

大数据背景下的社交网络谣言传播研究^①

陈旭华¹ 潘雅婷²

¹ (福州大学图书馆 福建 福州 350108)

² (福州大学图书馆 福建 福州 350108)

摘要 [目的] 大数据背景下, 在线社交网络的发展加剧了谣言信息的危害性, 因而在大数据背景下对谣言传播机制进行研究越来越重要。[方法] 在经典 SIR 谣言传播模型的基础上, 尝试建立了基于谣言传播者分类的 SIIR 模型, 并应用数值模拟的方法分析了 SIIR 模型所描述的谣言传播过程。[结果] SIIR 模型更符合谣言信息的实际传播过程, 且表明谣言传播周期与扩散范围的影响因素。[局限] 有待进一步扩充微博等社交平台的真实谣言数据规模进行实证研究。[结论] SIIR 谣言传播模型为控制大数据背景下的社交网络谣言传播影响提供了依据。

关键词 大数据; 谣言传播; SIR 模型; 社交网络

Research on Social Network Rumor Propagation in the Background of Big Data

Chen Xuhua¹ Pan Yating²

¹ (Library of Fuzhou University, Fuzhou, Fujian, 350108, China)

² (Library of Fuzhou University, Fuzhou, Fujian, 350108, China)

Abstract [Objective] In the context of big data, the development of online social networks intensifies the harmlessness of rumor information, so it is more and more important to study the rumor propagation mechanism in the context of big data. [Methods] On the basis of the classical SIR rumor propagation model, a SIIR model based on the classification of rumor disseminators is established, and the process of rumor propagation described by SIIR model is analyzed by means of numerical simulation. [Results] The SIIR model is not only more consistent with the actual propagation process of rumor information, but also shows the

①本文系福建省教育厅中青年课题基金项目“大数据背景下的社交网络谣言传播研究”(项目编号: JZ180198)的研究成果之一。

influencing factors of rumor propagation cycle and diffusion range. **[Limitations]** It is necessary to further expand the scale of real rumor data on weibo and other social platforms for empirical research. **[Conclusions]** SIIR rumor propagation model provides a basis for controlling the influence of rumor propagation on social networks in the context of big data.

Keywords Big Data; Rumor Spreading; SIIR Model; Social Networks

1 引言

在网络技术不断发展的今天, 社交网络已经成为了当代人生活中不可或缺的一部分。抖音、微博、微信等新社交形式的出现, 进一步推动了社交平台的影响力, 加速了人类活动中信息流的传递。作为沟通交流与发表观点的平台, 社交网络促进了信息流通和不同观点的传播, 扩宽了群众的视野, 同时带来了负面影响。谣言的存在与传播, 是在线社交网络在发展过程中, 无法避免的现象之一。谣言通过社交网络中个体间的相互沟通、交流进行扩散传播, 给人类社会带来了极大的危害性。在社交网络的信息传播过程中难免会存在谣言问题, 为了提高社交网络平台信息传播的真实性与有序性, 减少谣言在传播过程中带来的危害性, 应当研究时代背景下的社交网络谣言传播特性及其传播机制, 以便寻求合理的方式抑制谣言所带来的不良影响。

目前, 在现有的研究课题中关于谣言传播的研究, 主要集中在谣言的识别、成因、传播模型与传播机制这几方面, 其中, 传播模型的建立与研究是一个核心问题。通过对收集到的相关文献的梳理分析, 笔者发现各学者关于社交网络中谣言传播机制的研究很大一部分都是借鉴传染病模型或者从复杂网络理论的角度建立谣言传播模型。

传染病感染与扩散机制同现实中谣言在社交网络中的传播过程有着极大的相似性, 因而现有的模型中存在一定数量的谣

言传播模型是在传染病模型基础上建立起来的。在传染病模型中, 存在的个体被划分为几类, 每一类都处于一种典型状态, 包括: S (Susceptible) ——易感状态, 或健康状态; I (Infected) ——感染状态; R (Recovered) ——被移除状态或免疫/恢复状态^[1]。常见的传染病模型根据个体在几种不同状态间的转换关系, 分为 SI、SIR、SIRS 等。在谣言传播规律的研究中, 社交网络结构与微观个体间的差异性是两个重要的影响因素, 因此, 学者们将复杂网络理论引入到谣言传播研究中, 为谣言传播研究提供了新的思路。Zanette D H^[2], Buzana^[3], Moreno Y^[4]等人分别在小世界网络、无标度网络上进行谣言传播模型的研究。

2 大数据背景下社交网络谣言传播变化分析

“大数据”一词最早出现于美国知名公司麦肯锡, 是指海量、高增长率和多样性的信息数据, 此类数据集合无法轻易地用常规软件工具进行采集、管理和分析。在数字化时代下, 数据类型除了传统的结构数据以外, 还包括海量的非结构化数据和交易数据, 大数据技术的发展使得这些数据的价值在各行业领域内充分体现。大数据技术能够在较短的时间内, 对非关系型数据进行异质性处理, 通过数据分析与挖掘, 从海量的、异构的数据中提取价值, 大数据技术将是IT领域新一代的技术与架构^[5]。

谣言传播作为信息传播的一种特殊类型,与大数据有着密切的联系,基于大数据背景进行谣言传播研究是紧跟时代特点、有利于更深刻地理解社交网络谣言的传播机制的研究。结合大数据进行谣言传播研究,能够得到更为精准、贴近现实的谣言传播模型;通过对社交网络中的大数据分析,能够验证建立的谣言传播模型;此外,大数据技术能够对信息进行分类、追踪、识别和查询,将社交网络中的谣言从源头发布到传播扩散进行记录与追踪,对实时监控和预测社交网络中的谣言传播情况有重要作用。因此,将大数据与谣言传播结合起来,是符合时代特点的科学研究,也能够满足新时代下对于降低谣言危害性的要求。在大数据背景下,社交网络谣言传播发生了新变化,下文将从传播角色、传播形式、谣言传播成因三方面进行分析。

2.1 传播角色分析

在社交网络信息传播中,大多数个体既是传播者又是受众。Web 2.0技术、自媒体的发展使得谣言传播过程中,受与传双方角色有着更为频繁的交互性改变。受众不再仅是被动的信息接收者,更多的信息搜寻与传播行为融入了受众的信息行为中,个人主动传播、圈群式传播逐渐凸显,传与受的身份不再一成不变^[6]。网络环境下,传播过程的匿名性是传播者角色发生变化的另一关键所在。各大社交网络平台提供的匿名功能,给予了网络用户一定的言论自由,同时加速了谣言演变的广度与力度,用户被允许无负担地进行谣言传播。

“意见领袖”是指具有一定人际关系能力、能够获得群众或公众认可的、具有高关注度的少部分人。“意见领袖”是信息传播过程中非常重要的一个角色,能够成为舆论导向的引领者。在虚拟的社交网络平台上,消息灵通

的权威人士常常充当意见领袖,当他们在网络中发声时,往往有很多跟随者。如果“意见领袖”在社交网络谣言的传播过程中受到影响,并对谣言进行支持性的评论、转发等行为时,会使关注他们的用户意见趋向一致,并且进一步促进谣言的扩散。在大数据环境下,各种在线社交网络平台的兴起,“意见领袖”与“中心节点”在数量上与引导力方面都有了质的飞跃。

2.2 传播形式分析

传统媒体时代,谣言传播的呈现形式多以文字、图片、口头语言为主,在表现形式与传播渠道等方面都有着极大限制。而现代基于在线社交网络的谣言传播突破了传统媒体依靠单一符号或以某一符号为主要载体的局限性,实现了多符号交融,多选择的传播形式,谣言内容信息更具有直观性与说服力^[7]。现代网络社交平台,如:微博、知乎、抖音等,都是网络用户自主注册的空间,所体现的是一种以公众为中心的新型网络传播形式。社交媒体中具有高度自由性的用户,更容易被以短视频、短消息、动图、音频等为载体的谣言信息所蒙骗,从而成为谣言传播者中的一员。除此之外,社交媒体的非线性传播形式也是谣言得以大范围扩散的客观条件之一,较之线性传播,非线性传播充分揭示了人类传播的互动性质与双向性。因而,当谣言信息在大数据融入的社交媒体传播时,不仅在外在形式上有了更为多样化的呈现,而且也改变了传统时代单一的线性传播形式。

2.3 谣言传播成因分析

对谣言成因和动机进行分析是研究谣言传播的内在规律的起点。在线社交网络平台上是重要的舆论与信息传播阵地,伴随着互联网的发展与用户数量的上涨,通过互动传播模式引

发的网络谣言有着更为复杂的形成原因。主要从主观与客观两个方面审视谣言传播的成因。从主观的角度来看,传播主体基于自身各种需求而进行的谣言传播,其中包括出于缓解焦虑、释放压力的宣泄需求,同他人进行信息交流的沟通需求,以及传播主体的各种娱乐、猎奇心理等。从客观的角度而言,谣言传播的成因涵盖了几个方面,主要有:新媒体技术的发展为谣言提供了良好的物理环境;谣言频发领域(食品安全、自然灾害、经济政治……)本身具有的高关注度;社交网络中信息总量与用户信息量之间的不对称性。正是因为不同行为主体都有强烈发布或传播网络谣言的动力,所以谣言会呈现出在短时间内大面积扩散的情况^[8]。因而,在时代背景下,应当利用大数据技术及其特征对谣言传播的成因进行深入分析,通过剖析谣言成因,从源头上抑制谣言传播。

3 社交网络谣言信息传播模型的构建

SIR模型作为传染病经典模型,被用于谣言传播领域,主要描述了谣言信息的传播过程,在该过程中存在三种不同状态的个体——易感状态、感染状态、免疫状态,分别以Susceptible、Infected、Removed表示,处于易感状态的个体(S)未接触过谣言,但在接触谣言信息后易转变为感染状态;感染状态个体(I)对于谣言信息持相信态度,并具有传播谣言行为;免疫状态个体(R)接收过谣言信息,但停

止不再参与谣言传播。

SIR模型描述的谣言信息传播过程为:最初,所有未接触过谣言信息的个体都为S。接着,当个体S接触到谣言信息或处于传播状态的个体I时,个体S以一定概率转变为个体I,进入感染状态,进行谣言信息传播。随着时间的推移,由于辨别出信息的虚假性或对其失去兴趣等,个体I不再传播谣言信息,转变为具有免疫性的个体R。

3.1 一种基于传播者分类的改进 SIR 谣言传播模型

现代社交网络中,主要以在线社交平台为信息的主要传播地,在谣言信息的传播过程中,普通公众的参与率日益增加,信息传播者的类型也更为多样化,经典SIR模型中对于传播个体的简单抽象已无法满足对于实际谣言信息传播过程的详细分析。在实际的谣言信息传播中,参与个体本身具有一定的知识背景和生活经验,因而不同个体在面对同一谣言信息时所表现的态度与后续行为是不同的。对于社交网络中个体接触谣言信息后持有的不同态度与行为,笔者尝试对经典SIR模型进行改进,将接触谣言信息后的处于感染状态的个体分为两大类:积极传播者(I_a),接触并传播谣言信息;消极传播者(I_b),接触谣言信息后选择不传播谣言信息或传播辟谣信息。对于上述基于传播者分类的改进模型,将其称为SIIR模型,模型中所描述的谣言传播过程如图1所示。

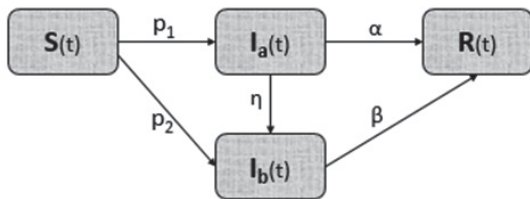


图1 SIIR模型的个体状态转换图

根据图1所示, SIIR模型中处于各状态之间的个体的转换情况为:

(1) 当处于易感状态的个体接触谣言信息或接触积极传播者时, 易感状态个体即未知者以概率 p_1 进行谣言传播, 转化为积极传播者。

(2) 当处于易感状态的个体接触谣言信息或接触消极传播者时, 易感状态个体即未知者以概率 p_2 转化为消极传播者, 消极传播者不传播谣言信息或传播辟谣信息。

(3) 当积极传播者接触传播辟谣信息的消极传播者时, 以概率 η 转化为消极传播者。

(4) 在积极传播者与消极传播者之间, 假设消极传播者转换为积极传播者的概率为0, 原因有二, 当消极传播者由于本身的不活跃性停止谣言传播时, 对传播过程无影响; 当消极传播者由于辨识出谣言虚假性传播辟谣信息时, 则此类消极传播者已经掌握真实信息, 转变为积极传播者的可能性极小, 此处忽略不计。

(5) 随着时间的推移, 积极传播者与消极传播者分别以概率 α 、 β 转化为免疫状态, 不再进行相关信息的传播, 且不再受谣言信息的影响。

3.2 参数确定与描述

对SIIR社交网络谣言传播模型中一些假设和参数进行描述, 如下:

(1) 假设在谣言初始阶段, 社交网络中存在一定比例的初始谣言信息传播者 I_a , 并且社交网络中的个体, 只受其邻居节点影响。

(2) p_1 : 在 t 时刻, 社交网络中个体 S 转变为个体 I_a 的概率, 该参数的影响因素包括: 个体的自身条件(包括: 生活经验、知识水平、性格)、谣言信息与个体间的相关性、传播者与接受个体间的社交关系度、媒体作用。 p_2 :

在 t 时刻, 社交网络中个体 S 转变为个体 I_b 的概率, 该参数的影响因素同参数 p_1 。

(3) η : 在 t 时刻, 社交网络中积极传播者 I_a 转化为消极传播者 I_b 的概率, 该参数印象因素包括: 个体自身条件、媒体作用、传播者间的社交关系度、真实信息内涵(即辟谣信息, 包括: 其信息量与可信度)。

(4) α : 社交网络中积极传播者 I_a 转变为免疫个体 R 的概率, 该参数的影响因素包括: 个体与谣言信息间的接触频率、记忆效应、个体自身条件。 β : 消极传播者 I_b 转变为免疫个体 R 的概率, 该参数的影响因素同参数 α 。

3.3 模型的传播动力学方程

(1) 方程的建立。

基于以上的参数及其描述, 构建SIIR模型的传播动力学方程, 描述如下:

$$\frac{dS(t)}{dt} = -p_1 S(t) I_a(t) - p_2 S(t) I_b(t);$$

$$\frac{dI_a(t)}{dt} = p_1 S(t) I_a(t) - \eta I_a(t) - \alpha I_a(t);$$

$$\frac{dI_b(t)}{dt} = p_2 S(t) I_b(t) + \eta I_a(t) - \beta I_b(t);$$

$$\frac{dR(t)}{dt} = \alpha I_a(t) + \beta I_b(t);$$

其中, t 代表 t 时刻, $S(t)$ 表示社交网络中易感个体的密度, $I_a(t)$ 表示积极传播者的密度, $I_b(t)$ 表示消极传播者的密度, $R(t)$ 表示免疫个体的密度。假定上述模型建立在一个总人数不变的社交网络系统中, 人口总数 N 为常数, 由于谣言传播周期较短, 该模型对于个体的死亡率、出生率、迁移率不做考虑, 则有: $N(t)=S(t)+I(t)+R(t)$, 其中 $I(t)=I_a(t)+I_b(t)$ 。

(2) 模型参数影响分析。

在上述建立的SIIR模型的基础上, 借助数值模拟方法分析谣言传播过程与机制, 使用MATLAB软件对SIIR谣言传播算法进行模拟。固定参数 $p_1=0.6$, $p_2=0.3$, $\eta=0.03$,

$\alpha=0.05$, $\beta=0.05$, 此时假定社交网络中不同个体的人数比例分别为 $s(0)=0.97$, $I_a(0)=0.02$, $I_b(0)=0.01$, $r(0)=0$ 。则图2展示了随着时间推

移, 社交网络中未知者、积极传播者、消极传播者、免疫者的人数比例变化。

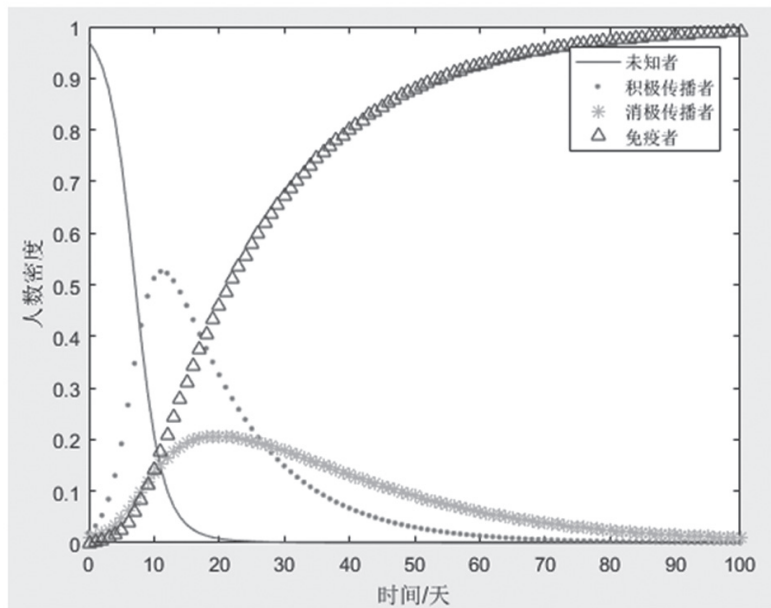


图2 SIIR模型: 不同个体密度变化曲线图

从图中可以看出, 在最初阶段社交网络中存在的大部分个体都处于未知状态, 随着积极传播者的谣言传播行为, 参与信息传播过程的个体逐渐增加, 社交网络中的传播者数量增加 (包括传播谣言信息的积极传播者和传播辟谣信息的消极传播者), 未知者 S 数量减少, 同时在媒体以及相关辟谣信息的作用下, 部分积极传播者 I_a 转化为消极传播者 I_b , 开始传播辟谣信息, 而部分 I_a 则变成免疫者, 不再进行相关信息的传播; 在第二阶段, 社交网络中的积极传播者数量到达最大值后逐渐下降, 而传播辟谣信息与不传播信息的个体继续增加, 且随着谣言信息的范围扩散, 网络中的个体 S 数量持续减少; 最终, 由于时间因素, 消极传播者 I_b 也转化为免疫者, 社交网络仅存在个体 R , 其他状态个体数量都趋向0。

3.4 模型检验

在实际的谣言传播过程中, 不同个体对于接触谣言后的反应存在多种状态, 以真实微博谣言信息的传播为例, 下表罗列了十条微博谣言接触者的反馈信息, 分别以评论量和转发量的形式呈现。如下表1所示, 可以看到微博用户在评论、转发谣言微博时所持态度不尽相同, 其中每条微博存在不同比例的质疑评论。持相信态度的部分用户转发谣言进一步促进谣言的传播, 对于已知真实信息的微博用户 (不相信谣言) 通过转发和评论微博影响未接触谣言或相信谣言的微博用户, 从而影响谣言的传播效果。因此, 基于转发评论内容, 分析微博谣言接触者所持态度, 验证了加入消极传播者的SIIR模型更接近实际的谣言传播过程, 证实了该模型的基本假设。

表 1 新浪微博 10 条谣言信息表

微博 编号	发表时间	转发量	点赞数	评论量		质疑评论出现时间
				总评论	质疑评论	
01	2018/2/11 8:41	26	137	367	159	2018/2/12 19:45
02	2018/7/16 19:53	126	117	62	7	2018/7/16 21:04
03	2018/12/16 17:54	675	143	163	18	2018/12/17 7:10
04	2018/12/10 10:38	10	98	268	23	2018/12/11 8:46
05	2019/1/7 19:46	297	3626	278	12	2019/1/7 21:30
06	2019/1/4 0:42	431	15946	855	189	2019/1/8 0:56
07	2019/2/6 11:35	1133	302	256	169	2019/2/6 12:27
08	2019/8/5 7:43	320	9639	438	67	2019/8/6 21:30
09	2019/8/5 19:36	44	3503	426	35	2019/8/6 17:52
10	2019/9/26 17:20	26	494	108	50	2019/9/26 23:23

为了论证提出的SIIR谣言传播模型的有效性,通过Matlab软件对数学模型进行仿真模拟,将本文提出的改进SIIR模型与传统传染病模型SIR模型进行对比检验,图3为传统SIR模型的数值仿真模拟图,对比图 2和图3可以发

现不同节点的密度变化中主要是传播节点(传播者)的变化差别较大,参见图4。由于消极传播者的存在SIIR模型中谣言的传播速度与传播范围都不及SIR模型。在改进模型SIIR中,积极传播者的比例为50%,传播周期在60天左右

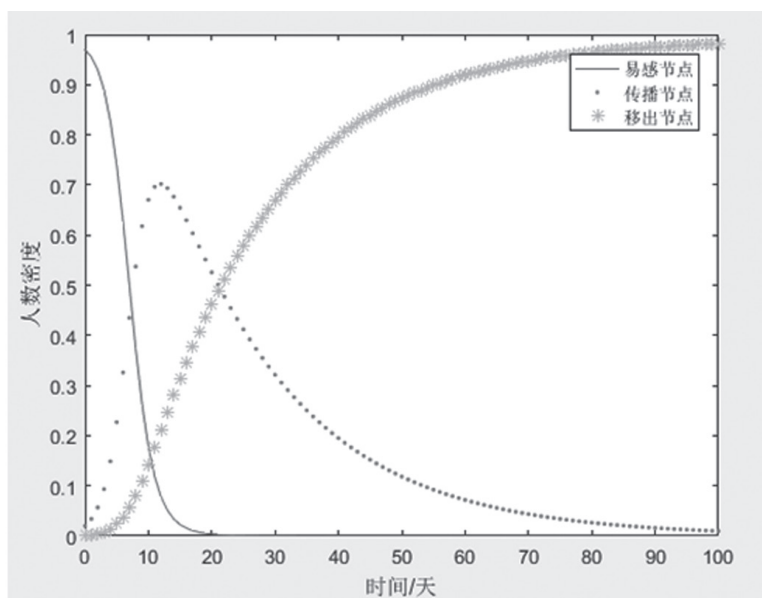


图 3 SIR 模型节点密度变化曲线

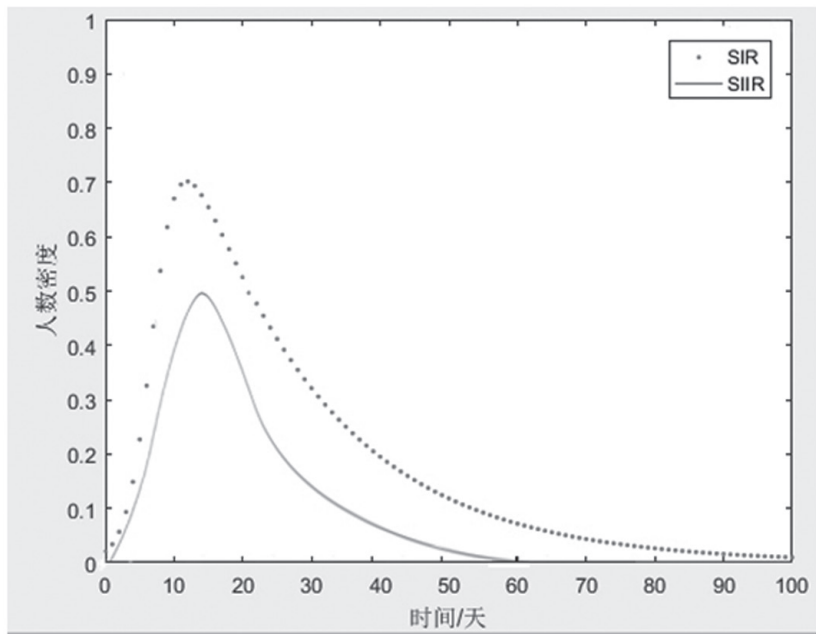


图4 SIIR模型与SIR模型传播谣言者对比

右,而传统SIR模型中,相信并传播谣言信息的传播者比例达70%,传播周期100天左右。这是由于在实际传播中,未知者在接触谣言信息时,受到了不实信息与真实信息的双重影响,并且当辟谣角色为官方人员或社会权威人士时,不实谣言的传播范围与生命周期将会受到更大影响,因此改进的SIIR模型较之传统模型传播,反映的谣言传播范围与生命周期更贴近真实传播过程。

3.5 结论

SIIR模型的构建基于对谣言传播者的分类,加入了传播辟谣信息与不传播信息的消极传播个体,更符合现实中的谣言传播情况,在实际的谣言传播中,由于新闻媒体、政府、权威机构等的作用,部分个体会传播辟谣信息,对于抑制谣言的传播有着一定作用。另一方面,从SIIR模型中可以得出影响谣言传播周期以及谣言扩散范围(积极传播者在社交网络中的占比峰值)的因素不仅包括移出率,还受到参数 η 的影响,因而在谣言传播生命周期内,

可以通过影响参数 α 和 η 控制谣言的扩散范围与时间。

4 结语

在大数据背景下,社交网络中的谣言信息的传播速度与扩散范围是十分惊人的,谣言信息的受众呈几何式增长,谣言传播的参与者类型与状态也愈加多样化,对谣言传播者进行分类能够更好地理解谣言传播机制,鉴于此,本文在充分分析现有的谣言传播模型的基础上,尝试构建了SIIR模型,并利用数值模拟的方法对其进行了分析。

首先,本文对研究对象所涉及的相关概念进行了解释与阐释,同时从传播角色、传播形式与谣言成因三个方面入手,分析了大数据背景下的社交网络谣言传播的新变化。其次,在经典谣言传播模型SIR模型的基础上,将处于感染状态的个体,即谣言传播者,分为积极传播者与消极传播者,构建了更加符合实际情况的谣言传播模型SIIR。最后,采用数值模拟

方法,使用MATLAB给出了谣言在社交网络中传播时,参与传播的个体的变化曲线,分析了影响谣言传播周期和范围的相关因素,并对SIIR模型进行检验论证,为控制谣言传播提供依据。

参考文献

- [1]张芳,司光亚,罗批.谣言传播模型研究综述[J].复杂系统与复杂性科学,2009,6(04): 1—11.
- [2]Zanette D H. Dynamics of Rumor Propagation on Small-world Networks[J]. Physical Review E, 2002, 65(4): 041908.
- [3]L. Buzna, K. Peters, D. Helbing. Modeling the Dynamics of Disaster Spreading in Networks. Physica A, 2006, 363: 132—140.
- [4]Moreno Y, Nekovee M, Pacheco A F. Dynamics of Rumor Spreading in Complex Networks[J]. Physical Review E, 2004, 69(6): 066130.
- [5]陶雪娇,胡晓峰,刘洋.大数据研究综述[J].系统仿真学报,2013,25(S1): 142—146.
- [6]张梦.大数据背景下国防科技信息传播新发展[J].情报理论与实践,2019,42(08): 50—53.
- [7]吴彦孜.微信谣言的扩散机制研究[D].西南大学,2016.
- [8]王国华,方付建,陈强.网络谣言传导:过程、动因与根源——以地震谣言为例[J].北京理工大学学报(社会科学版),2011,13(02): 112—116.