在累加时赋予一定的衰减因子, 即 $e^{-\lambda}$.

3) 信息潜伏

有潜在信息传播倾向的用户, 或许会由于网络下线、信息甄别等原因, 并非立刻向外传播信息, 即处于状态 E 的网络节点会在随机时长的潜伏期后, 转化为 I 状态节点.

E 节点状态下的用户在某一时刻只有传播和不传播两种行为反应, 记用户在单位时间内进行传播的概率为 p , 则不传播信息的概率为 $1-p$. 记用户由 E 状态转化为 I 状态经过的时长为随机变量 T , 其服从参数为 p 的几何分布. 由几何分布的性质可知, $P(T=t) = (1-p)^{t-1} p$, $E[T] = 1/p$. 则刚转化为 E 状态的用户节点, 平均会在时长 $1/p$ 后向下游节点传播信息.

经过以上对信息传播的几个关键环节的分析, 整个信息传播的流程如下.

网络初始时全部节点的状态都是易感节点 S. 当信息开始在网络中出现时, 信息爆发原点出现, 其节点状态转化为感染节点 I, 并开始向下游节点传播信息; 在不同时段下, 下游网络节点依据自身被影响程度决定保持缄默或参与传播, 并在随机时间 T 的潜伏后向更下游节点传播该信息, 在完成传播后短期内不再参与该信息的传播. 更下游节点不断重复接触与传播操作, 直到遍及整个网络或信息脉冲消失, 如图 2 所示.

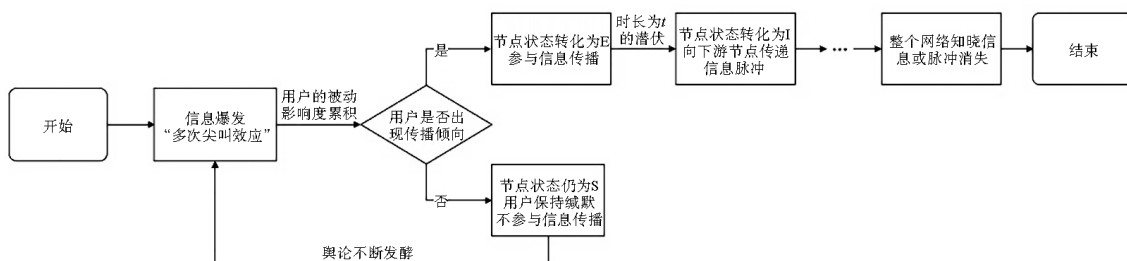


图 2 信息传播流程

对于特定信息, 根据网络节点分类, 单个网络用户在社交网络中的表现如下:

1) 若该网络用户为信息爆发原点, 其状态在信息出现时立刻转化为感染节点 I, 并向下游节点传

递信息,下游节点逐渐转化为 E 节点与 I 节点,或保持 S 节点状态不变.若某一时刻其下游节点已无 S 节点,该网络用户转变为 R 节点,即已完成信息传播过程.

2)若该网络用户为非信息爆发原点,自某一时刻开始,若接触到从其上游节点传播而来的信息并不断累积被动影响度,当超过阈值后,该网络用户转化为 E 节点,并在潜伏期后转化为 I 节点,并开始向下游传递信息;下游节点逐渐转化为 E 节点与 I 节点,或保持 S 节点状态不变,若该网络用户会在下游节点已无 S 节点时转变为 R 节点,或者自信息出现到被埋没整个过程中,该网络用户一直没接收到上游节点传播而来的信息,则其会一直保持 S 节点状态不变,即自始至终不参与信息传播.

3 信息控制策略

3.1 提高兴趣阈值

一个好的社交网络秩序离不开全体用户的维护,需要大多数用户形成良好的信息甄别意识,具有独立思考判断的能力,不盲从他人观点,自觉抵制不良信息.当接触到不良信息时,尤其在尖叫效应的情形发生时,用户不会轻易对不良信息产生兴趣,进而衍生出向下游用户传播该信息的想法,即对此类信息不敏感,兴趣阈值较高.当信息爆发原点传来信息时,绝大多数下游节点的被动影响度函数值低于兴趣阈值,即 $0 < \mu_i(t) < 1$,由 $\chi_i(t)$ 函数的取值范围可知, $\chi_i(t) = 0$,由公式(15)得,实际主动影响力函数 $\widehat{w_i^+}(t) = 0$.不良信息的传播路径受到阻断,如图 3 所示.

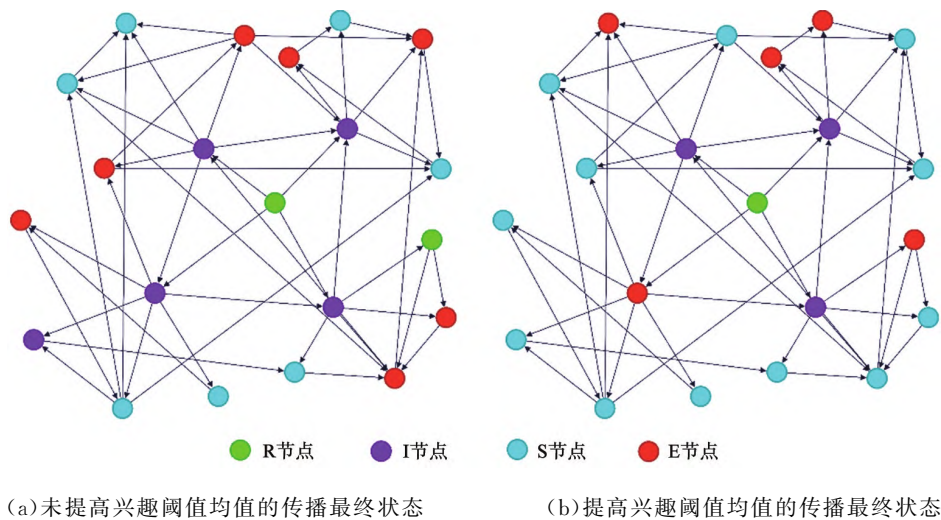


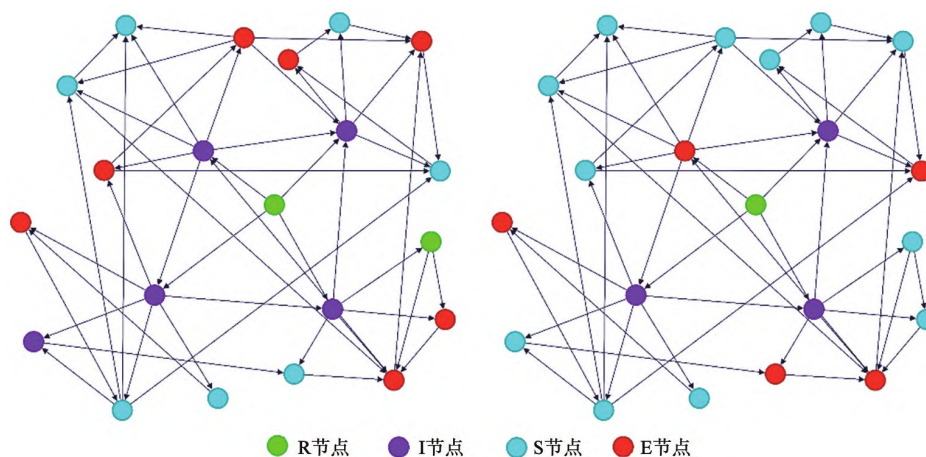
图 3 原始传播与提高兴趣阈值均值后的传播情况对比示意图

提高兴趣阈值,使 S 节点倾向于保持节点状态不变的能力更强,进而遏止不良信息的传播.图 3(b)蓝色 S 节点的数量显著增多,其中诸多 S 节点对应在图 3(a)中为 I 节点或 E 节点,由于 S 节点不具备对外传播信息的功能,使得本应出现的信息传播路径遭到阻断,从而阻止了信息在社交网络中的蔓延.

3.2 降低脉冲强度

可以通过降低“尖叫效应”的核心脉冲函数的强度,降低信息的传播扩散速度,减少信息的传播路径,以延缓信息在社交网络中的蔓延.

当脉冲函数族中脉冲强度 α_i 降低时,由公式(15)可知,信息爆发原点的实际主动影响力 $\widehat{w_i^+}(t)$ 会减弱;由公式(18)可知,其下游节点的被动影响度函数值的增长速度会减慢, $\mu_i(t)$ 的值降低;由公式(16)可知,下游节点实际主动影响力函数 $\widehat{w_i^+}(t)$ 的值降低,因而信息会在传播中被不断延缓,信息无法在全体用户中广泛传播,如图 4 所示.



(a) 未降低脉冲函数强度的传播最终状态 (b) 降低脉冲函数强度的传播最终状态

图 4 原始传播与降低脉冲函数强度后的传播情况对比示意图

降低脉冲强度,使得 S 节点的兴趣积累速度减慢,与先前相比需要经过更长时间才能获得信息传播倾向,甚至一直达不到其阈值.图 4(b)蓝色 S 节点的数量显著增多,其中诸多 S 节点对应在图 4(a)中为 I 节点或 E 节点,由于 S 节点不具备对外传播信息的功能,且新的信息传播路径尚未生成,从而延缓了信息在社交网络中的蔓延.

3.3 屏蔽特定用户

前面两种方式都是从不良信息与尖叫效应本身出发,寻找一种主动进攻式方法阻断社交网络中尖叫效应的外溢,另外也可以通过被动式方法弱化信息传播过程.

定义 6 (k -核)^[10] 在有向图 G 中, H 是 G 的一个极大连通导出子图,如果对于图中的任意一个节点 v , $d_v^+ + d_v^- \geq k$, 则称 H 是 G 的一个 k -核.

在现实的社交网络里,自媒体大 V 节点对外传播信息时,其下游节点中很多是其粉丝.当大 V 节点成为信息爆发原点时,在尖叫效应爆发之初,粉丝们往往在助推信息扩散中发挥了重要作用.大 V 节点与下游 k 个粉丝节点相连, k 个粉丝节点也联系密切,大 V 节点与这 k 个粉丝们共同组成了一个 k -核.

由于现实中的社交网络是一种无标度网络,集群分布、幂律度分布的现象使得少数网络节点在信息传播的过程中发挥了至关重要的作用.它们往往是连接不同社群的“桥”节点,或节点度极大的核心节点.在尖叫效应爆发之初,此类节点会将信息快速传播至所在的 k -核中,继而造成信息的快速传播.因此,当系统检测到有不良信息开始引发尖叫效应并大范围传播时,针对此信息,选取特定节点群定向进行“限流”操作,是一种控制舆论的有效方式,即将一些 k -核设置为 R 节点状态,使其无法接收到该类信息,相当于将 k -核从社交网络中剔除,不良信息无法到达部分用户社群或无法从小社群中传播进入整个社交网络.

4 仿真结果与分析

仿真使用的社交网络结构来自于斯坦福大学的 SNAP(Stanford Network Analysis Platform)项目中的社交网络数据集^[4].对其进行了结构预处理,预处理后的社交网络包含 324 个网络节点和 5028 条有向边.

4.1 尖叫效应形成过程

初始设定下,特定信息从出现到被淹没共经历了 7 次大大小小的舆论浪潮,即脉冲函数族共有 $m=7$ 支,脉冲函数族强度 α_i 的最大值为 3,时间区间半径 $\epsilon_i (i=1, 2, \dots, m)$ 均为 3.社交网络所有用户的兴趣阈值 $\overline{w_i}$ 的均值为 0.6.对节点按照点度的大小进行排序,从前 10% 的点中随机选择一点作为信息爆发原点,且每次选择信息爆发原点时用户 i 对用户 j 的影响力 $w_{ij} \in [0, 1]$ 按照均匀分布随机选取.多次仿真的结果表明,随机选择对信息传播过程中各类节点的比例变化影响不大.

在不加限制的条件下,从尖叫效应爆发开始,整个社交网络中的各类人群比例变化图和各时间点增长的传播信息人群占比变化图分别如图 5 和图 6 所示。

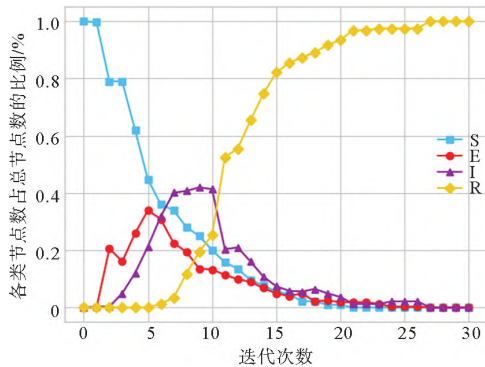


图 5 各类人群比例变化图

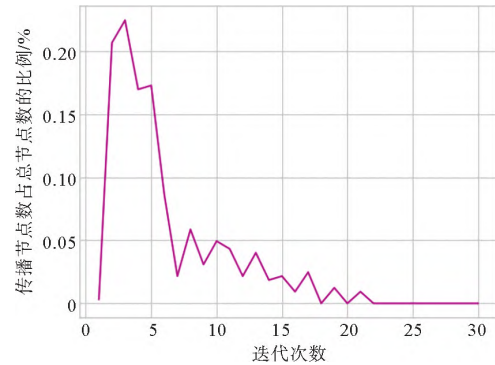


图 6 各时间传播信息人群占比变化图

由图 5 可知,整个网络中的用户节点最初全部为易感节点(S 节点状态),当尖叫效应开始爆发后,潜伏节点和传播节点的数量开始快速上升,并先后达到了最大值.其中,在某一时刻全社交网络中有大约 50%的用户都在进行对外信息传播,达到了传播节点在网络中占比的最大值,此时也是信息传播最为迅速的时期.但也正是从此开始,完成信息传播并退出信息传播过程的用户开始逐渐增加,直到最后全部用户几乎都参与过信息传播的历程,说明信息传播的热度在经历鼎盛期后逐渐开始消退,直到最后被完全淹没。

由图 6 可知,每当有信息脉冲式在社交网络中开始传播时,参与信息传播的人群(E 节点+I 节点)的比例都会出现相应的“波峰”.但随着信息在社交网络中逐步传播,“波峰”的强度会逐渐下降,直到随着脉冲函数族的消失而消失。

4.2 信息控制策略

分别应用 3 种不同的策略控制尖叫效应在社交网络中的爆发,基于控制方法调整参数取值,测试在不同取值条件下的网络信息传播效果,并以 R 状态节点的占比作为衡量指标。

1) 提高兴趣阈值

在其他条件不变的情况下,将初始设定为 0.6 的阈值不断提升,结果如图 7 所示。

由图 7 可知,如果社交网络用户兴趣阈值的均值提高,信息在社交网络中的爆发就有了一定的延迟,这是因为用户不会轻易地对信息感兴趣,不会不假思索地向外传播信息.阈值的均值提高后,信息最终并不能完全蔓延到整个社交网络,相当一部分网络用户不参与到信息传播过程中,尖叫效应受到一定的遏制。

2) 降低脉冲强度

在其他条件不变的情况下,将初始设定为 3 的最大脉冲强度不断削弱,如图 8 所示。

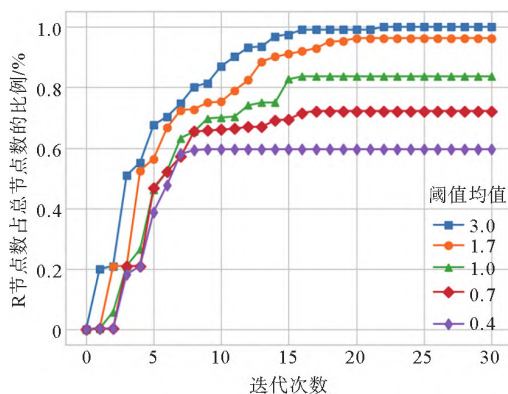


图 7 不同兴趣阈值下信息传播效果对比图

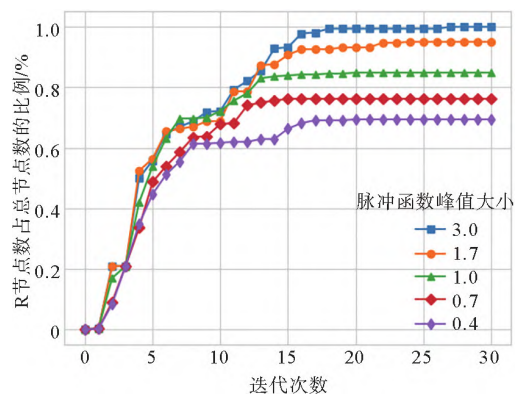


图 8 削弱脉冲函数强度下信息传播效果对比图

由图 8 可知,通过削弱脉冲的最大强度,信息在社交网络中的爆发就有了一定的延迟,这是因为用户即使向外传播信息,其传播力度也会降低.最大强度调低后,信息最终并不能完全蔓延到整个社交网络,相当一部分网络用户不参与到信息传播过程中,尖叫效应受到一定的遏制.

3)屏蔽特定用户

在其他条件不变的情况下,将不同 k 值的 k -核从社交网络中删除,如图 9 所示.

由图 9 可知,与前两种控制策略显著不同,删除 k -核后不仅能够有效降低信息传播速度,且对信息传播的控制效果也极为明显.删除的 k -核的 k 值越低,就更多地屏蔽了一些传播能力强的用户节点,对信息传播的遏止能力也就越强.

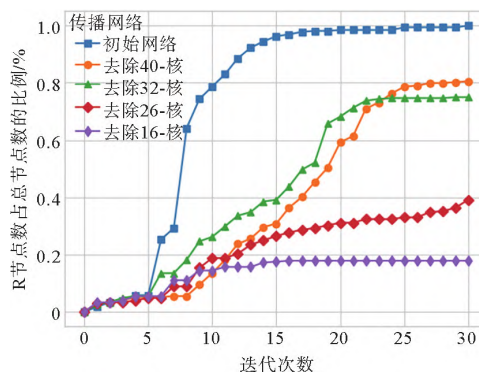


图9 去除 k -核后的信息传播图

5 结论

基于社交网络,利用建立的一系列网络指标,借鉴 SEIR 模型,引入脉冲函数族研究尖叫效应的形成过程,刻画了信息传播中各类人群动态变化过程及变化趋势;研究了控制或消除尖叫效应的策略,从多个角度提出了控制和消除尖叫效应的方法.

最后,基于真实的社交网络结构,随机选择信息爆发原点仿真模拟了尖叫效应机制的形成及演变的过程,验证了本文提出的控制或消除尖叫效应的策略.多次仿真结果表明,本文建立的尖叫效应机制的形成及演变的机理模型是与实际吻合的、合理有效的;控制或消除尖叫效应的策略是实际可行的.在实践中,可以使用其中一种信息控制策略,也可以多种信息控制策略组合使用以取得更快更好的效果.

参考文献

- [1] 深圳市尚龙教学技术与交叉学科产业化研发中心. 2022 年“深圳杯”数学建模挑战赛 A 题——破除“尖叫效应”与“回声室效应”,走出“信息茧房”[Z/OL]. [2022-07-23]. <http://www.m2ct.org/view-page.jsp?editId=12&uri=0D00263&gobackUrl=modular-list.jsp&pageType=smxly&menuType=flowUp1>.
- [2] 韩中庚. 数学建模方法及其应用[M]. 北京: 高等教育出版社, 2005.
- [3] 司守奎, 孙玺. Python 数学实验与建模[M]. 北京: 科学出版社, 2020.
- [4] 张国防, 徐建民. 积极交互作用视角下网络信息传播及节点影响力排序[J]. 现代情报, 2023, 43(5): 90-103.
- [5] 周飞, 高茂庭. 基于 PageRank 的网络社区意见领袖发现算法[J]. 计算机工程, 2018, 44(2): 203-209+219.
- [6] 宋会敏. 社区网络中信息传播控制技术研究[D]. 沈阳: 沈阳航空航天大学, 2017.
- [7] 杜鸿毅, 杨华, 刘艳红, 等. 基于网络媒体的非线性动力学信息传播模型[J]. 计算机科学, 2022, 49(S1): 280-284.
- [8] 顾亦然, 朱梓嫣. 基于 LeaderRank 和节点相似度的复杂网络重要节点排序算法[J]. 电子科技大学学报, 2017, 46(2): 441-448.
- [9] Batagelj V, Zaversnik M. An $O(m)$ algorithm for cores decomposition of networks: American, 10.1007/BF01074693 [P]. 2003-10-25.
- [10] Stanford Network Analysis Project. Stanford large network dataset collection, social networks[Z/OL]. [2023-04-15]. <https://snap.stanford.edu/data/ego-Facebook.html>.

(下转第 106 页)

Seventy Years of Scheduling Models

LIN Ling¹, TAN Zhiyi²

(1. School of Computer and Computing Science, Hangzhou City University, Hangzhou, Zhejiang 310015, China;
2. School of Mathematical Science, Zhejiang University, Hangzhou, Zhejiang 310027, China)

Abstract: Scheduling is an important class of problems in the field of operations research and combinatorial optimization. This article outlines the basic framework, common terminology, and main types of scheduling models. It provides an overview of the origins and a brief developmental history of scheduling research, and introduces the mainstream research paradigms and typical extensions of scheduling models.

Key words: scheduling; combinatorial optimization; operations research

作者简介

林 凌(1980—), 女, 博士, 讲师, 主要研究方向为运筹学组合优化、排序与调度。

谈之奕(1975—), 男, 博士, 教授, 主要研究方向为运筹学组合优化、算法设计与分析、算法博弈论。

(上接第 72 页)

Research of the Propagation Mechanism of Scream Effect and its Simulation

JIANG Zhengxu¹, YAO Jiaqi¹, MENG Lingjun², WANG Jian², CAO Yongchang²

(1. Haide College, Ocean University of China, Qingdao, Shandong 266071, China;
2. School of Mathematics and Science, Ocean University of China, Qingdao, Shandong 266071, China)

Abstract: Based on the social network model, a series of network indicators such as edge weight bias factor, active influence function, passive influence function and interest threshold are established. Based on the SEIR model, the pulse function family is introduced to study the propagation mechanism of the "Scream Effect", and the dynamic change process and change trend of various groups of people in information dissemination are described. Strategies to control or eliminate the "Scream Effect" such as increasing the interest threshold, reducing the pulse intensity and shielding specific users are proposed. Based on the real social network structure data, the propagation mechanism of the "Scream Effect" is simulated, and the validity of the proposed strategy of controlling or eliminating the "Scream Effect" is verified.

Key words: social network; scream effect; pulse function; SEIR model; rumor blocking

作者简介

姜政旭(2003—), 男, 中国海洋大学 2021 级本科生。

姚嘉祺(2002—), 男, 中国海洋大学 2021 级本科生。

孟令钧(2003—), 女, 中国海洋大学 2021 级本科生。

王 建(1980—), 男, 博士, 副教授, 主要研究方向为偏微分方程。

曹永昌(1983—), 男, 博士, 讲师, 主要研究方向为优化理论和算法。