

引用格式: 樊静妍, 胡枫, 郭磊, 等. 基于个体属性异质的微博信息超网络传播模型 [J]. 电子科技大学学报, 2026, 55(1): 137-148.

FAN J Y, HU F, GUO L, et al. A Weibo information hypernetwork propagation model based on individual attribute heterogeneity[J]. Journal of University of Electronic Science and Technology of China, 2026, 55(1): 137-148.

基于个体属性异质的微博信息 超网络传播模型



樊静妍^{1,2}, 胡枫^{1,2*}, 郭磊^{1,2}, 杨煜升^{1,2}, 宋玉蓉^{3,4}

(1. 青海师范大学计算机学院, 西宁 810008; 2. 藏语智能信息处理及应用国家重点实验室, 西宁 810008;

3. 南京邮电大学自动化学院, 南京 210023; 4. 南京邮电大学人工智能学院, 南京 210023)

摘要: 异质网络能够有效建模现实世界的诸多复杂应用场景。基于微博平台个体的多样性, 该文提出构建个体属性异质的微博信息超网络模型, 模型以用户、话题为两类异质节点, 用户参与话题讨论为超边, 构建无标度异质超网络模型。在此基础上, 结合 SEIR 传播模型, 对异质节点的个体属性进行量化分析, 通过元路径方法设计用户影响力、感染率和免疫率的计算方法。此外, 通过仿真实验分析不同网络结构下信息传播的动态过程和规律, 研究用户影响力、置信度、兴趣价值、信息时效性对该模型信息传播过程的影响。进一步, 通过“日本核污水排放”事件验证模型的有效性和准确性。结果表明, 该模型能够较为准确地描述真实社交网络中的信息传播趋势和过程。该工作对异质超网络的模型构建及超网络信息传播的研究有一定的借鉴意义, 有助于深入研究更复杂多元的信息传播机制。

关键词: 超图; 异质超网络; 微博信息传播; 个体属性异质; 元路径

中图分类号: TN92; O41

文献标志码: A

DOI: 10.12178/1001-0548.2024330

A Weibo information hypernetwork propagation model based on individual attribute heterogeneity

FAN Jingyan^{1,2}, HU Feng^{1,2*}, GUO Lei^{1,2}, YANG Yusheng^{1,2}, and SONG Yurong^{3,4}

(1. College of Computer, Qinghai Normal University, Xining 810008, China;

2. The State Key Laboratory of Tibetan Intelligent Information Processing and Application, Xining 810008, China;

3. College of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210023, China;

4. College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Heterogeneous networks can effectively model various complex real-world application scenarios. Based on the diversity of individuals on the Weibo platform, this paper proposes a heterogeneous Weibo information hypernetwork model that incorporates individual attributes. In this model, users and topics are treated as two types of heterogeneous nodes, while user participation in topic discussions forms hyperedges, resulting in a scale-free heterogeneous hypernetwork model. Building on this framework, the SEIR propagation model is integrated to quantitatively analyze the individual attributes of heterogeneous nodes. Meta-path methods are employed to design calculation approaches for user influence, infection rate, and immunity rate. Furthermore, simulation experiments are conducted to analyze the dynamic processes and patterns of information dissemination across different network structures, investigating the impacts of user influence, confidence levels, interest values, and information timeliness on the dissemination process. Additionally, the validity and accuracy of the model are verified using the case study of the "Japan Nuclear Wastewater Discharge" incident. The results demonstrate that this model effectively captures the trends and processes of information dissemination in real social networks. This work provides valuable insights for research on the model construction of heterogeneous hypernetworks and hypernetwork-based information dissemination, contributing to a deeper understanding of more complex and diverse information propagation mechanisms.

收稿日期: 2024-12-05

基金项目: 国家自然科学基金 (62566052); 青海省自然科学基金 (2023-ZJ-916M)

作者简介: 樊静妍, 主要从事超网络、信息传播动力学方面的研究。

*通信作者 E-mail: qhuf@163.com

Key words: hypergraph; heterogeneous hypernetwork; Weibo information dissemination; individual attribute heterogeneity; meta-path

随着 Web2.0 时代的到来, 在线社交网络、微博、微信等众多社交媒体平台迅速普及, 极大提高了信息传播速度与范围^[1]。在微博网络中, 信息传播不仅与用户的兴趣偏好、活跃程度等个体属性相关, 还受到用户与话题间复杂关联关系的影响。个体异质对信息传播路径和效果产生显著影响, 导致传播规律的多样性与复杂性。因此, 研究异质节点在超网络中的信息传播规律, 不仅有助于揭示社交网络中的信息扩散机制, 而且在诸多领域中有显著的应用价值, 如疾病传播^[2-3]、兴趣推荐^[4]、社区检测^[5]、舆情控制^[6-7]等。

在信息传播研究领域, 学者们构建了大量模型来描述信息传播机制及规律, 并对其展开了深入研究。文献 [8] 最早提出 SIR 数学模型研究流行病传播过程, 随后文献 [9] 在此基础上引入暴露状态 (E 状态), 提出 SEIR 模型, 为分析具有潜伏期的传染病传播提供理论基础。鉴于社交网络中的信息传播过程与传染病传播过程相似, 研究者借鉴流行病传播模型, 模拟信息、谣言、社交媒体内容等的传播过程。文献 [10] 使用平均场理论对 SEIR 传播模型进行扩展, 以探究传播概率、网络结构、初始传播者和社会效应对传播过程的影响。文献 [11] 提出基于模糊逻辑的 SEICRS 谣言传播模型, 该模型在传统的 SEIR 模型的基础上引入谣言延迟和反击机制, 有效降低了谣言的传播速率。然而, 在现实生活中个体具有复杂的交互关系, 传统复杂网络难以准确表示完整的网络关系及特征^[12]。因此, 基于超图的超网络应运而生^[13]。近年来, 针对超网络的信息传播模型研究已取得显著进展。文献 [14] 基于超图理论提出一种改进的 SSEIR 模型, 将传统易感状态细分为活跃易感状态和非活跃易感状态, 更真实地模拟用户在信息传播中的行为差异。文献 [15] 基于用户和信息属性提出在线社交超网络信息传播模型, 揭示了不同结构参数下信息的传播机制。文献 [16] 提出一种基于超网络的非线性耦合动态模型, 并通过马尔科夫链方法推导了动态系统的演化过程。

考虑到个体的异质性和复杂的交互关系, 采用异质网络来研究信息传播变得日益重要, 因为它比传统的同质网络模型能更准确地描述信息的传播过程。文献 [17] 构建具有多种类型节点和边组成的异质网络, 使用不同元路径之间的相互作用来预测

传播过程。文献 [18] 构建具有不同兴趣或行为的个体组成的异质网络, 探讨了在社交网络中信息传播的再现现象。文献 [19] 通过潜在的表示学习分析异质网络中的信息扩散, 探究不同结构参数下信息的传播机制。然而, 现有的异质信息传播主要基于普通复杂网络, 但这类网络并不能准确描述社交网络中的群聚特性。如在微信或 QQ 中, 用户既能将信息私发给某个好友, 也可以转发到多个社交群组中。为了能够精准描述真实社交网络中用户之间复杂的交互关系, 学者们构建大量异质超网络模型来描述网络结构对信息传播产生的影响。文献 [20] 采用优先连接机制构建具有双重异质性的生成超图模型。文献 [21] 采用 3 种异构超图模型对校企合作网络数据进行研究。然而, 传统的信息传播模型通常基于单一节点类型或同质性假设, 难以准确刻画真实微博网络中用户和话题的异质性及复杂的交互关系, 且真实的社交网络更契合于用超图表示的超网络结构。针对上述问题, 本文基于微博平台个体的多样性, 提出构建个体属性异质的微博信息超网络传播模型。该模型以用户和话题为两类异质节点, 用户参与话题讨论为超边, 充分体现网络结构的异质性特征。在此基础上, 结合 SEIR 传播模型, 对异质节点的个体属性进行量化分析, 进一步增强对网络中复杂交互关系的刻画能力。该模型不仅为更准确地描述信息传播规律提供理论依据, 也为社交网络中信息扩散的实际应用奠定基础。

1 相关理论

定义 1 超图

文献 [22] 首次提出超图的定义, 设 $V = \{v_1, v_2, \dots, v_n\}$ 是一个有限集, 如果 $e_i \neq \emptyset (i = 1, 2, \dots, m)$ 且 $\cup_{i=1}^m e_i = V$, 则二元关系 $H = (V, E)$ 称为超图, 其中 $v_i (i = 1, 2, \dots, n)$ 称为超图的顶点, $E = \{e_1, e_2, \dots, e_m\}$ 是超图的边集合, $e_j (j = 1, 2, \dots, m)$ 称为超图的超边。超图 $H = (V, E)$ 的邻接矩阵 $A(H)$ 是一个 $n \times n$ 的对称方阵, 其元素 a_{ij} 为同时包含顶点 v_i 和 v_j 的超边数量, 对角线元素为 0。关联矩阵 $B(H)$ 是一个 $n \times m$ 的矩阵, 如果顶点 v_i 在超边 e_j 中, 则 $b_{ij} = 1$, 否则为 0。节点 v_i 的超度定义为包含该节点的超边数量, 记为 $d_H(v_i)$ 。节点的超度分布 $P(d_H)$ 定义为超度为 d_H 的节点数量 N_{d_H} 与节点总数的比值, 即:

$$P(d_H) = \frac{N_{d_H}}{N} \quad (1)$$

定义 2 异质超网络

超图通过不同类型的节点或超边来捕捉底层超网络的异质性。如果超网络中仅存在同一类型节点和同一类型超边, 则此时的超网络为同质超网络。如果超网络中存在多种不同类型的节点, 或存在多种不同类型的超边, 则将其称为异质超网络, 此时超边可以表示不同类型的事件^[23]。如图 1 所示, 社会化标签网络由文档、标签和注释关系这 3 类异质节点构成异质超网络, 其中蓝色的超边表示用户用书签注释文档, 橙色的超边表示文档和内容相似性关系, 绿色超边代表标签关系^[24]。

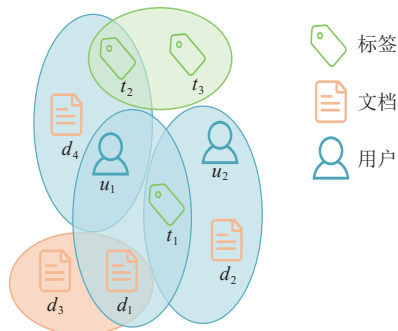


图 1 异质超网络示意图

定义 3 元路径

元路径^[25] P 被定义为 $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} A_{l+1}$ (缩写为 A_1, A_2, \dots, A_{l+1}), 表示对象 A_1, A_2, \dots, A_{l+1} 间的复合关系 $R_1 \circ R_2 \circ \dots \circ R_l$, 其中, \circ 表示关系上的合成运算符。元路径 P 下存在多个特定路径, 称为路径实例, 表示为 p 。如图 2 所示, 对于用户 u_1 和 u_2 , 两个节点之间可以被多条路径实例相连, 如 $u_1 - t_1 - u_2$ 和 $u_1 - t_6 - u_4 - t_7 - u_2$ 。仅通过一个中介体连接的元路径称为单介体元路径, 如在路径 $u_1 - t_1 - u_2$ 中, 用户 u_1 仅通过话题 t_1 连接到用户 u_2 。通过多个中介体连接的元路径称为多介体元路径或 n 介体元路径。如在路径 $u_1 - t_6 - u_4 - t_7 - u_2$ 中, 用户 u_1 通过话题 t_6 、 t_7 连接到用户 u_2 。

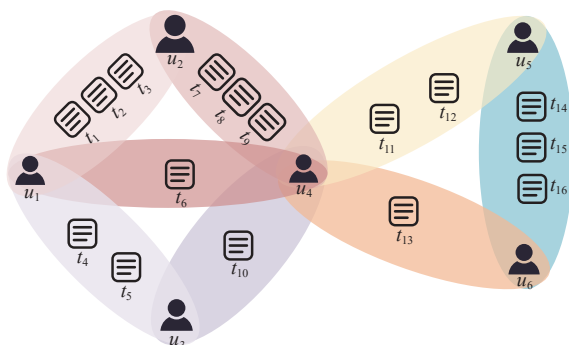


图 2 异质网络中元路径实例

表 1 为图 2 中所使用的部分元路径实例。元路径 UTU (user-topic-user) 表示两个用户通过参与相同话题的讨论而形成的直接联系, 表明用户在某些话题具有共同的兴趣或关注度, 能够衡量用户在局部网络的影响力。而 UTUTU (user-topic-user-topic-user) 反映了用户之间通过共同邻居产生的间接联系, 揭示了用户之间更深层次的兴趣链, 能够揭示用户在更大范围网络的潜在影响力。

表 1 图 2 的部分元路径实例

类型	缩写	元路径	路径实例
单介体元路径	UTU	user-topic-user	$u_1 - t_1 - u_2$
			$u_1 - t_2 - u_2$
			\vdots
多介体元路径	UTUTU	user-topic-user-topic-user	$u_1 - t_4 - u_3 - t_{10} - u_4$
			$u_1 - t_6 - u_4 - t_7 - u_2$
			\vdots

定义 4 信息熵

信息熵^[26]利用概率与统计方法来表征样本空间所体现的系统无序化程度, 进而反映节点在网络中的重要性。信息熵被定义为:

$$H = - \sum_{i=1}^N P_i(x_i) \log P_i(x_i) \quad (2)$$

式中, H 为事件整体的信息熵; x_i 为某一可能出现的随机事件; P_i 为事件 x_i 出现的概率。

2 微博信息超网络传播模型

微博中个体由用户和话题组成, 这两类节点具有显著的异质性: 用户在网络中的活跃程度、兴趣偏好和影响力存在较大差异, 大多数用户仅参与少量话题讨论, 而少数活跃用户频繁参与大量热门话题; 同时, 话题的受欢迎程度、生命周期具有多样性, 少数话题吸引广泛讨论, 而多数话题仅受到少量用户关注。因此, 用异质超网络能更准确地表达微博网络的异质性。现有的超网络模型, 如文献 [27] 中提出的通用无标度超网络模型, 虽然较为准确地描述了微博网络中用户行为及信息传播的群体互动特性, 但该模型主要依赖单一类型的节点, 未能充分考虑社交平台中个体属性的多样性及用户与话题之间复杂的交互关系。为此, 本文在该模型基础上进行改进, 引入用户和话题两类异质节点, 并结合不同个体属性的影响因素, 提出构建具有异质特性的微博超网络模型。该模型能更准确地模拟微博平台上信息传播的过程和规律。

2.1 微博超网络的构造过程

本文在 BA 无标度超网络模型^[27]基础上改进后构建微博超网络, 作为本文所提模型的底层拓扑结构, 构造过程如下。

1) 初始化: 在初始时刻, 超网络由 m_0 个用户节点、一个话题节点组成一条超边, 表示初始用户参与某个话题的讨论, 形成最初的社交群组。

2) 增长: 在每个时间步中, 向网络中添加新的用户节点或话题节点, 以模拟用户不断加入平台并发起新话题的过程。在真实社交网络中, 用户更倾向于参与讨论话题而非主动发布话题。因此, 本文引入用户数与话题数之间的权重关系, 并设置其比例为 $N_u : N_t$ 。在节点添加过程中, 通过以下两种方式生成超边: 若生成节点为用户节点, 则选择已有的 m_1 个话题节点, 使新用户节点优先连接到参与度较高的话题节点, 即更频繁地参与讨论的话题。若生成节点为话题节点, 则随机选择已有 m_2 个用户节点, 使新话题节点优先连接到活跃用户, 从而模拟新话题首先被活跃用户参与的情境。

3) 优先连接: 新加入的节点与已有节点建立连接的概率与已有节点的超度成正比。即新发布的话题节点, 越容易吸引活跃用户进行讨论。新节点连接到已有节点 v_i 的概率为:

$$\Pi d_H(i) = \frac{d_H(i)}{\sum_j d_H(j)} \quad (3)$$

式中, 分子 $d_H(i)$ 表示节点 i 的超度; 分母表示所有节点的超度之和。图 3 为微博超网络的演化过程示例图, 其中, 黑色节点表示已有的用户或话题, 蓝色节点表示新加入的用户或话题, 虚曲线表示新生成的参与讨论关系。

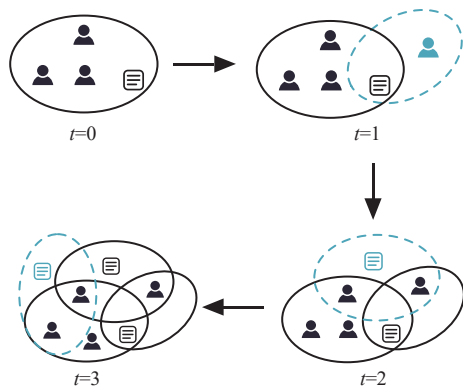


图 3 微博超网络演化过程示意图

2.2 微博超网络中信息传播的影响因素及量化

在微博网络中, 用户可以参与讨论多个不同的

话题, 而信息的接收与进一步传播因话题热度或参与用户的兴趣价值等因素动态变化。基于超图的超网络能够更精准地描述微博用户与话题之间的复杂交互行为。因此, 本文在构建底层微博超网络的基础上, 通过量化超网络中异质节点的个体属性, 研究微博超网络中信息的传播过程及规律。

1) 基于邻居和元路径实例熵的用户影响力

用户影响力是指用户在信息传播过程中的影响范围。不同的用户具有不同的影响力, 影响力较大的用户发布或分享的信息可以广泛传播, 扩散速度较快。在异质网络中, 节点的影响力不仅依赖于直接邻居, 还受其在不同元路径下与其他节点的多样化关系的影响。本文利用元路径和熵对微博超网络中用户节点的影响力进行量化。通过计算每个元路径中节点对不同类型邻居的影响概率计算用户的影响力。该方法结合超网络中两个重要的维度: 局部信息和全局信息。局部信息考虑了直接邻居的信息, 全局信息通过元路径反映与间接邻居之间的影响。

对于单介体元路径, 节点 A_i 到达节点 A_j 的概率根据 A_i 的邻居节点以及两个节点之间的元路径实例数量进行计算, 定义为:

$$P_{A_i A_j} = \frac{C_{A_i A_j}}{\sum_{E \in N_{A_i}} C_{A_i E}} \quad (4)$$

式中, N_{A_i} 表示节点 A_i 的邻居节点; $C_{A_i A_j}$ 表示节点 A_i 与节点 A_j 的元路径实例数量。

对于多介体元路径, 每条 n 介体元路径可拆解为 n 条单介体元路径, 即元路径 $A_1 R_1 A_2 R_2 A_3 \cdots A_{l-1} R_{l-1} A_l$ 可拆解为 $l-1$ 条单介体元路径 $A_1 R_1 A_2$, $A_2 R_2 A_3, \cdots, A_{l-1} R_{l-1} A_l$, 节点 A_1 到达节点 A_l 的概率 $P_{A_1 A_l}$ 为所有可能路径的概率之和, 每条路径的概率为拆解后单介体元路径的乘积, 定义为:

$$P_{A_1 A_l} = \sum_{\text{path from } A_1 \text{ to } A_l} P_{A_1 A_2} P_{A_2 A_3} \times \cdots \times P_{A_{l-1} A_l} \quad (5)$$

节点 A_i 在不同元路径中的熵值为:

$$H_{A_i} = - \sum_{E \in N_{A_i}} P_{A_i E} \log P_{A_i E} \quad (6)$$

式中, $P_{A_i E}$ 表示节点 A_i 在不同元路径下到达其邻居节点的概率。式 (6) 衡量了节点 A_i 在不同元路径下的不确定性。为了综合节点 A_i 在不同元路径下的影响力, 通过对节点 A_i 在不同元路径下的熵值进行归一化处理, 定义节点 A_i 在元路径 K 下的归一化影响

力值:

$$\text{Info}_{A_i}^K = \frac{H_{A_i}}{\sum_{E \in N_{A_i}^K} C_{A_i E}^K} \quad (7)$$

式中, $C_{A_i E}^K$ 表示节点 A_i 与邻居节点在元路径 K 下的实例数量。最后, 将节点 A_i 在所有元路径下的归一化影响力值相加, 得到该节点的总影响力指标:

$$\text{Info}_{A_i} = \sum_K \text{Info}_{A_i}^K \quad (8)$$

以包括 6 个用户和 16 个话题的小型超网络为例 (如图 2 所示), 表 2 列出不同元路径下节点的路径实例数量 (C)、到达邻居节点的概率 (P)、信息熵 (H) 的值, 空值表示用户之间未通过本文所述的两种元路径相连。

表 2 不同元路径下用户之间的熵值

Nodes	UTU			UTUTU		
	C	P	H	C	P	H
User1- User2	3	0.500	0.151	3	0.063	0.075
User1- User3	2	0.333	0.159	1	0.021	0.035
User1- User4	1	0.167	0.130	11	0.361	0.158
User1- User5	—	—	—	2	0.042	0.058
User1- User6	—	—	—	1	0.021	0.035
User2- User3	—	—	—	9	0.230	0.147
User2- User4	3	0.500	0.151	3	0.083	0.090
User2- User5	—	—	—	6	0.125	0.113
User2- User6	—	—	—	3	0.063	0.075
User3- User4	1	0.333	0.159	2	0.111	0.106
User3- User5	—	—	—	2	0.083	0.090
User3- User6	—	—	—	1	0.042	0.058
User4- User5	2	0.250	0.151	3	0.094	0.096
User4- User6	1	0.125	0.113	6	0.150	0.124
User5- User6	3	0.200	0.133	2	0.050	0.065

2) 用户自信度

用户信心是指用户判断信息的能力。用户的生活环境、宗教信仰、文化程度、职业等都会影响其对信息的本质判断, 从而影响其传播信息的行为。随着教育的普及度逐渐提高, 城市之间的人员流动扩大, 人们接触的文化逐渐多元化, 越来越多的人对获取的信息会有相对理性的判断。文献 [15] 将用户自信度量公式定义为:

$$\text{Cfds}(v_i) = 1 - e^{-w(v_i)} \quad (9)$$

式中, $w(v_i)$ 是用户 v_i 的自身自信度量, 即为一个常数。用户所接受的教育程度越高、所接触的文化越多元化等, $w(v_i)$ 取值越大, 用户自信度越高。

3) 用户兴趣价值

信息在网络中的传播受到信息热点和内容的影

响。用户对不同热点和内容的信息有不同的传播行为, 这与用户的兴趣密切相关。用户兴趣值反映了信息的热度和内容与用户兴趣的匹配程度。微博通过用户点赞、转发、收藏等一系列行为推测用户偏好, 并推荐用户可能感兴趣的内容。由于用户的兴趣不同, 平台为用户推荐的内容也有差异。文献 [28] 假设平台为不同用户推荐相同信息的概率服从正态分布, 定义为:

$$\text{VOI} \sim N(\text{voi}, \sigma_{\text{voi}}^2) \quad (10)$$

式中, voi 表示用户兴趣价值的均值; σ_{voi} 表示用户兴趣价值的标准差。VOI 越大, 用户在信息传播过程中所呈现的价值越高。

4) 信息时效性

根据信息的生命周期理论, 信息在社交网络中的传播可能经历 3 个阶段: 出现、爆发和下降。信息具有时间敏感性, 随着时间的推移, 用户对信息的关注和兴趣逐渐衰减。文献 [29] 假设信息传播者逐渐退出信息传播过程的免疫率随时间呈指数变化, 定义为:

$$1 - e^{-\lambda t} \quad (11)$$

式中, λ 表示网络信息的时效性特征。

2.3 微博超网络上的信息传播模型

微博超网络通过超边结构连接异质节点, 有效刻画了用户与话题间复杂的交互关系, 同时考虑用户影响力、兴趣价值等因素对用户状态转化过程的影响。文献 [15] 提出 UHIR 模型利用超网络结构有效描述用户之间的多维交互关系, 并结合用户属性与信息属性动态刻画信息传播过程中用户状态的转化机制, 该模型能较为准确地描述微博超网络中用户行为及信息传播的群体互动特性。因此, 本文以该模型作为信息传播的底层模型研究微博超网络中的传播机制和动态演化过程。

2.3.1 模型构造

在微博超网络的信息传播过程中, 网络用户被划分为以下 4 种状态。

- 1) 尚未获取信息的不知情者 (S 状态);
- 2) 已经知晓信息, 但暂未进行传播的潜伏者 (E 状态);
- 3) 知晓信息, 并且正在传播信息的知情者 (I 状态);
- 4) 知晓信息, 但无兴趣继续传播的免疫者 (R 状态)。

基于个体属性异质的微博超网络信息传播过程

如图 4 所示，描述如下。

1) 初始时，从超网络中选择一个用户节点为知情节点，该节点为超网络中第 1 个获取信息的知情者，处于 I 状态，其他节点均处于 S 状态；

2) 每个时间步内，超网络中知情节点向其所在超边中的所有邻接用户节点传播信息。其中，不知情节点在知情节点的影响力、自信度和自身兴趣价值的影响下，以概率 α 知晓信息，但暂未传播该信息，转化为潜伏节点；或以概率 β 知晓信息，并向不知情节点传播该信息，转化为知情节点。潜伏节点在其自身自信度、自身兴趣价值的影响下，以概率 η 传播信息转化为知情节点；或以概率 ξ 知晓信息，但不传播该信息，转化为免疫节点。随着时间的推移，知情节点对信息的新鲜度会逐渐下降，以概率 γ 终止对信息的传播行为，转化为免疫节点。信息传播过程中，节点的最终状态为免疫状态，处于免疫状态的节点不再发生任何变化；

3) 随着信息在超网络中不断传播，处于各种状态的节点所占密度将会达到一个相对稳定值，并在稳定值附近进行小幅度波动，表明超网络中的信息传播过程达到稳态。

图 5 所示为微博信息超网络传播示意图。其中，黑色节点、蓝色节点、红色节点、灰色节点分别代表处于 S 状态、E 状态、I 状态、R 状态的用户节点。 $t=0$ 时刻，超网络中的所有用户节点为 S 状态； $t=1$ 时刻，从超网络中任意选择一个用户节点 u_4 作为初始传播节点； $t=2$ 时刻，由于节点 u_4 参与讨论话题 t_1 、 t_2 、 t_3 ，因此其邻接节点 u_2 获得话题信息，但暂未传播处于 E 状态，节点 u_6 获得话题信息并正在传播信息处于 I 状态，而 u_4 失去了传播兴趣或能力处于 R 状态； $t=3$ 时刻，节点 u_2 传播信息处于 I 状态，由于 u_2 参与讨论话题 t_1 、 t_2 ，故其邻接节点 u_1 获得信息，但暂未传播信息处于 E 状态，节点 u_6 失去了传播兴趣或能力处于 R 状态。

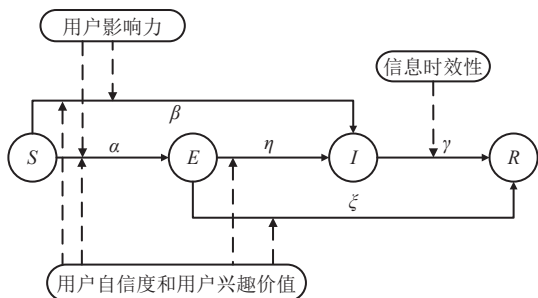


图 4 SEIR 模型

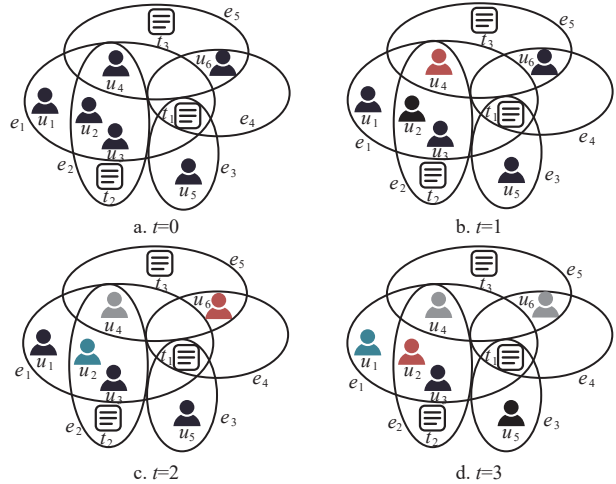


图 5 微博信息超网络传播示意图

2.3.2 模型参数分析

状态转移概率是网络信息传播过程中用户状态变化的重要参数。微博超网络中用户影响力、用户自信度、用户兴趣价值和信息时效性等影响因素的差异性，导致了用户状态变化的复杂性。因此，将考虑影响因素的状态转移概率定义如下。

δ : 考虑用户影响力、用户自信度、用户兴趣价值等影响因素，微博超网络中不知情节点转化为潜伏节点的概率表示为：

$$\alpha(v_i, v_j) = \begin{cases} \text{VOI} \times \text{Cfds}(v_j) \times \text{Info}(v_i), & \text{VOI} \times \text{Cfds}(v_j) \times \text{Info}(v_i) \leq 1 \\ 1, & \text{其他} \end{cases} \quad (12)$$

β : 微博超网络中不知情节点转化为知情节点的概率表示为：

$$\beta(v_i, v_j) = \begin{cases} 1 - \alpha(v_i, v_j), & 1 - \alpha(v_i, v_j) > 0 \\ 1, & \alpha(v_i, v_j) = 0 \end{cases} \quad (13)$$

η : 考虑用户自信度和用户兴趣价值等影响因素，微博超网络中潜伏节点转化为知情节点的概率表示为：

$$\eta(v_j) = \begin{cases} \text{VOI} \times \text{Cfds}(v_j), & \text{VOI} \times \text{Cfds}(v_j) \leq 1 \\ 1, & \text{其他} \end{cases} \quad (14)$$

ξ : 微博超网络中潜伏节点转化为免疫节点的概率表示为：

$$\xi(v_j) = \begin{cases} 1 - \eta(v_j), & 1 - \eta(v_j) > 0 \\ 1, & \eta(v_j) = 0 \end{cases} \quad (15)$$

γ : 由于信息时效性，微信超网络中知情节点转化为免疫节点的概率表示为：

$$\gamma(t+1) = 1 - e^{-\lambda t} \quad (16)$$

式中, 参数 λ 表示网络信息时效性特征。

3 仿真结果分析

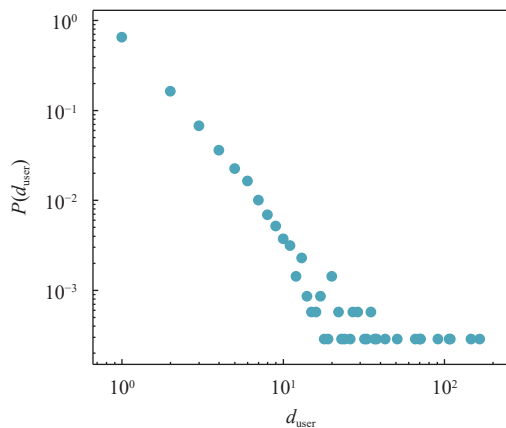
本文通过计算机仿真构建微博超网络模型, 将仿真结果与实证结果进行对比, 验证所提模型的有效性。此外, 仿真实验进一步分析超度分布、不同网络结构、用户影响力、置信度、兴趣价值和信息时效性下信息传播的动态过程和规律。为消除随机性影响, 每组仿真结果均为独立重复运行 50 次的平均值。

3.1 超网络的超度分布

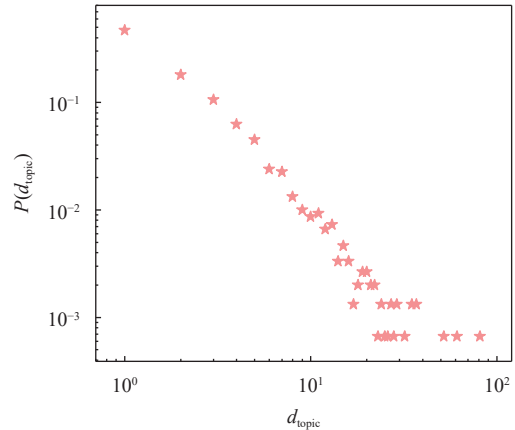
为验证超度分布的正确性, 本文构建了一个参数为 $N = 5000$, $N_u : N_t = 7 : 3$, $m_0 = 5$, $m_1 = 1$, $m_2 = 3$ 的超网络。由于微博超网络中包含用户、话题两类异质节点, 故超度分布需从用户节点的讨论量和话题节点的参与用户数两个维度进行分析。

图 6a 为双对数坐标下微博超网络中用户参与讨论量的超度分布图, 结果表明用户参与讨论量遵循幂律分布, 具有无标度特征。在真实社交网络中, 少数活跃用户 (如意见领袖、名人) 参与大量讨论, 推动了信息的快速传播, 而大多数用户仅参与少量讨论, 扮演信息的接收者, 呈现出社交网络信息传播的“长尾效应”。

图 6b 为双对数坐标下微博超网络中话题参与量的超度分布图。结果表明话题参与量遵循幂律分布, 具有无标度特征。在真实社交网络中, 少数热门话题能够吸引大量用户参与讨论, 而大多数话题的参与量较少, 导致传播范围受限。



a. 用户讨论量的超度



b. 话题参与量的超度分布

图 6 微博超网络的超度分布图

3.2 不同网络结构动态仿真分析

为研究微博超网络中信息传播的过程及规律, 选择网络中超度值最大的用户节点为初始知情节点。模拟微博超网络中最大超度值为 165; BA 无标度网络中最大度值为 241; NW 小世界网络中最大度值为 16, 其他仿真参数设置如式 (17) 所示:

$$\begin{aligned} N = 5000, N_u : N_t = 7 : 3, \alpha = 0.3, \\ \beta = 0.7, \eta = 0.7, \xi = 0.3 \end{aligned} \quad (17)$$

图 7 为不同网络结构下各状态节点密度随时间变化的信息传播曲线图。仿真结果表明, 在基于超图的微博超网络和 BA 无标度网络中, 知情者 (I 态) 数量的变化呈现出快速增加后缓慢减少的变化趋势, 而在 NW 小世界网络中, 知情者数量的变化趋势呈现正态分布的特征, 究其原因, 可能由于小世界网络中节点度分布为泊松分布的缘故。此外, 不同网络结构达到传播峰值所需的时间逐渐增加, 传播峰值逐渐降低, 原因在于: 微博超网络中超边的所有用户均被视为邻居节点, 信息能够快速广泛地传播; BA 无标度网络由于存在少量高连接度节点, 在初期传播迅速, 但普通节点之间的弱连接性导致传播范围受限; NW 小世界网络的局部聚类性限制了信息传播的速度和范围。

在真实微博网络中, 信息的传播具有突然爆发、快速消退及生命周期短的特点^[30]。当某个话题刚被发布时, 绝大部分用户尚未关注到该话题, 参与讨论的用户比例较低; 随着时间推移, 话题的传播范围扩大, 更多用户开始关注并参与讨论, 话题热度迅速攀升并达到峰值; 然而, 由于话题具有一定的时效性, 许多用户逐渐失去兴趣并停止讨论, 话题热度逐渐衰减, 最终趋近于零或完全无人讨

论；随着话题在网络中传播时间延长，失去兴趣或停止参与的用户比例增加，网络中大部分用户变为“免疫”状态，对该话题不产生任何反应。由此可

知，微博超网络的传播规律更契合真实的微博信息传播特征。因此，本文后续仿真实验均在此模型上进行实验。

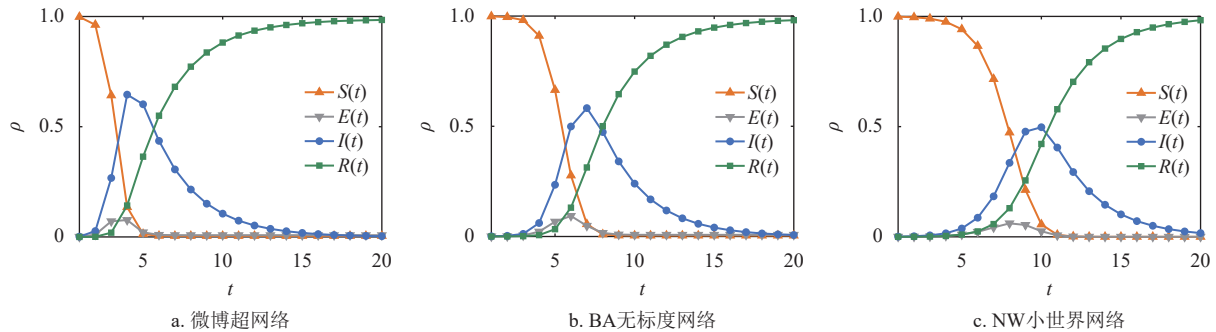


图 7 不同网络结构下的信息传播曲线

3.3 信息时效性对信息传播的影响

为研究信息时效性对网络信息传播过程的影响，选择超网络中节点超度值为平均超度的一个用户作为初始传播者，将网络信息时效性特征 λ 分别设置为 0.1、0.5、1，其他仿真参数设置如式 (18) 所示：

$$N = 5\,000, N_u : N_t = 7 : 3, \alpha = 0.3, \beta = 0.7, \eta = 0.7, \xi = 0.3 \quad (18)$$

图 8 所示为超网络中不同信息时效性特征 λ 下各状态节点密度随时间变化的信息传播曲线。结果

表明，随着 λ 的增大，知情节点 (I 态) 到达传播峰值所占密度逐渐减小，最终达到稳态所占密度为 0；免疫节点 (R 态) 到达稳态所需时间逐渐减小，分别在第 13、11、10 步达到稳态，达到稳态所占密度逐渐减小。分析其原因，由于 λ 的增大提高了知情节点转化为免疫节点的概率，信息传播的生命周期缩短，用户对话题的关注度迅速下降，导致信息传播快速衰退。在真实微博网络中，随着时间的推移，用户对同一话题的关注度会逐渐下降，话题的热度也会逐渐下降，最终可能无人讨论。

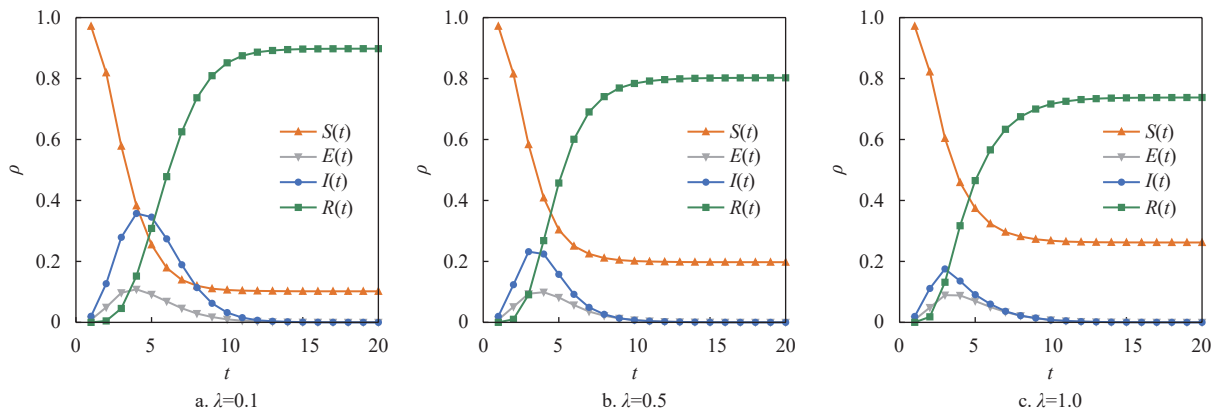


图 8 不同信息时效性特征 λ 下的信息传播曲线

3.4 用户自信度对信息传播的影响

为研究用户自信度对信息传播过程的影响，随机选择微博超网络中平均超度值的用户为初始知情者，将自信度大于 0.6 的用户所占比例分别设置为 8%、50%、90%，其他仿真参数设置如下所示：

$$N = 5\,000, N_u : N_t = 7 : 3, \lambda = 1, \text{VOI} \sim (1, 0.1^2) \quad (19)$$

图 9 所示为超网络中不同用户自信度下各状态

节点密度随时间变化的信息传播曲线。结果表明用户自信度对网络信息传播过程具有显著影响，随着网络中高自信度用户比例的增加，不知情节点 (S 态) 和潜伏节点 (E 态) 密度逐渐增大，知情节点 (I 态) 密度峰值逐渐减小，免疫节点 (R 态) 到达稳态时所占密度逐渐减小。其原因在于高自信度用户能够更加理性地判断信息的价值和真实性，从而减少低质量信息的传播，导致信息传播速度减

缓, 信息传播范围缩小。在微博网络信息传播过程中, 高自信度用户通常为具有较高理性判断能力和较强信息筛选能力的用户, 如意见领袖、专家用户

或受过较高教育的用户。当高自信度用户比例增大时, 虚假信息和无意义信息扩散被抑制, 从而降低信息传播的概率和速度, 且传播范围减小。

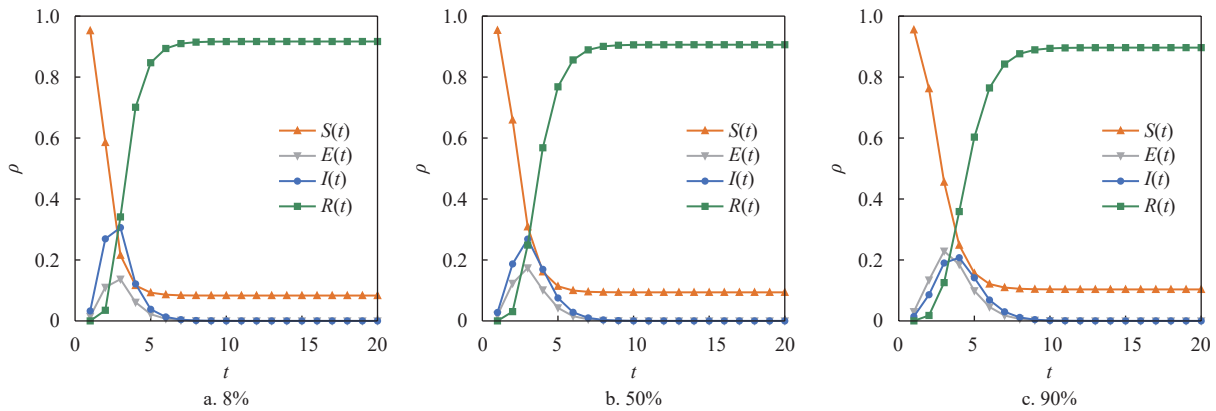


图 9 不同用户自信度所占比例下的信息传播曲线

3.5 用户兴趣价值对信息传播的影响

为研究用户兴趣价值对网络信息传播过程的影响, 随机选择微博信息超网络中平均超度值的用户为初始知情者, 将自信度大于 0.6 的用户所占比例设置为 8%, 将网络用户兴趣价值的均值 voi 分别设置为 0.1、0.5、1, 其他仿真参数设置如下所示:

$$N = 5\,000, N_u : N_t = 7 : 3, \lambda = 1 \quad (20)$$

图 10 所示为超网络中不同用户兴趣价值下各状态节点密度随时间变化的信息传播曲线。结果表明在信息传播过程中, 用户兴趣价值对信息传播的速度和传播范围的影响较大。随着网络用户兴趣价

值逐渐增大, 潜伏节点 (E 态) 达到峰值所占密度逐渐减小; 知情节点 (I 态) 所到达的峰值逐渐增大; 免疫节点 (R 态) 到达稳态的时间逐渐减小, 其所占最终密度逐渐增大, 且略低于 1。该现象产生的原因在于用户的兴趣价值通常与其对信息的关注程度和参与度密切相关。当用户对某一话题具有较高的兴趣价值时, 他们更倾向于关注、接受和分享信息, 因此对信息具有较高兴趣值的用户更容易成为信息的传播源、快速从易感节点转化为知情节点, 促进信息的传播。在真实微博网络中, 用户发布或分享与某话题相关的信息时, 对该话题感兴趣的邻居用户越多, 信息被分享的概率越大, 信息传播速度越快, 信息传播的范围越广。

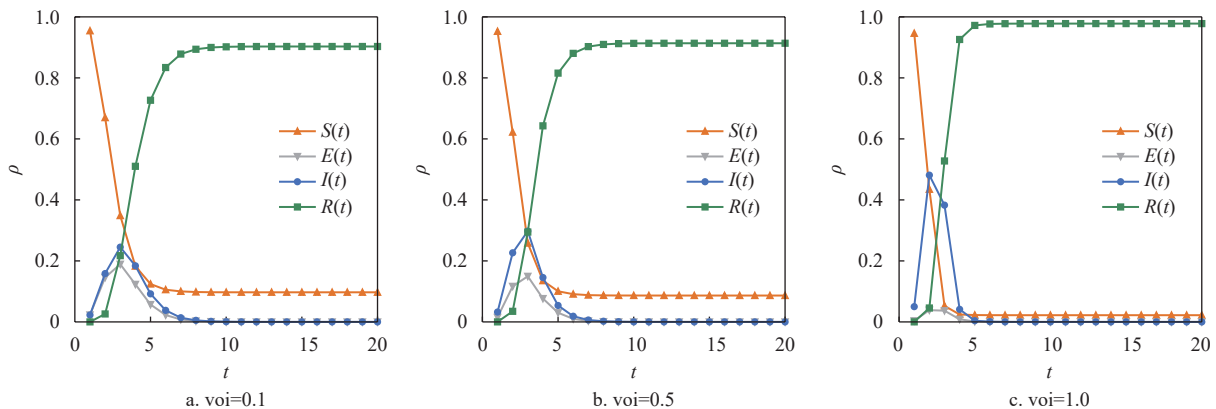


图 10 不同用户兴趣价值下的信息传播曲线

3.6 与同质超网络的信息传播对比

对比信息在同质超网络与微博信息超网络结构下不同状态节点密度随时间变化的传播过程, 按照

文献 [27] 构建一个参数分别为 $N = 5\,000$ 、 $m_0 = 5$ 、 $m_1 = 1$ 、 $m_2 = 3$ 的同质超网络, 图 11 所示为不同超网络结构下的信息传播曲线。结果表明在微博信息

超网络传播初期，知情节点 (I 态) 所到达的峰值逐渐增大，表明信息能迅速通过异质节点 (如用户、话题) 的多元连接传播更多群体，且免疫节点 (R 态) 所占最终密度增大，表明信息能够覆盖更

广泛的用户群体。在真实社交场景中，用户往往因共同关注特定话题而形成传播群体，这种现象在传统同质网络中难以呈现。因此，微博信息超网络模型能更好地模拟真实社交网络的信息传播。

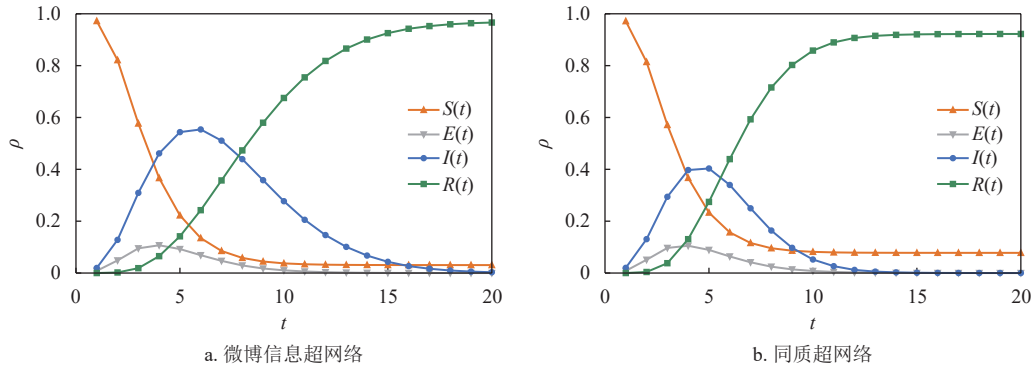


图 11 不同超网络结构下的信息传播曲线

4 实证结果分析

4.1 数据收集及预处理

2023 年 8 月日本政府宣布正式开始排放核污水，此举引发国际社会的广泛关注和争议 (<https://www.chinanews.com.cn/gj/2023/08-24/10066383.shtml>)，特别是在社交媒体平台上，微博用户围绕该事件展开激烈讨论，涵盖环境保护、国际政治、健康风险等话题，体现个体多样性及话题传播路径的复杂性，为系统动态地观察公众对事件讨论的变化提供了良好的研究基础，故本文选择该事件进行实证分析。通过搜索与“日本核污水”相关的 975 个话题，以讨论量为有效字段并抓取每个话题下参与讨论的用户信息进行仿真实验，并去除无意义内涵的高频词汇、标点、数字、字母。该数据集涵盖事件的开始、高峰和消退的过程，作为后续数据处理、分析和论述的基础语料库。部分微博话题数据及讨论量如表 3 所示。

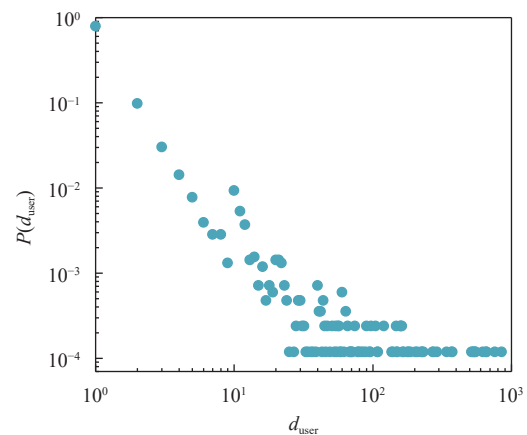
表 3 部分微博话题数据及讨论量

话题	讨论量
日本核污水	5.5万
日本东电首次开放核污水处理设施	1 625
日本核污水57天将污染半个太平洋	13.1万
日本核污水将于下月抵达中国海域	150
日本核污水排海意味着什么	387
⋮	⋮

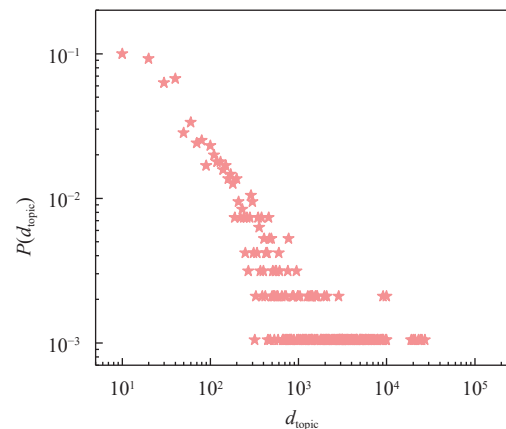
4.2 对比分析

根据数据绘制出如图 12 所示双对数坐标下用户节点和话题节点的超度分布图，该结果遵循幂律分布，具有无标度特性，与图 6 的模拟结果一致，

说明微博超网络符合无标度超网络特性，契合于作为信息传播的底层网络，因此本文在该数据集上进行了信息传播研究。



a. 用户讨论量的超度分布



b. 话题参与量的超度分布

图 12 用户节点和话题节点的超度分布在 SEIR 模型中， I 状态节点代表一个传播信

息的节点, 对应于现实生活中参与话题讨论的用户, 因此, 本文统计在事件期间每天参与话题讨论的用户比例, 结果如图 13a 所示, 事件传播者的比例呈现出先快速增长, 后快速下降, 再缓慢消散的趋势。事件初期, 信息仅在小范围内传播, 随着拥有大量粉丝的用户参与, 信息传播速度迅速增加, 最后逐步衰减直至消散。对比 SEIR 模型仿真结果中的 I 状态曲线 (图 13b) 可知, 两者的变化趋势和峰值具有高度相似性, 从而验证本文模型的有效性。

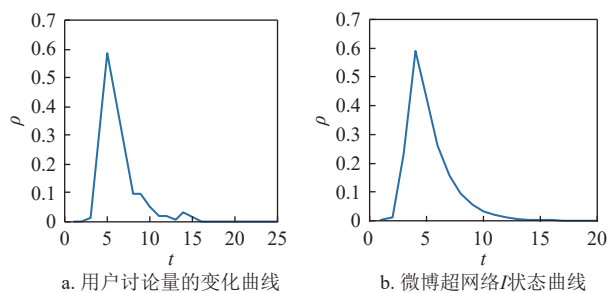


图 13 用户讨论量与 SEIR 模型 I 状态节点的比较

5 结束语

为描述真实微博网络中用户与话题间复杂的交互关系, 基于改进的 BA 无标度超网络模型, 考虑用户影响力、自信度、兴趣价值和信息时效性等个体属性, 结合 SEIR 传播模型, 提出构建个体属性异质的微博信息超网络传播模型。通过仿真实验分析不同影响因素对信息传播过程及规律的影响, 结果表明用户讨论量和话题讨论量的超度分布均遵循幂律分布, 具有无标度特性; 信息内容与用户兴趣值一致性越高, 高自信度用户所占比例越小, 信息传播速度更快, 范围更广; 网络中旧信息的快速更新或被新信息覆盖会减小信息的扩散范围。最后, 以“日本核污水”话题为例, 绘制传播趋势图, 并与仿真结果进行比较, 两者在传播趋势和峰值上具有较高的一致性, 进一步验证了该模型的有效性。该研究为舆情监控、传播干预等实际应用提供了理论依据。然而, 该模型尚未充分考虑更复杂的个体行为特征 (如情感倾向、个体兴趣等), 未来研究可以在此基础上分析多种因素影响下的信息全局传播过程, 以提升模型对信息传播动态变化的准确性。此外, 研究多层次传播规律和不同类型节点间的交互关系, 并探索信息传播在时间维度上的演化, 将为揭示信息传播的全局规律提供更全面的视角。

参考文献

- [1] 沈旺, 时倩如, 王俊尧, 等. 基于超图的在线社交网络信息传播模型研究[J]. 情报学报, 2023, 42(3): 354-364.
SHEN W, SHI Q R, WANG J Y, et al. Information dissemination model used in online social networks based on hypergraphs[J]. Journal of the China Society for Scientific and Technical Information, 2023, 42(3): 354-364.
- [2] LUO T Y, XU D, CAO Z D, et al. Modeling the coupling propagation of information, behavior, and disease in multilayer heterogeneous networks[J]. IEEE Transactions on Computational Social Systems, 2024, 11(3): 4058-4070.
- [3] 康泽, 王建波, 阳成, 等. 控制策略对流行病首达时间的影响研究[J]. 电子科技大学学报, 2024, 53(1): 119-128.
KANG Z, WANG J B, YANG C, et al. Study on the influence of control strategy on the first arrival time of epidemics[J]. Journal of University of Electronic Science and Technology of China, 2024, 53(1): 119-128.
- [4] YAN D C, XIE W X, ZHANG Y W. Heterogeneous information network-based interest composition with graph neural network for recommendation[J]. Applied Intelligence, 2022, 52(10): 11199-11213.
- [5] MOSCATO V, SPERLÌ G. A survey about community detection over on-line social and heterogeneous information networks[J]. Knowledge-Based Systems, 2021, 224: 107112.
- [6] LI T T, ZENG Z M, SUN S Q, et al. Constructing a multi-layer heterogeneous networks model to explore the public opinion evolution pattern of key users in public health emergencies[J]. Journal of Information Science, 2023, DOI: 10.1177/01655515231169953.
- [7] 李若晨, 肖人彬. 基于改进狼群算法优化 LSTM 网络的舆情演化预测[J]. 复杂系统与复杂性科学, 2024, 21(1): 1-11.
LI R C, XIAO R B. Public opinion evolution prediction based on LSTM network optimized by an improved wolf pack algorithm[J]. Complex Systems and Complexity Science, 2024, 21(1): 1-11.
- [8] KERMACK W O, MCKENDRICK A G. A contribution to the mathematical theory of epidemics[J]. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character, 1927, 115(772): 700-721.
- [9] ANDERSON R M. The population dynamics of infectious diseases: theory and applications[M]. [S.l.]: Springer, 2013.
- [10] ZHANG M L, QIN S M, ZHU X X. Information diffusion under public crisis in BA scale-free network based on SEIR model: Taking COVID-19 as an example[J]. Physica A: Statistical Mechanics and Its Applications, 2021, 571: 125848.
- [11] HOSSEINI S, ZANDVAKILI A. Information dissemination modeling based on rumor propagation in online social networks with fuzzy logic[J]. Social Network Analysis and Mining, 2022, 12(1): 34.
- [12] NAGURNEY A, DONG J E, ZHANG D. A supply chain network equilibrium model[J]. Transportation Research Part E: Logistics and Transportation Review, 2002, 38(5): 281-303.

- [13] ESTRADA E, RODRÍGUEZ-VELÁZQUEZ J A. Subgraph centrality and clustering in complex hyper-networks[J]. *Physica A: Statistical Mechanics and Its Applications*, 2006, 364: 581-594.
- [14] XIAO H B, HU F, LI P Y, et al. Information propagation in hypergraph-based social networks[J]. *Entropy*, 2024, 26(11): 957.
- [15] GONG Y C, WANG M, LIANG W, et al. UHIR: An effective information dissemination model of online social hypernetworks based on user and information attributes[J]. *Information Sciences*, 2023, 644: 119284.
- [16] WANG J, WANG Z P, YU P, et al. The SEIR dynamic evolutionary model with Markov chains in hyper networks[J]. *Sustainability*, 2022, 14(20): 13036.
- [17] MOLAEI S, BABAEI S, SALEHI M, et al. Information spread and topic diffusion in heterogeneous information networks[J]. *Scientific Reports*, 2018, 8: 9549.
- [18] YI Y X, ZHANG Z F, YANG L T, et al. Reemergence modeling of intelligent information diffusion in heterogeneous social networks: The dynamics perspective [J]. *IEEE Transactions on Network Science and Engineering*, 2021, 8(2): 828-840.
- [19] MOLAEI S, ZARE H, VEISI H. Deep learning approach on information diffusion in heterogeneous networks[J]. *Knowledge-Based Systems*, 2020, 189: 105153.
- [20] LI Z, ZHANG J, ZHENG G Z, et al. A generative hypergraph model for double heterogeneity[J]. *Journal of Complex Networks*, 2023, 12(1): cnad048.
- [21] MOLTMAN G A W. Heterogeneous hypergraphs models for collaboration networks: A bridge between mathematics and business[D]. Enschede: University of Twente, 2023.
- [22] BERGE C. Packing problems and hypergraph theory: A survey[J]. *Annals of Discrete Mathematics*, 1979, 4: 3-37.
- [23] BAYTAS I M, XIAO C, WANG F, et al. Heterogeneous hyper-network embedding[C]//Proceedings of the IEEE International Conference on Data Mining. New York: IEEE, 2018: 875-880.
- [24] ZHU Y, GUAN Z Y, TAN S L, et al. Heterogeneous hypergraph embedding for document recommendation[J]. *Neurocomputing*, 2016, 216: 150-162.
- [25] FU X Y, ZHANG J N, MENG Z Q, et al. MAGNN: Metapath aggregated graph neural network for heterogeneous graph embedding[C]//Proceedings of the Proceedings of the Web Conference 2020. New York: ACM, 2020: 2331-2341.
- [26] SHANNON C E, WEAVER W. The mathematical theory of communication[M]. Urbana: University of Illinois Press, 1949.
- [27] HU F, ZHAO H X, MA X J. An evolving hypernetwork model and its properties[J]. *Scientia Sinica Physica, Mechanica & Astronomica*, 2013, 43(1): 16.
- [28] GENG L X, YANG S R, WANG K, et al. Modeling public opinion dissemination in a multilayer network with SEIR model based on real social networks[J]. *Engineering Applications of Artificial Intelligence*, 2023, 125: 106719.
- [29] WANG C X, WANG G S, LUO X P, et al. Modeling rumor propagation and mitigation across multiple social networks[J]. *Physica A: Statistical Mechanics and Its Applications*, 2019, 535: 122240.
- [30] LI L F, ZHANG Q P, WANG X, et al. Characterizing the propagation of situational information in social media during COVID-19 epidemic: A case study on weibo[J]. *IEEE Transactions on Computational Social Systems*, 2020, 7(2): 556-562.

责任编辑 叶 芳