



基于时变阻尼运动的社交网络信息传播动力学分析

刘小洋^{1*}, 何道兵¹, 刘超¹, 张宜浩^{2,3}

1. 重庆理工大学计算机科学与工程学院, 重庆 400054, 中国

2. 重庆理工大学人工智能学院, 重庆 401135, 中国

3. School of Computing, National University of Singapore, Singapore 119077, Singapore

* 通信作者. E-mail: lxy3103@cqut.edu.cn

收稿日期: 2020-05-21; 修回日期: 2020-07-29; 接受日期: 2020-10-13; 网络出版日期: 2021-11-12

国家社科基金 (批准号: 17XXW004) 和国家自然科学基金 (批准号: 61702063) 资助项目

摘要 针对目前社交网络上信息传播动力学的研究成果中, 未能从社会物理学层面揭示微观个体上的随机性和无序性与宏观群体上的可知性和可控性之间的关系问题, 本文提出了一种基于时变阻尼运动的信息传播动力学模型 (time-varying damped motion, TVDM). 首先, 深入分析社交网络中信息传播现象, 从时间、空间、主体 (行为者)、客体 (信息) 4 个维度提取关键特征; 然后, 从带阻尼的简谐运动视角、内外力系合力作用视角、信息能量转化视角、时变系统与信号变换视角, 揭示了社交网络上信息传播的物理学本质规则和作用机制; 最后, 通过数值模拟实验验证了网络上个体行为服从统计规律性. 实验结果表明: 提出的时变阻尼运动信息传播动力学模型精度为 91%, 构建的社交网络信息传播动力学模型 TVDM 是合理有效的.

关键词 社交网络, 动力学, 信息能量, 阻尼运动, 时变系统

1 引言

随着 5G 通信技术的普及和自媒体的不断发展, 以 Twitter、Facebook、微博为代表的在线社交网络^[1,2] 迅猛发展, 成为人类社会社会中社会关系维系和信息传播的重要渠道与载体. 它一方面极大地改变了人们的生活方式, 由简单信息搜索和网页浏览转向网上社会关系的构建与维护, 基于社会关系的信息创造、交流和共享; 另一方面, 也给社会带来了负面影响, 例如, 信息传播、虚假信息、网络诈骗等. 因此, 开展在线社交网络分析及网络信息传播的基础理论研究, 有着重要的应用价值和社会意义.

信息传播动力学研究是在线社交网络上信息传播过程研究的核心问题之一^[3,4]. 由于信息传播过程一般不能用线性系统表述, 所以当前大部分研究都是用非线性动力学方法来进行建模. 动力学的核心使命是预测系统的变化, 非线性动力学在这点上也是一样的. 根据一个系统过去的历史数据, 结合

引用格式: 刘小洋, 何道兵, 刘超, 等. 基于时变阻尼运动的社交网络信息传播动力学分析. 中国科学: 信息科学, 2021, 51: 1867–1884, doi: 10.1360/SSI-2020-0142
Liu X Y, He D B, Liu C, et al. Social network information propagation dynamic analysis based on time-varying damping motion (in Chinese). Sci Sin Inform, 2021, 51: 1867–1884, doi: 10.1360/SSI-2020-0142

当前状态信息, 预测系统的未来, 需要知道它在微小时间尺度里的性质并列动力学方程. 目前, 国内外学者所构建的信息传播动力学模型, 根据侧重点的不同, 主要有基于传染病模型、博弈论模型和物理系统模型 3 类方法^[5,6].

基于传染病模型是信息传播领域公认比较成熟的模型. 它将人群分为易感者 S (susceptible)、感染者 I (infected) 和治愈者 R (recovered), S 状态转变到 R 状态的过程中可能存在潜伏期状态 E (exposed), 于是产生了 SEIR (susceptible exposed infected recovered) 模型. 该类模型发展过程中, 近几年随着应用场景的不同出现许多变种, 比如具有两个时滞和垂直转移的 SEIRS (susceptible exposed infected recovered susceptible) 模型、基于用户情绪的 SISa (susceptible infected susceptible automatic) 模型等^[7,8]. 现阶段也有利用博弈论 (game theory) 和进化博弈模型等来研究动态社交网络中的不同类型的用户行为^[9,10], 此类模型中的节点通过观点交互形成和更新自身观点. 在社交演化博弈中, 每个个体都有两个关注点, 即短期效用和长期声誉, 相对应存在着两种更新机制, 即为了获得更好效用的策略更新和更好声誉环境的社交关系更新^[11]; 基于物理系统模型^[12,13] 是分析在线社交网络用户特征和网络属性, 借用物理学中有关理论和方法, 模拟在线社交网络中社交力量对信息传播的推动过程^[14,15].

通过对以上最新研究成果的系统分析和归纳总结发现, 基于传染病模型属于宏观层面运用平均场理论建模, 忽略用户个体特征, 不关注具体用户行为, 只能反映信息传播的大致趋势, 模型精度并不高. 博弈论模型重视个体间交互行为, 注重用户在互动中的动机, 具体采用的策略简单划分为背叛与合作, 相应带来不同的收益, 这种划分过于武断, 而且人的行为具有不确定性, 模型较难准确模拟真实传播过程. 基于物理系统的模型是在深入研究社会物理现象后, 借鉴经典物理理论和方法, 构建出信息传播模型, 它能从某个方面较为真实地反映信息传播的客观规律和作用机制. 但是, 目前基于物理运行系统对信息传播问题的研究, 特征提取过于简单, 忽略很多比较重要的因素, 并没有系统地、深层地分析和解决此类问题, 模型精度较低, 适合性不强.

鉴于以上分析结果, 解决在线社交网络上信息传播类问题还需要从人、物和环境角度和系统观点, 分析更加客观、更加深层次的物理过程, 从而诠释信息传播作用机制. 本文在深入分析在线社交网络中信息传播物理现象的基础上, 首先从时间、空间、行为人和信息 4 个维度上提取特征参量, 进行形式化定义和归一化处理, 全方位多视角地剖析传播活动. 然后借鉴运动学、动力学和系统论知识, 对信息传播过程从信息传播速度及加速度、内外受力和信息能量守恒方面进行研究, 构建出基于时变阻尼运动的信息传播动力学模型 TVDM. 最后通过一系列实验, 与真实数据进行对比, 检验了本文模型的有效性和精确性.

2 基于时变阻尼运动的信息传播动力学模型 TVDM

2.1 个体特征分析

基于传播学和复杂网络理论对在线社交网络进行分析, 从时间、空间、行为人和信息 4 个维度提取出影响信息传播的基本特征, 如图 1 所示. 下面逐一对每个特征进行解释和描述.

(1) 在线频率. 表征用户登录在线社交网络的时间行为习惯, 用在线角频率 ω 定量刻画. 如果用户每天有 n 小时在线, 那么在线角频率可表示为

$$\omega = \frac{2\pi n}{24} = \frac{\pi n}{12}. \quad (1)$$

(2) 上线时差. 表征用户登录在线社交网络的时间先后关系, 用初相位 φ 表示. 在图 2 中, 用户 A 和 B 都向用户 C 转发一条相同的信息, 但是从时间序列上看, 用户 C 与用户 A 同相, 与用户 B 反相,

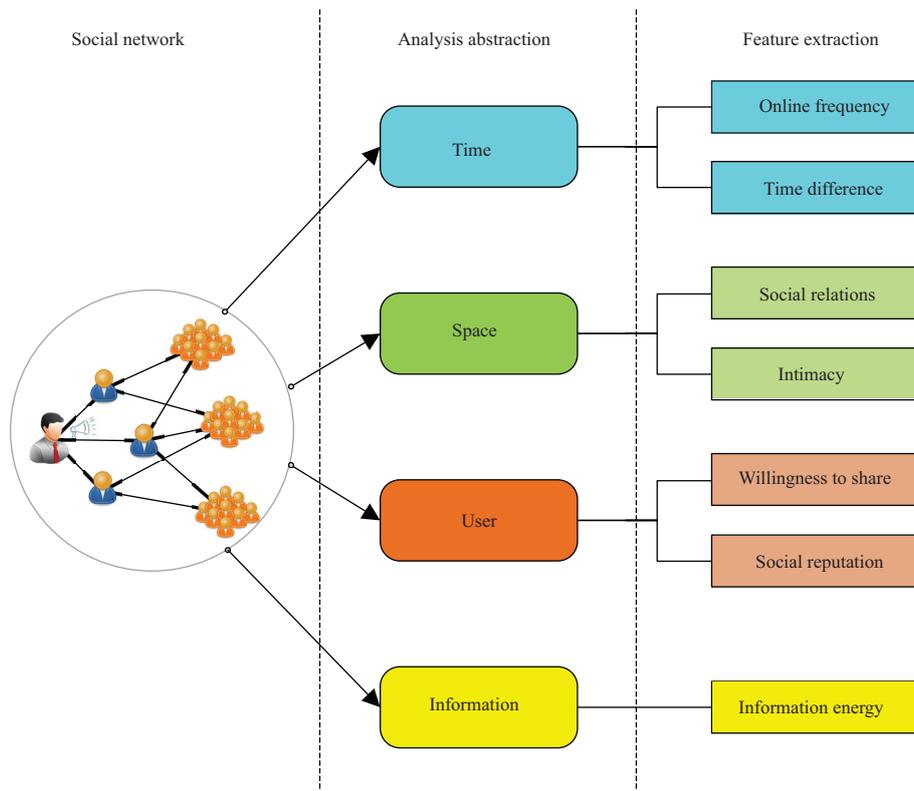


图 1 (网络版彩图) 特征提取图

Figure 1 (Color online) Feature extraction diagram

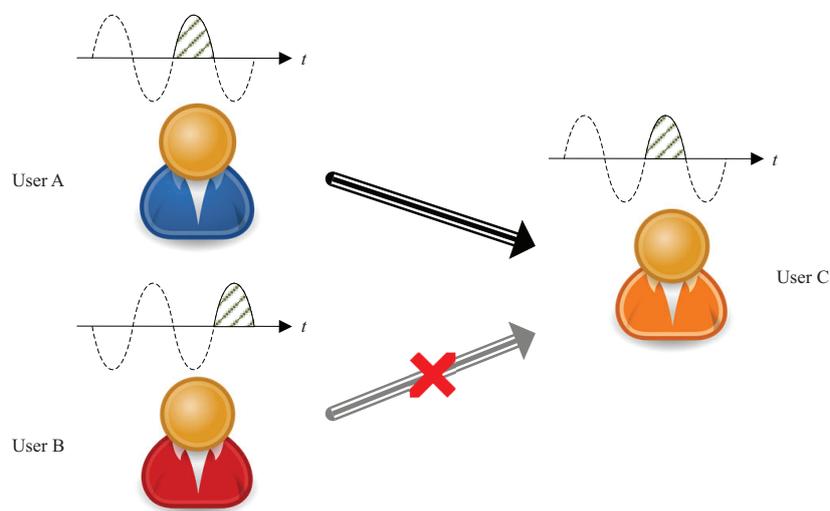


图 2 (网络版彩图) 在线时差影响示意图

Figure 2 (Color online) Schematic diagram of the impact of online time difference

以至于用户 C 能及时收到用户 A 发来的信息, 而用户 B 的信息被忽略.

(3) 社交关系. 表示用户在社交网络中的全局重要性程度, 用节点质量 m 进行度量. 为了避免小

表 1 特征描述表
Table 1 Characterization table

No.	Feature name	Variable name	Sign	Meaning
1	Online frequency	Online angular frequency	ω	Online time rule
2	Online time difference	Initial phase	φ	The difference in online time
3	Social relation	Node quality	m	User network weight
4	Intimate degree	Close weight	w	Group closeness
5	Share willingness	Sharing coefficient	S	Communication initiative
6	Social stigma	Reputation quotient	R	User social influence
7	Information energy	Information potential energy, kinetic energy	E_K, E_P	Information content and dissemination value

度数节点被忽略, 因此对节点度数据进行对数平滑处理. 即

$$m = \lg(1 + d), \tag{2}$$

其中 d 为节点度.

(4) 亲密程度. 表示用户在社交关系网中的局部重要性程度, 反映在线社交网络的群体特性. 用亲密密度权重 ω 来刻画两个用户之间的这种信息取信程度, 在社交网络上, 绝大多数用户间的亲密密度很低, ω 在 $(0, 1)$ 上服从参数为 λ 的指数分布, 即 $\omega \sim e(\lambda)$.

(5) 分享意愿. 描述用户面向其他用户进行信息发布、信息转发的主观能动性大小, 用分享系数 S 予以刻画. 它表示用户间一对多的群体交互信息行为, 具有方向性. S 受到许多相互独立的随机因素的影响, 比如用户的性格、忙闲状态、心情等, 总的影响可以看作是服从正态分布的, 即 $S \sim N(\mu_1, \sigma_1^2)$.

(6) 社会声誉. 一个用户的社会声誉是指其他用户对该用户的综合评价, 社交网络中具有良好声誉的用户总是会受到其他用户的欢迎. 用声誉系数 R 进行表示, 社交网络上声誉很高和声誉很差的用户都占极少数, 大多处于中等水平, 整体上 R 服从正态分布, 即 $S \sim N(\mu_2, \sigma_2^2)$.

(7) 信息能量 E . 信息之所以能够在网络上传播, 是因为它具有能量, 能量又可分为信息势能 E_P 和信息动能 E_K 两种. 在数值上, 信息能量等于信息势能与信息动能之和, 即

$$E = E_P + E_K. \tag{3}$$

为了方便阅读和理解, 将上述 7 个特征描述与符号表示进行汇总, 如表 1 所示.

2.2 信息传播运动分析

在一条传播路径上, 每个用户上线后获知上一用户传递而来的信息, 解读后转发给下一用户, 此过程迭代进行, 形成一条传播链. 后面用户获知信息的时延取决于前面用户的在线频率和上线时差. 如图 3 所示, 假若存在 A, B, C, D 4 个用户, 具有相同的特征, 用户 A 转发一条信息能量为 E 的信息, 它究竟能传播多远.

其物理过程为: (1) 用户 A 在 T_1 周期内上线期间向邻居节点传播扩散一条信息; (2) 用户 B 在 T_2 周期内上线, 阅读并向自己的邻居节点转发该信息; (3) 类似地, 用户 C 在 T_3 周期内上线, 阅读并转发该信息; (4) 用户 D 在 T_4 周期内上线阅读该信息, 但不再转发该传播链上的信息传播活动结果. 由此可见, 信息传播过程与信息本身的能量和用户特征的分享意愿、社会声誉, 以及用户行为中的在线

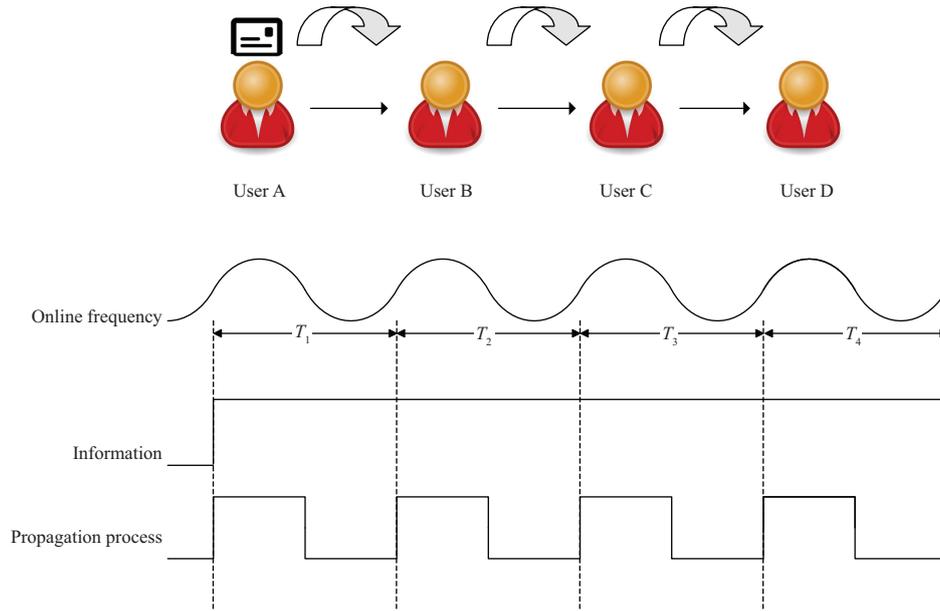


图 3 (网络版彩图) 单路径上信息传播示意图
 Figure 3 (Color online) Feature extraction diagram

频率、上线时差关系密切, 阶跃性地向前传播. 结合式 (1) 可将传播距离表示为

$$\begin{aligned}
 x &= \sqrt{ESR} \sin(\omega t + \varphi) r(\omega t + \varphi) \\
 &= \sqrt{ESR} \sin\left(\frac{\pi n}{12} t + \varphi\right) r\left(\frac{\pi n}{12} t + \varphi\right),
 \end{aligned} \tag{4}$$

其中, x 为传播距离; E 为传播能量; S 为分享系数; R 为声誉系数; ω 为在 $r(t)$ 为斜坡函数, 表示节点在一个周期内解读信息的时间越长, 对信息理解越深刻, 之后评论和转发的信息越权威, 信息传播距离也就越远. $r(t)$ 在最小正周期内的定义为

$$r(t) = \begin{cases} t, & 0 \leq t < \pi, \\ 0, & t \geq \pi. \end{cases} \tag{5}$$

将式 (4) 对时间求一阶导数, 可得到信息传播的速度 v 为

$$\begin{aligned}
 v &= \frac{dx}{dt} = \omega \sqrt{ESR} \cos(\omega t + \varphi) r(\omega t + \varphi) + \omega \sqrt{ESR} \sin(\omega t + \varphi) u(\omega t + \varphi) \\
 &= \omega \sqrt{ESR} [\sin(\omega t + \varphi) u(\omega t + \varphi) + \cos(\omega t + \varphi) r(\omega t + \varphi)] \\
 &= \frac{\pi n}{12} \sqrt{ESR} \left[\sin\left(\frac{\pi n}{12} t + \varphi\right) u\left(\frac{\pi n}{12} t + \varphi\right) + \cos\left(\frac{\pi n}{12} t + \varphi\right) r\left(\frac{\pi n}{12} t + \varphi\right) \right],
 \end{aligned} \tag{6}$$

其中, 函数 $u(t)$ 为阶跃函数, 在最小正周期内定义为

$$u(t) = \frac{dr(t)}{dt} = \begin{cases} 1, & 0 \leq t < \pi, \\ 0, & t \geq \pi. \end{cases} \tag{7}$$

再将式 (9) 对时间求二阶导数, 可得到信息传播的加速度 a 为

$$a = \frac{d^2x}{dt^2} = \omega \sqrt{ESR} [\omega \cos(\omega t + \varphi) u(\omega t + \varphi) + \omega \sin(\omega t + \varphi) \delta(\omega t + \varphi)]$$

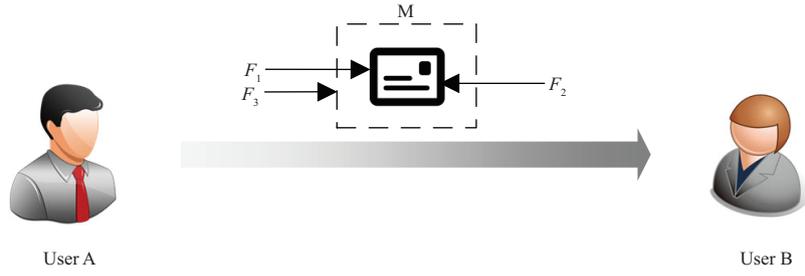


图 4 (网络版彩图) 信息受力分析图
Figure 4 (Color online) Information stress analysis diagram

$$\begin{aligned}
 & -\omega \sin(\omega t + \varphi)r(\omega t + \varphi) + \omega \cos(\omega t + \varphi)u(\omega t + \varphi)] \\
 & = \omega^2 \sqrt{ESR} \{ \sin(\omega t + \varphi)[\delta(\omega t + \varphi) - r(\omega t + \varphi)] + 2 \cos(\omega t + \varphi)u(\omega t + \varphi) \} \\
 & = \frac{\pi^2 n^2}{144} \sqrt{ESR} \left\{ \sin\left(\frac{\pi n}{12}t + \varphi\right) \left[\delta\left(\frac{\pi n}{12}t + \varphi\right) - r\left(\frac{\pi n}{12}t + \varphi\right) \right] + 2 \cos\left(\frac{\pi n}{12}t + \varphi\right) u\left(\frac{\pi n}{12}t + \varphi\right) \right\}, \quad (8)
 \end{aligned}$$

其中, 函数 $\delta(t)$ 为单位冲激函数, 其定义为

$$\begin{cases} \delta(t) = \frac{du(t)}{dt} = \frac{d^2r(t)}{dt^2}, & t \neq 0, \\ \int_0^{2\pi} \delta(t) = 1. \end{cases} \quad (9)$$

2.3 信息传播受力分析

在在线社交网络上, 信息的传播会同时受到多个力的作用, 它们的合力共同决定信息能否向前传播.

如图 4 所示, 用户 A 向用户 B 发送一条信息 M, M 受到 3 个力的作用, 如下所述. (1) 传播驱动力 F_1 , 描述传播者传递信息能力的大小, 传播驱动力在数值上等于传播者节点质量 m_1 与传播加速度 a 的乘积, 即

$$\begin{aligned}
 F_1 & = m_1 a \\
 & = \frac{\pi^2 n^2}{144} m_1 \sqrt{ESR} \left\{ \sin\left(\frac{\pi n}{12}t + \varphi\right) \left[\delta\left(\frac{\pi n}{12}t + \varphi\right) - r\left(\frac{\pi n}{12}t + \varphi\right) \right] \right. \\
 & \quad \left. + 2 \cos\left(\frac{\pi n}{12}t + \varphi\right) u\left(\frac{\pi n}{12}t + \varphi\right) \right\}, \quad (10)
 \end{aligned}$$

其中, m_1 为用户 A 的节点质量. (2) 受众理性排斥力 F_2 , 描述信息受众对信息进行甄别的能力大小, 它与自身和发送者节点的节点质量、亲密度权重及传播驱动力 F_1 有关, 受众理性排斥力 F_2 可表示为

$$F_2 = F_1 \sqrt{\frac{m_2}{m_1 + m_2} (1 - w)}, \quad (11)$$

其中, m_1, m_2 分别为用户 A、用户 B 的节点质量; w 为用户 A、用户 B 之间的亲密度权重; F_1 为信息传播驱动力. (3) 社会舆论场力 F_3 . 舆论场力大小与政府部门或媒体的节点量和影响因子及信息传播驱动力有关, 关系表达式为

$$F_3 = \text{sgn}(d) \frac{1}{N_1} F_1 \sum_{i=1}^{N_1} m_i f_i, \quad (12)$$

其中, N_1 为关于此信息在社交网络上发表评论意见或进行信息转发的政府部门或媒体节点的总数量; m_i, f_i 为媒体节点质量和对应的影响因子. $\text{sgn}(d)$ 为符号函数, d 表示媒体所持的观点和意见, 根据支持、中立、反对的意见决定其力的作用方向, 其定义如下:

$$\text{sgn}(d) = \begin{cases} 1, & \text{支持,} \\ 0, & \text{中立,} \\ -1, & \text{反对.} \end{cases} \quad (13)$$

由以上讨论可知, 合力 F 可表示为

$$\begin{aligned} F &= F_1 - F_2 + F_3 \\ &= F_1 - F_1 \sqrt{\frac{m_2}{m_1 + m_2}}(1 - w) + \text{sgn}(d) \frac{1}{N_1} F_1 \sum_{i=1}^{N_1} m_i f_i \\ &= F_1 \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}}(1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right] \\ &= \frac{\pi^2 n^2}{144} m_1 \sqrt{ESR} \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}}(1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right] \\ &\quad \times \left\{ \sin\left(\frac{\pi n}{12}t + \varphi\right) \left[\delta\left(\frac{\pi n}{12}t + \varphi\right) - r\left(\frac{\pi n}{12}t + \varphi\right) \right] + 2 \cos\left(\frac{\pi n}{12}t + \varphi\right) u\left(\frac{\pi n}{12}t + \varphi\right) \right\}. \end{aligned} \quad (14)$$

在合力 F 作用下, 信息传播的加速度 a 和速度 v 分别为

$$\begin{aligned} a &= \frac{F}{m_1} \\ &= \frac{\pi^2 n^2}{144} \sqrt{ESR} \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}}(1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right] \\ &\quad \times \left\{ \sin\left(\frac{\pi n}{12}t + \varphi\right) \left[\delta\left(\frac{\pi n}{12}t + \varphi\right) - r\left(\frac{\pi n}{12}t + \varphi\right) \right] + 2 \cos\left(\frac{\pi n}{12}t + \varphi\right) u\left(\frac{\pi n}{12}t + \varphi\right) \right\}, \end{aligned} \quad (15)$$

$$\begin{aligned} v &= \int_0^{\frac{24}{n}} a dt \\ &= \int_0^{\frac{24}{n}} \frac{\pi^2 n^2}{144} \sqrt{ESR} \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}}(1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right] \\ &\quad \times \left\{ \sin\left(\frac{\pi n}{12}t + \varphi\right) \left[\delta\left(\frac{\pi n}{12}t + \varphi\right) - r\left(\frac{\pi n}{12}t + \varphi\right) \right] + 2 \cos\left(\frac{\pi n}{12}t + \varphi\right) u\left(\frac{\pi n}{12}t + \varphi\right) \right\} dt \\ &= \frac{\pi n}{12} \sqrt{ESR} \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}}(1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right] \\ &\quad \times \left[\sin\left(\frac{\pi n}{12}t + \varphi\right) u\left(\frac{\pi n}{12}t + \varphi\right) + \cos\left(\frac{\pi n}{12}t + \varphi\right) r\left(\frac{\pi n}{12}t + \varphi\right) \right]. \end{aligned} \quad (16)$$

2.4 信息传播演化过程

(1) 信息在节点上, 信息处于静止的状态, 此时信息动能 $E_K = 0$, 信息能量表示为

$$E = E_P + E_K = E_P. \quad (17)$$

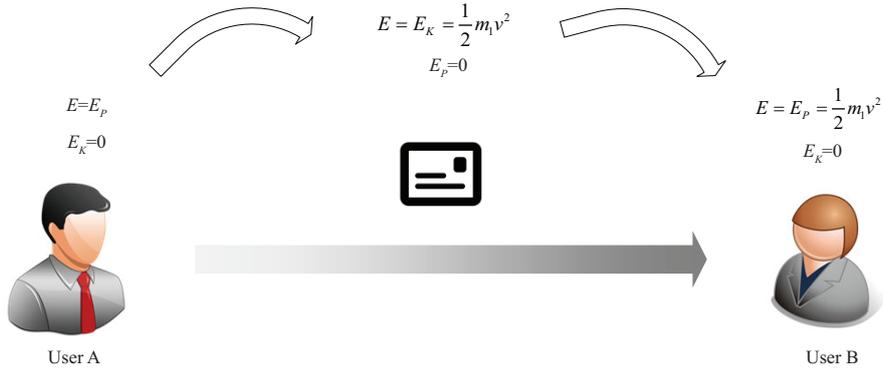


图 5 (网络版彩图) 信息能量转化示意图

Figure 5 (Color online) Schematic diagram of information energy transformation

(2) 信息在两节点之间, 处于传播运动状态, 此时信息势能 $E_P = 0$, 信息能量表示为

$$E = E_P + E_K = E_K. \quad (18)$$

当某一节点拥有一条信息势能 $E_P \neq 0$ 的信息, 由于势能差的影响, 必将向邻居节点进行传播, 使信息进入运动传播状态. 在传播过程中, 信息能量全部转化为信息动能 E_K , 转化关系如下所示:

$$\begin{aligned} E_K &= \frac{1}{2} m_1 v^2 \\ &= \frac{\pi^2 n^2}{288} m_1 ESR \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}} (1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right]^2 \\ &\quad \times \left[\sin\left(\frac{\pi n}{12} t + \varphi\right) u\left(\frac{\pi n}{12} t + \varphi\right) + \cos\left(\frac{\pi n}{12} t + \varphi\right) r\left(\frac{\pi n}{12} t + \varphi\right) \right]^2. \end{aligned} \quad (19)$$

当信息到达接收节点后, 信息进入静止状态, 信息动能全部转化为信息势能 E_P . 两个用户之间信息传递和信息能量转换的关系如图 5 所示.

在传播过程中, 为了克服受众理性排斥力和社会舆论场力的作用, 需要消耗一定的信息能量, 随着传播链的不断延长, 信息能量 E 逐渐减小, 当各条传播路径上信息能量都衰减到足够小 (趋于零) 时, 信息传播活动停止. 整个信息传播过程如图 6 所示.

在图 6 中, t_0 时刻, 1 号节点发布信息; t_1 时刻, 1 号节点执行节点信息传递算法, 将信息传播给 2 号和 10 号节点; t_2 时刻, 2 号节点将信息转发给 3 号节点, 10 号节点将信息转发给 9 号和 11 号节点; t_3 时刻, 信息由 3 号节点转发给 4 号节点, 9 号节点转发给 7 号节点, 11 号节点转发给 5 号和 12 号节点, 此时, 由于 4 条传播路径上的信息能量衰减到非常小, 趋于零值, 因而传播行为结束. 网络中 6 号和 8 号节点一直未接收到信息, 最终信息覆盖率达到 83% (10/12).

3 结果与分析

下面将从数值分析、仿真分析、实证仿真 3 个方面对提出的社交网络信息传播动力学模型进行合理性、有效性验证.

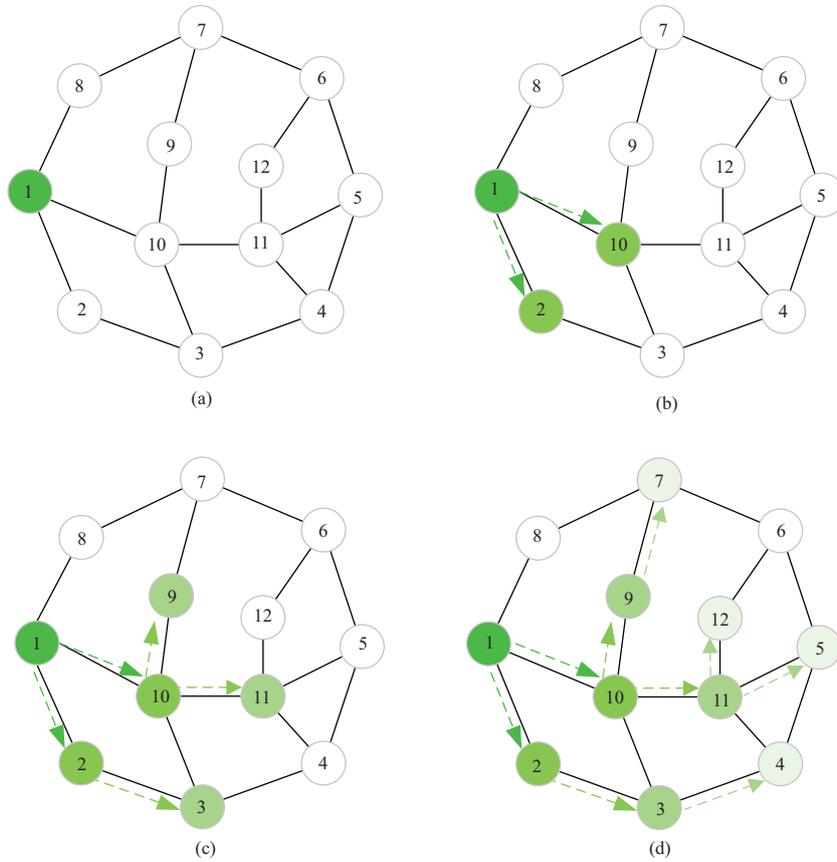


图 6 (网络版彩图) 信息传播过程示意图. (a) $t = t_0$, 信息发布; (b) $t = t_1$, 信息传播; (c) $t = t_2$, 信息转发; (d) $t = t_3$, 传播结束

Figure 6 (Color online) Schematic diagram of information transmission process. (a) $t = t_0$, information release; (b) $t = t_1$, information dissemination; (c) $t = t_2$, information forwarding; (d) $t = t_3$, end of spread

3.1 数值模拟

由信息传播演化过程和节点信息传递算法可见, 信息传播的关键在于按式 (17) ~ (19) 进行信息能量的转化. 假设节点 i 为信息发送节点, 节点 j 为信息接收节点, 令

$$f_1(i) = \frac{\pi^2 m_1}{288} n^2 SR, \tag{20}$$

$$f_2(i, j) = \left[1 - \sqrt{\frac{m_2}{m_1 + m_2}} (1 - w) + \text{sgn}(d) \frac{1}{N_1} \sum_{i=1}^{N_1} m_i f_i \right]^2, \tag{21}$$

$$f_3(i, t) = \left[\sin\left(\frac{\pi n}{12} t + \varphi\right) u\left(\frac{\pi n}{12} t + \varphi\right) + \cos\left(\frac{\pi n}{12} t + \varphi\right) r\left(\frac{\pi n}{12} t + \varphi\right) \right]^2, \tag{22}$$

则式 (19) 可改写为

$$E_K = f_1(i) f_2(i, j) f_3(i, t) E. \tag{23}$$

再令

$$Z(i, j, t) = f_1(i) f_2(i, j) f_3(i, t), \tag{24}$$

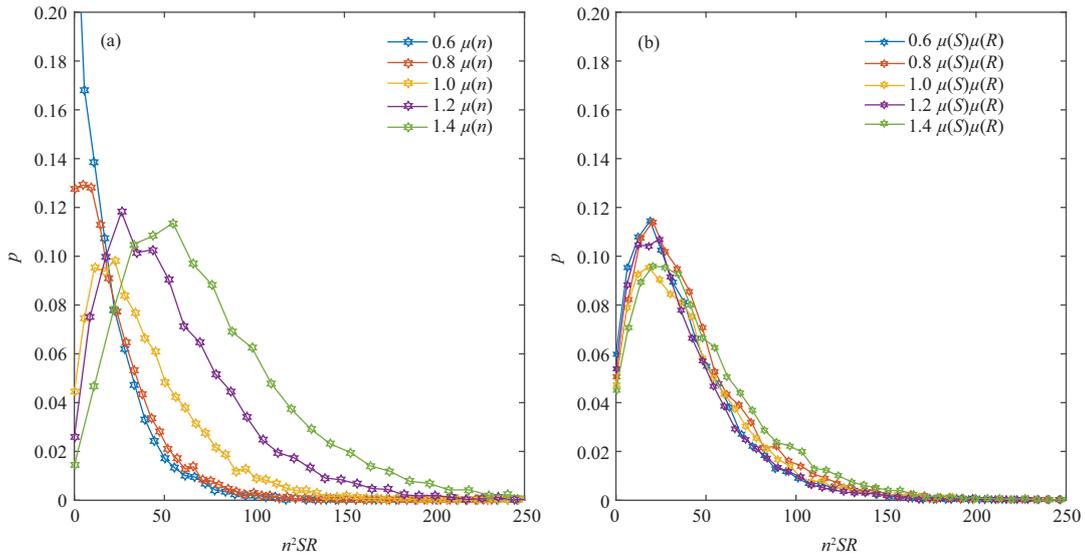


图 7 (网络版彩图) 随机变量 n, S, R 数学期望变化影响效果图. (a) 在线时长期望值 $\mu(n)$ 对 $f_1(i)$ 的影响; (b) 分享系数、声誉系数期望值 $\mu(S)\mu(R)$ 对 $f_1(i)$ 的影响

Figure 7 (Color online) Feature extraction diagram. (a) The impact of online time expectation $\mu(n)$ on $f_1(i)$; (b) the impact of sharing coefficient and reputation coefficient expected value $\mu(S)\mu(R)$ on $f_1(i)$

结合式 (18), 两个节点间的信息传播实际上就是一个非线性时变能量转换系统^[12~15], 系统表达式为

$$E(j, t + 1) = Z(i, j, t)E(i, t). \quad (25)$$

为了进一步考察输入随机变量 n, S, R, ω 的特性变化对信息能量输出值的影响, 下面逐一考察随机变量不同的变化情况. (1) 考察 $f_1(i)$ 随在线时长 n 的期望值的变化关系, 以 n^2SR 为横坐标、 $f_1(i)$ 的概率 p 为纵坐标, 绘制概率密度图, 如图 7 所示. 由图 7(a) 中可看出, $f_1(i)$ 的概率密度随着 $\mu(n)$ 的增大明显右移, 数学期望和方差同时增大, 说明用户平均在线时间越长, 信息传播时, 传递给后继节点的信息能量越大、传播路径越长、信息覆盖面越大. 用同样的方法观察分享系数、声誉系数期望值 $\mu(S)\mu(R)$ 变化情况, 结果如图 7(b) 所示, 对 $f_1(i)$ 的概率分布并无显著影响, 说明社会上分享系数、声誉系数平均水平在一定范围内变化对信息传播没有太大影响.

(2) 考察 $f_1(i)$ 随在线时长 n 的方差的变化关系, 分别取在线时长方差 $\sigma^2(n)$ 的 0.6, 0.8, 1.0, 1.2, 1.4 倍的样本作为参变量输入系统, 以 n^2SR 为横坐标、 $f_1(i)$ 的概率为纵坐标, 绘制概率密度图, 如图 8 所示.

由图 8(a) 可看出, $f_1(i)$ 的概率密度随着 $\sigma^2(n)$ 的改变没有太大变化, 说明信息传播与人们上线时长是否完全相等无关, 只取决于大家的平均时长. 用同样的方法观察分享系数、声誉系数方差 $\sigma^2(S)\sigma^2(R)$ 的变化, 如图 8(b) 所示, 发现对 $f_1(i)$ 的概率分布影响显著, 概率密度曲线随着参变量样本离散程度的增长向右平移, $f_1(i)$ 的数学期望和方差均有增加. 说明社会上个人分享意愿、声誉水平差别较大时, 认知水平的差异也较大, 信息传播中容易形成舆论场, 提高信息势能差, 促使信息快速传播.

(3) 为了简化问题的分析, 暂时忽略社会媒体的影响, 即令社会舆论场力 $F_3 = 0$, 考察 $f_2(i)$ 随亲密度 w 的变化关系, w 服从指数分布, 即 $w \sim e(\lambda)$, 分别取亲密度期望值 $\lambda = 0.1, 0.2, 0.3, 0.4, 0.5$ 的样本值作为参变量输入系统, 以 w 为横坐标、 $f_2(i)$ 的概率 p 为纵坐标, 绘制概率密度图, 如图 9 所

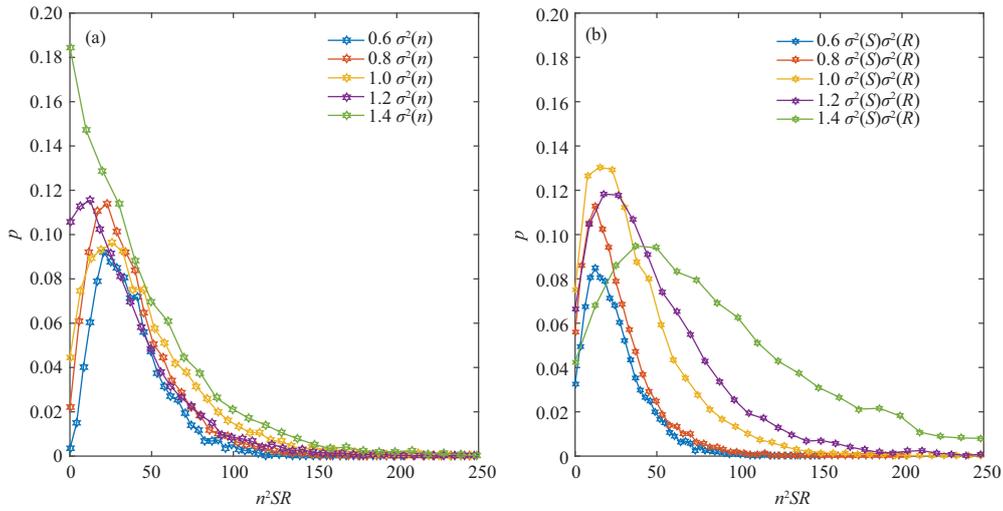


图 8 (网络版彩图) 随机变量 n, S, R 方差变化影响效果图. (a) 在线时长方差 $\sigma^2(n)$ 对 $f_1(i)$ 的影响; (b) 分享系数、声誉系数方差 $\sigma^2(S)\sigma^2(R)$ 的影响

Figure 8 (Color online) Feature extraction diagram. (a) The impact of online time variance $\sigma^2(n)$ on $f_1(i)$; (b) the impact of sharing coefficient and reputation coefficient variance $\sigma^2(S)\sigma^2(R)$

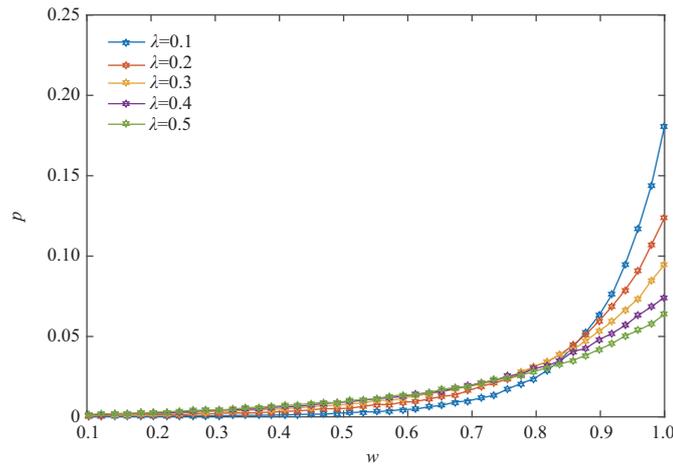


图 9 (网络版彩图) 随机变量 w 数学期望、方差变化影响效果图

Figure 9 (Color online) Effect diagram of the mathematical expectation and variance of random variable w

示. 由图 9 可看出, 个体之间亲密度系数 $w \leq 0.85$ 时, w 分布的改变对 $f_2(i)$ 的概率分布并无影响; 但当 $w > 0.85$ 时, 对 $f_2(i)$ 的概率分布影响显著, 而且这种影响力随着 w 的继续增长呈几何级数增长. 说明在现实社会中, 人与人之间的关系亲密程度并不是线性增长的, 只有当亲密程度足够高时, 比如朋友关系、亲人关系等, 才会影响人的行为, 进而影响信息传播.

(4) 用户上网习惯对信息传播的影响呈周期性变化, 图 10 显示了 $f_3(i, t)$ 随时间 t 的变化规律, 首先对 $f_3(i, t)$ 进行复合函数分解, 绘制了每一个分量的变化规律, 如图 10 中深蓝色、天蓝色、粉红色、绿色曲线, 然后再将各分量进行叠加合成, 得到 $f_3(i, t)$ 的变化趋势, 如图 10 中红色曲线. 网络节点只有处于红色曲线上凸部分的时域内, 才能向后继节点传播信息和辐射信息能量. 说明用户在在线社交

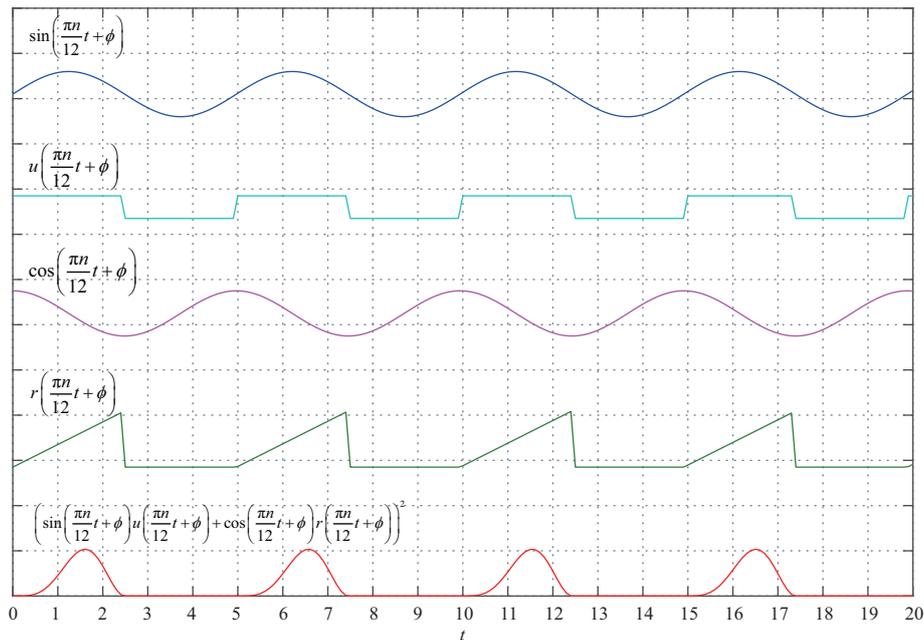


图 10 (网络版彩图) 信号合成示意图

Figure 10 (Color online) Signal composition diagram

网络上的信息传播行为是间歇性和阵发性的, 只有在整个周期中那一小段上凸部分时域内才能收到信息, 才能向其他用户转发信息; 否则, 必须等待下一个周期的到来, 同时由于信息具有时效性, 如果未能及时将信息转发给自身的邻居节点, 那么邻居节点很有可能已从其他路径获悉信息, 导致该用户在下一周期内转发的信息被忽略, 最终信息传播失败.

3.2 仿真分析

将基于时变阻尼运动的信息传播动力学模型 TVDM 应用于网络节点总数为 $N = 1000$ 的无标度网络上开展仿真实验. 网络节点度分布服从 $\gamma = 3$ 的 $p(k) \sim k^{-\gamma}$ 幂律分布, 分布图如图 11 所示. 任意选取一个节点作为种子节点, 将初始信息能量 $E_0 = 20$ 的信息在 $t = 0$ 时刻注入网络, 随即在网络扩散传播, 信息传播效果如图 11 所示.

图 11(a) 反映了网络上信息转发情况, 即信息转发率随时间 t 的变化关系, 传播初期信息转发率呈指数级增长, 在 $t = 5$ 时达到峰值, 瞬时转发率达到 0.27. 之后由于传播过程中信息能量的衰减, 部分路径上的传播行为终止, 导致转发率逐渐下降, 在 $t = 8$ 时低至 0.01. 最后在 $t \geq 9$ 的周期上, 网络趋于稳定, 只有极少数节点存在信息转发行为. 图 11(b) 反映了网络中的信息覆盖情况, 即信息覆盖率随时间 t 的变化关系, 整体上呈现出“S”形曲线变化规律, 起初阶段大致是按指数增长, 然后开始变得饱和, 增加变慢, 最后在 $t = 9$ 时增加停止, 最终稳定在 0.91. 传播过程结束后, 网络中仍有 9% 的节点, 即 900 个节点, 未收到信息, 说明在在线社交网络中, 一条信息的传播, 无法让所有的用户都能接收到. 信息传播还受到社会舆论场力的影响, 对一般突发公共事件来说, 媒体引导使人们趋于理性, 不要盲目转发信息, 因此施加了负的舆论场力, 在一定程度上抑制了信息的传播. 在实验网络中媒体节点数量分别取 $Media = 0, 10, 20, 30, 40, 50, 60$ 进行仿真, 将处于不同数量媒体情况下的信息转发效果和覆盖效果绘制成图表, 如图 12 所示.

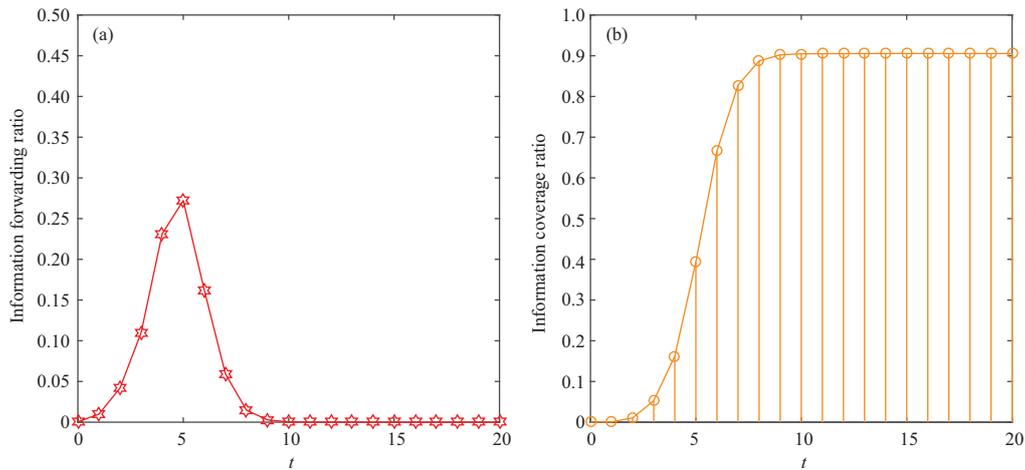


图 11 (网络版彩图) 信息传播趋势图. (a) 信息转发率; (b) 信息覆盖率

Figure 11 (Color online) Information dissemination trend chart. (a) Information forwarding ratio; (b) information coverage ratio

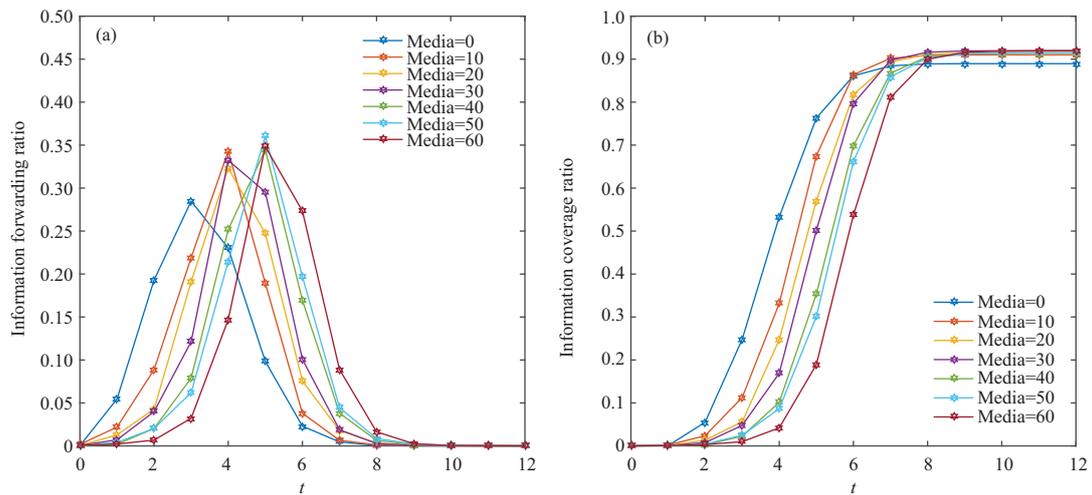


图 12 (网络版彩图) 媒体影响效果图. (a) 信息转发率; (b) 信息覆盖率

Figure 12 (Color online) Media influence rendering. (a) Information forwarding ratio; (b) information coverage ratio

随着媒体节点数量的增加, 信息转发活动增长变慢, 转发峰值时刻来得较晚, 如图 12(a) 所示. 同时, 信息覆盖面的增长也比较慢, 但最终都能覆盖网络的大部分, 如图 12(b) 所示. 说明媒体的理性评论只会延缓信息传播的速度, 而对信息覆盖面没有太大影响.

3.3 实证对比分析

为了进一步验证传播模型的有效性和可靠性, 根据 2018 年 12 月份发生的热点事件, 选取 6 条不同领域事件的典型信息数据, 详细信息如表 2 所示.

“新榜”平台 (www.newrank.cn) 依托数据挖掘分析能力, 建立用户画像和效果监测系统, 它向众多企业、政府机构提供线上、线下数据产品服务, “新榜认证”、“分钟级监测”应用广泛. 本文利用“新榜”数据服务的趋势查询功能, 获取上述信息 A ~ F 的传播监测统计数据, 采集信息传播周期内每天

表 2 六条不同领域事件
Table 2 Six different domain events

Serial number	Category	Content	Browse maximum value
A	Diplomacy	China and the United States agree not to impose new tariffs	1159691
B	Military	U.S. withdrawal from Syria	314462
C	Politics	Speech at the 40th anniversary of reform and opening up	4462115
D	Technology	Blackmail virus requiring WeChat to pay ransom	1829246
E	Sports	Ke Jie defeated Korean players	286674
F	Livelihood	Yueqing boy lost incident	3783792

信息的原创篇数、浏览数、点赞数、活跃天数等数据, 取其中的阅读数进行归一化处理后作为实证数据, 该数据体现了单位网络上浏览信息的用户数量, 也就是真实传播过程中每天的信息覆盖率. 取本文模型动态模拟演化结果的信息覆盖率的时间序列值为分析数据, 其不同数据间的对比分析如图 13 所示.

从图 13 中可以看出, 6 条不同领域事件的真实数据与仿真模拟数据的变化在不同的时间段均表现出了很好的趋势一致性. 同时, 用余弦相似性 Similarity 表示两组数据的接近程度; 用相关系数 ρ 表示两组数据间的相关程度; 误差峰值 deviation 表示两组数据之间局部最坏情况下的离散程度; 分别定义如下:

$$\text{Similarity} = \frac{\sum_{i=1}^{20} (y_i \times y'_i)}{\sqrt{\sum_{i=1}^{20} y_i^2} \sqrt{\sum_{i=1}^{20} (y'_i)^2}}, \quad (26)$$

$$\rho = \frac{\sum_{i=1}^{20} (y_i - \bar{y})(y'_i - \bar{y}')}{\sqrt{\sum_{i=1}^{20} (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^{20} (y'_i - \bar{y}')^2}}, \quad (27)$$

$$\text{deviation} = \max |y_i - y'_i|. \quad (28)$$

Similarity, ρ 与 deviation 在两组数据之间的分析结果如表 3 所示.

六条不同领域事件的真实数据与仿真模拟数据之间的关系的信息受力分析如图 14 所示. 图 14 中, 不同数据之间的波动较小, 较大的波动主要发生在信息大规模转发的阶段, 此时系统受外界影响比较大, 一个较小的扰动都会引起较大的变化, 比如信息当事人的澄清、相关人员的信息补充等. 从实证对比的结果来看, 仿真结果与真实数据的吻合程度高, 相似度达 0.90 以上、相关系数达 0.91 以上、误差峰值在 0.2 的范围以内. 说明模型体现了信息传播的关键因素, 反映了信息传播过程的客观规律, 模型是准确的、有效的.

4 结语

本文以在线社交网络中信息传播在微观上的随机性和无序性与宏观上的可认知和可观测之间的关系为主要研究对象. 首先对国内外针对信息传播动力学模型进行了总结, 介绍了基于传染病、博弈论和物理系统的 3 类主要模型; 其次, 深入分析在线社交网络上信息传播现象, 分别从时间、空间、主体、客体 4 个维度提取关键特征, 进而进行客观化的定量描述; 然后, 依据运动学、动力学和时变系统与信号变换理论, 对社交网络上的信息传播过程进行分析, 构建出基于时变阻尼运动的信息传播动力学模型; 最后通过数值模拟实验分析了模型中参数变化的影响和个体特征分布所服从的统计规律, 发

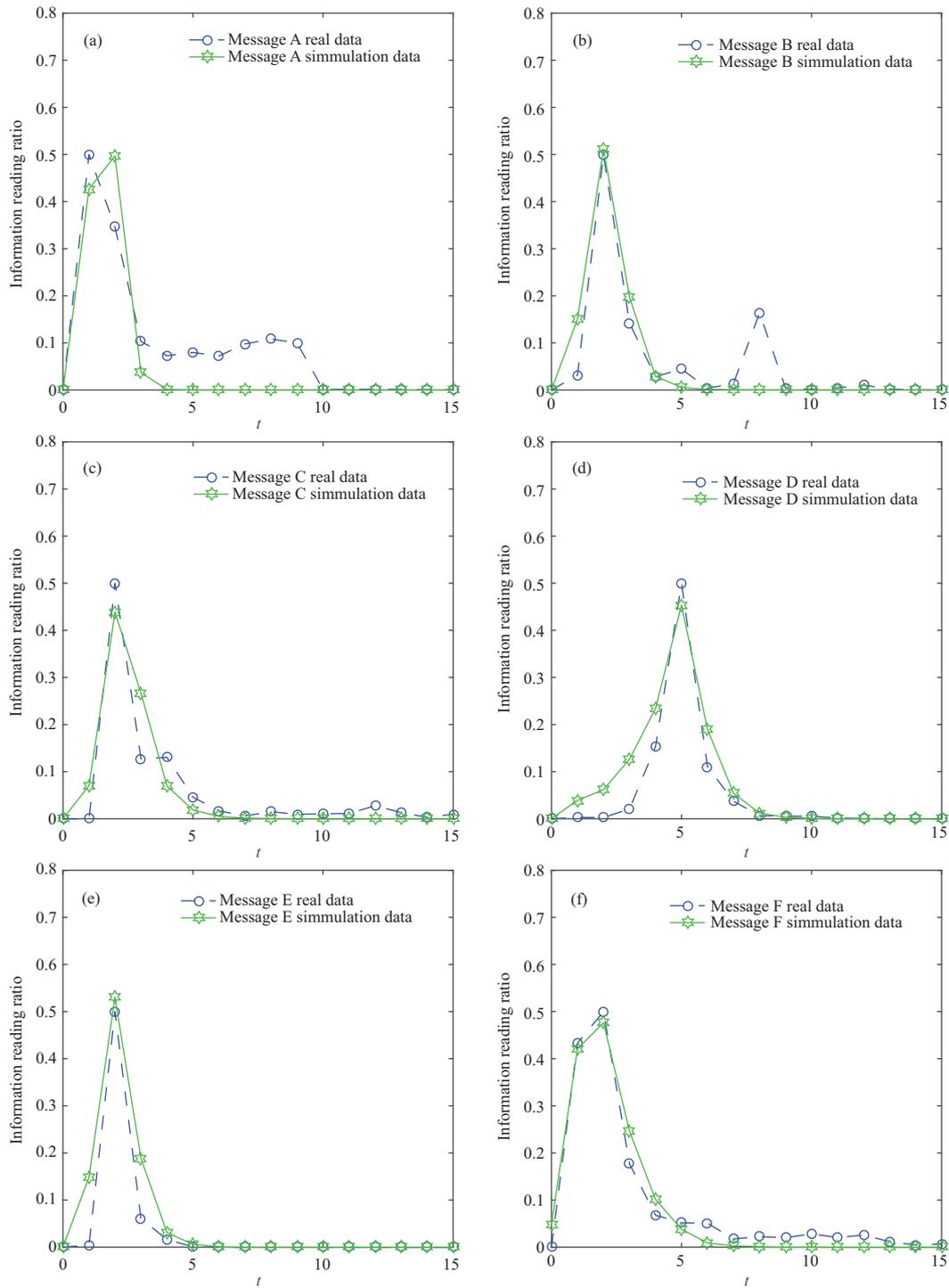


图 13 (网络版彩图) 仿真与实证数据对比图. (a) 外交; (b) 军事; (c) 政治; (d) 科技; (e) 体育; (f) 民生

Figure 13 (Color online) Comparison diagram of simulation and empirical data. (a) Diplomacy; (b) military; (c) politics; (d) technology; (e) sports; (f) livelihood

表 3 三种评价指标分析表

Table 3 Analysis table of three evaluation indexes

Content	Similarity	ρ	deviation
China and the United States agreed not to impose new tariffs	0.9064	0.9117	0.1509
U.S. withdrawal from Syria	0.9271	0.9131	0.1636
Speech at the celebration of the 40th anniversary of reform and opening	0.9390	0.8258	0.1393
The ransomware that requires WeChat to pay a ransom	0.9501	0.9488	0.1056
Ke Jie defeated South Korean players	0.9452	0.9443	0.1444
The Yueqing boy lost incident	0.9855	0.9815	0.0693

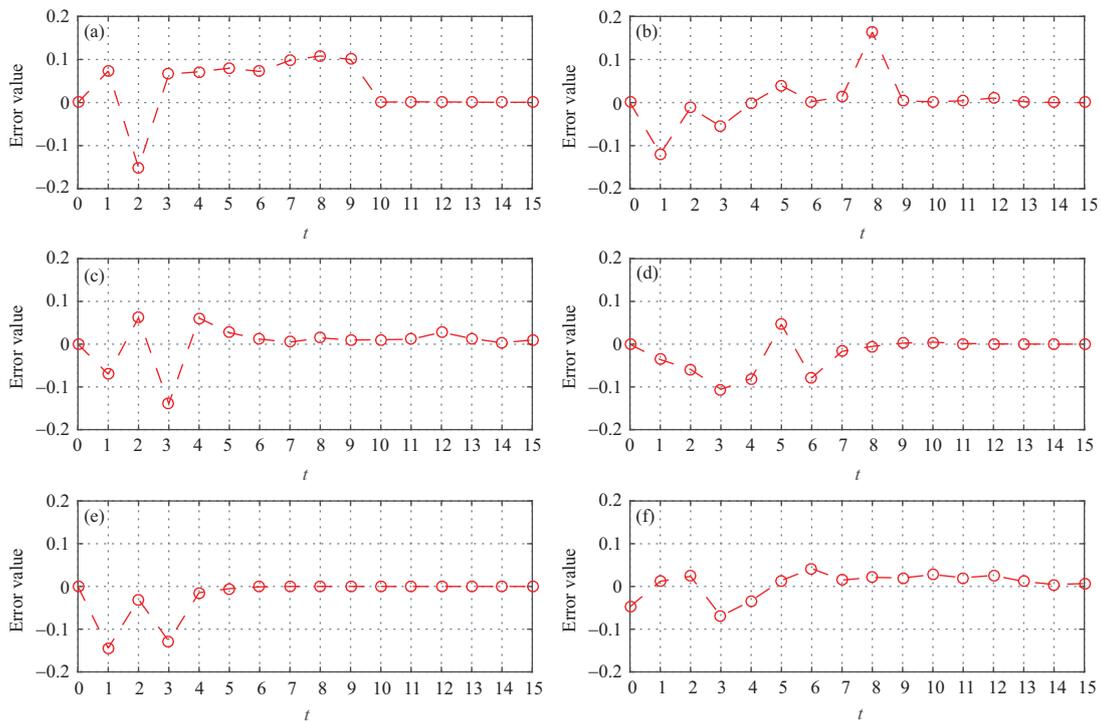


图 14 (网络版彩图) 仿真与实证数据误差图. (a) 外交; (b) 军事; (c) 政治; (d) 科技; (e) 体育; (f) 民生

Figure 14 (Color online) Error chart of simulation and empirical data. (a) Diplomacy; (b) military; (c) politics; (d) technology; (e) sports; (f) livelihood

现其相似度达 0.90, 相关系数达 0.91, 精度高. 一系列实验结果 (数值模拟、仿真分析、实证对比 3 种验证方法) 表明, 本项研究更能科学合理解释社交网络信息传播活动的本质, 实验设计更加客观, 指标更加具有说服力. 因此, 提出的 TVDM 模型是合理、有效的, 为社交网络上信息传播的研究提供了一种新的研究思路和方法; 同时, 提出的信息传播模型也可应用在广义社交网络信息传播与人类动力学行为分析研究领域.

参考文献

1 Arnaboldi V, Conti M, Passarella A, et al. Online social networks and information diffusion: the role of ego networks. Online Soc Networks Media, 2017, 1: 44-55

- 2 Wu Q, Lou Y, Zhu W. Epidemic outbreak for an SIS model in multiplex networks with immunization. *Math Biosci*, 2016, 277: 38–46
- 3 Wang X, Fang B, Zhang H, et al. A competitive model of popularity prediction using correlation matrix. *Cluster Comput*, 2019, 22: 2747–2759
- 4 Gao J, Zhang T, Science S O, et al. Analysis on an SEIRS epidemic model with pulse vaccination and two time delays. *Chin Quart J Math*, 2019, 34: 75–87
- 5 Ma J, Si F. Complex dynamics of a continuous bertrand duopoly game model with two-stage delay. *Entropy*, 2016, 18: 266
- 6 Wang X J, Gu C L, Lv S J, et al. Evolutionary game dynamics of combining the Moran and imitation processes. *Chin Phys B*, 2019, 28: 020203
- 7 Lu J, Wang L, Qian J, et al. Research on information dissemination in online social network based on human dynamics. *J Electron Inform Technol*, 2017, 39: 785–793
- 8 Xiao Y-P, Li S-Y, Liu Y-B. An information diffusion dynamic model based on social influence and mean-field theory. *Acta Phys Sin*, 2017, 66: 030501
- 9 Xiao R, Han J, Bi Y, et al. Dynamics model of information transmission in online social network based on the gravitational potential. *J Shanxi Univ (Nat Sci Ed)*, 2014, 37: 101–106
- 10 Zhu C, Yin G. On hybrid competitive Lotka-Volterra ecosystems. *Nonlin Anal-Theor Methods Appl*, 2009, 71: 1370–1379
- 11 Hoffman L H. Is Internet content different after all? A content analysis of mobilizing information in online and print newspapers. *Jism Mass Commun Q*, 2006, 83: 58–76
- 12 Benyoussef M, Ez-Zahraouy H, Benyoussef A. Optimal topology to minimizing congestion in connected communication complex network. *Int J Mod Phys C*, 2017, 28: 1750073
- 13 Liu C, Natarajan B. Modeling and analysis of simultaneous information and energy transfer in Internet of Things. *Trans Emerging Tel Tech*, 2017, 28: e3177
- 14 AL-Shudeifat M A. Time-varying stiffness method for extracting the frequency-energy dependence in the nonlinear dynamical systems. *Nonlin Dyn*, 2017, 89: 1463–1474
- 15 Li N, Das S K. Efficiently discovering users connectivity with local information in online social networks. *Online Soc Netw Media*, 2020, 16: 100062

Social network information propagation dynamic analysis based on time-varying damping motion

Xiaoyang LIU^{1*}, Daobing HE¹, Chao LIU¹ & Yihao ZHANG^{2,3}

1. School of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China;

2. School of Artificial Intelligence, Chongqing University of Technology, Chongqing 401135, China;

3. School of Computing, National University of Singapore, Singapore 119077, Singapore

* Corresponding author. E-mail: lxy3103@cqut.edu.cn

Abstract According to the current research achievements in online social network (OSN) information propagation dynamics, the relationship between the randomness and disorder of micro individuals and the knowability and controllability of macro group is not revealed through the social physics method. This paper proposes a novel information propagation dynamics model based on time-varying damping motion (TVDM). Firstly, the phenomenon of information propagation in OSN is analyzed in depth, and key features are extracted. Then, from the perspectives of simple harmonic motion with damping, combined forces of internal and external force systems, information energy transformation, time-varying system and signal transformation, this paper reveals the physical governing equations and action mechanism of information propagation. Finally, the numerical simulation experiment verifies that the individual behavior on the network obeys the statistical regularity. The empirical comparison experiment verifies that the model simulation results are highly consistent with the real event statistical data. Experimental results show that the precision of the proposed TVDM model is about 91%, the novel model is reasonable and effective.

Keywords social network, dynamics, information energy, damping motion, time-varying system



Xiaoyang LIU was born in 1980. He received his Ph.D. degree in communication and information systems from Northwestern Polytechnical University, Xi'an China, in 2013. From 2015 to 2017, he has completed his first post-doctoral fellow in Chongqing University, China. Then he completed his second post-doctoral fellow in Department of Electrical of Computer Engineering, The University of Alabama from 2017

to 2018 in Tuscaloosa, Alabama, USA. He is currently a professor in Chongqing University of Technology, China. His main research interests are in the areas of online social network, information diffusion, complex network.



Daobing HE was born in 1980. He is a master student in Chongqing University of Technology. His research interests are in the areas of mathematics model and social network.



Chao LIU was born in 1983. He received his Ph.D. degree in computer science and technology from Chongqing University in 2012. Currently, he is an associate professor at Chongqing University of Technology. His research interest includes dynamical system.



Yihao ZHANG was born in 1982. He received his Ph.D. degree in computer science from Chongqing University in 2014. Currently, he is an associate professor of Chongqing University of Technology. His research interests include recommendation system, natural language processing, and machine learning.