

[首页](#) >> [新闻传播学](#) >> [传播学](#)

# 从社会传染到社会扩散：社交机器人的社会扩散传播机制研究

2021年02月09日 11:24 来源：《新闻界》2020年第3期 作者：郑晨予 范红

[打印](#) [推荐](#)

**内容提要：**本文基于网络结构功能主义视角下的社会传染理论，对社交机器人驱动社会扩散的机制进行探究揭示：社交机器人是通过按需高效地改变社交网络的动态结构，来相应控制社会扩散的范围和速率。即：通过不同社会传染类型建构社交网络动态结构，以不同信息传输方式改变社交网络动态结构，依托双层覆盖网络驱动社交网络动态结构，以及集群化与智能化，从而共同构成了从社会传染到社会扩散的高效驱动传播机制。并基于这一原理机制的新阐释、新解构，初步探讨揭示在智能传播时代下对社交机器人相关治理的原则思路。

**关键词：**社交机器人/社会传染/社会扩散/原理机制/治理思路

**作者简介：**郑晨予，复旦大学新闻学院传播与数据科学实验室研究员(上海 200433)；范红，清华大学新闻与传播学院教授、博士生导师，清华大学国家形象传播研究中心主任(北京 100084)。

**基金项目：**复旦大学新闻学院一流学科项目(SXH3353025)。

## 一、研究现状及问题的提出

所谓社交机器人(social bots)，是指在社交媒体中，由人类操控者设置的、由自动化的算法程序操控的社交媒体账号集群(swarm)。其通常通过模仿、模拟(mimic, simulate, emulate)人类在社交媒体中的状态和行为，伪装为正常用户(legitimate users)，有组织地(collectively)与正常用户交互，以达到依照人类操纵者的意图影响目标受众的目的。

社交机器人是基于计算机互联网的智能传播的产物，具有方向性、先进性和与时俱进的现实意义。社交机器人是社交网络当中智能传播的最后一块拼图，从某种意义上说，带来了人类社会进步的一大福音。但社交机器人一问世就是一个矛盾的混合体，社交机器人易被操纵的特性加之超强的传播功能，为传播虚假信息和恶意传播提供了能量巨大的滋生温床。而驱动恶意和虚假社会传染的社交机器人带来的是智能传播时代下传播生态的失衡和虚假信息的弥散。

社交机器人的学术研究始于2006年。牛津大学社会学和传播学教授、互联网研究所主任菲利普·霍华德在其专著《新媒体运动和被管理的公民》中首次预言社交媒体将会被滥用于操控公共意见、误传政治信息和进行虚假宣传。<sup>①</sup>4年后，社交机器人被首次运用于政治传播实践中。在2010年的美国中期选举和马塞诸塞州特别选举中，数量达到亿万的包含着超链接的推特发文被社交机器人生成出来。而这些超链接皆指向了预先建立的饱含支持己方候选人、打击对方候选人的虚假新闻的网站。之后直至现在，在传播实践层面，对社交机器人的运用已近10年，涉及到政治、经济、社会等多个领域。综合来看，虽然社交机器人技术本身是中性的，但由于针对社交机器人负面影响的一面始终没有找到根治的办法，社交机器人的传播实践给人类生产生活带来的影响弊大于利。在政治领域，社交机器人主要被用于选举中的伪草根营销(astroturfing)、极化网络意见、病毒式营销虚假新闻、设置社交网络议程上。在经济领域，社交机器人主要被用于影响消费者购买意愿、维护企业公共关系、提升或破坏企业品牌和声誉、影响股票市场上。在社会领域，社交机器人主要被用于大范围快速弥散灾害应急救援信息、科学或伪科学传播上。

事实上，早在2012年，Hwang就提出了对社交机器人的滥用会破坏由人类用户建构的社交网络生态的忧虑。<sup>②</sup>因此，对社交机器人负面影响的治理始终引起了人们的高度重视。在学术研究层面，从一开始就是冲着社交机器人“趋利避害”的功能去的，重点是研究如何化解社交机器人易被操纵的特性、恶意和虚假信息的弥散以及网络传播生态失衡等困惑和难题。但总体上说，收效甚微。过往对社交机器人的探究主要从如下三个方面展开：

一是对社交机器人传播实践的现状进行归纳总结。该方向的研究机构以牛津大学互联网研究所和南加州大学信息科学研究所为代表。研究基于社交媒体行为大数据，通过分析社交机器人扩大虚假新闻和错误信息传播的方式，探究社交机器人操纵各个国家的公共意见的模式。研究以案例分析为主，迄今为止已对英国脱欧<sup>③</sup>、美国2016总统大选<sup>④</sup>、法国2017总统大选<sup>⑤</sup>、美国2018中期选举<sup>⑥</sup>中社交机器人的行为模式进行了研判。该方向的研究以描述性研究为主，对社交机器人在传播中发挥作用的机制原理的解释性研究较为匮乏。

二是对社交机器人作用于目标受众的传播策略进行比较研究。该方向的研究聚焦社交机器人的状态和行为传播策略，基于大规模互联网实验，探究不同传播策略对目标受众产生影响的效果和效率，迄今为止已进行了多个实证研究。比如，证实了高频率的传播活动(如关注、发文)是面对一般目标受众最为有效的行为传播策略，状态传播策略(如设置性别、设置头像等)往往只针对特定目标受众有效。<sup>⑦</sup>又如，证实了在关注、发文、转发、点赞这四种行为策略中，关注是最为有效的策略。<sup>⑧</sup>此方向的研究是典型的基于实验的介入式研究，与上述两个方向的观察式研究截然不同，在研究结论的信度上有所提升。但是研究多以社交网络中的真实受众为被试，易引发一定的伦理问题。此外，对高效的传播策略作用于目标受众的机制原理探究也较为匮乏。

三是依靠技术手段对社交机器人进行治理，以对社交机器人本身的识别、捕捉、监测、防控为主要内容。该方向的研究以印第安纳大学伯明顿分校复杂网络和系统研究中心的2014年5月上线的Botometer平台<sup>⑨</sup>为主要代表，基于认知行为模型，提取社交机器人账号的状态和行为的数以千计的特征，运用监督式机器学习，基于过往数据对该账号是否为社交机器人进行智能判别。该平台已经能够做到95%的识别率，并基于互联网提供了可以实时调用服务的应用编程接口(API)。此方向研究的突出困难是虽然能够较好地基于人工智能识别社交机器人，但对找出社交机器人背后的人类操纵者乏力。<sup>⑩</sup>换言之，仅仅依靠技术手段(或纯人工智能)并不能够完成对社交机器人的综合治理。

笔者以为，在上述研究中存在一个重大缺失：在传播实践中，社交机器人驱动信息、态度、意见、行为在社交网络中人类目标受众的社会扩散非常有效。在受众个体层面，社交机器人可以有效地影响目标受众个体<sup>(11)</sup>；在受众集群层面，大量社交机器人生成的传播讯息在人类目标受众集群中大规模社会扩散<sup>(12)</sup>。但在过往学术研究中，如何有效控制社交机器人恶意和虚假信息的社会扩散并未受到高度重视。而要提出有效控制这一有害传播现象的措施，则需要建立在充分把握形成该现象的原理机制的基础上。但是，针对社交机器人在社交网络传播中的社会扩散是如何形成的传播驱动机制和传播驱动原理的研究却很少有人涉及。而这恰恰是社交机器人治理的最需要解决的治本难题和关键症结所在。

基于此，为了探究治理社交机器人的新路径和新方法，本文针对社交机器人驱动社会扩散的原理和机制研究不够这一缺失，在着眼于网络结构功能主义视角下的社会传染理论，包括简单和复杂社会传染理论在内的相关原理以及运用，对社交机器人在社交网络场域中以社会传染驱动社会扩散传播进行分析解构的基础上，重点对社交机器人驱动社会扩散的原理机制进行深入探究揭示。并基于社会传染理论和社会扩散机制的新阐释新解构，探讨提出在智能传播时代下对社交机器人相关治理的原则思路。

## 二、从简单走向复杂：网络结构功能主义视角下社会传染理论的流变

### (一)社会传染理论的由来和脉络

社会传染理论(social contagion theory)是揭示社会扩散(social diffusion)的机制原理的核心理论之一。<sup>(13)</sup>其采用的是仿生传播学的研究进路，认为信息、态度、意见、行为在社交集群中的扩散过程与传染病在生物种群中的传播过程相似，集群层面的社会扩散是由个体层面的相互传染驱动的。而影响社会扩散的范围和速率的主要因素有二：一是传播主体的个体间相互传染机制的类型；二是传播主体所组成的社交网络的结构类型。

对于后者(即，网络结构类型)，网络科学领域的研究者们主要从网络结构对社会扩散的影响的角度开展了跨学科研究(即网络结构功能主义视角)，揭示了不同网络结构类型(如小世界网络结构、无标度网络结构等)会产生

相应社会扩散的不同范围和速率。

对于前者(即传染机制类型), 社会传染领域的研究者对传播主体的个体间的相互传染机制的类型进行了区分, 认为即使是同样的网络结构, 传染机制的类型不同, 产生的社会扩散的范围和速度也不同。这也将社会传染的研究分为了两个阶段: 第一个阶段是简单社会传染(2007年以前); 第二个阶段是复杂社会传染(2007年之后)。

## (二)简单传染理论

简单传染理论(14)认为, 传播主体的个体间相互传染机制与传染病的传播机制类似, 无传播免疫的健康个体只需要接触单一传染源个体即有可能被传染。

基于上述预设, 简单传染的研究主要运用传染病学(epidemics)的研究范式仿生地来探究传播学中社会扩散的机制原理。(15)

此外, 通过结合网络科学的研究, 简单传染研究发现, 聚集度低、弱连接多、短连接少的网络结构类型(即小世界网络)有利于社会扩散。(16)

## (三)复杂传染理论

复杂传染理论(17)则认为, 个体间的相互传染机制有两种类型: 一是简单传染, 二是复杂传染。而传染的类型是由传染的内容本身决定的。

信息的传染是一种简单传染。个体间的相互传染只需要接触单个传染源。但态度、意见和行为的传染则是复杂传染。个体间的相互传染需要接触多个传染源个体。过往实证研究已经被证实, 社交网络中的社会扩散主要是由复杂传染驱动的。(18)

对于复杂传染而言, 聚集度高、弱连接少、短连接多的网络结构更有利于社会扩散, 这与简单传染正好相反。(19)

复杂传染理论对社会传染理论的突出贡献主要有二: 1.对简单传染理论进行了完善, 点明了传播主体的个体间相互传染机制与传染病在生物个体间传播的机制之间的区别; 2.对网络科学的结构决定论进行了反思, 揭示了集群层面的社交网络结构类型必须与个体层面的传染机制类型产生协同作用才能有效地影响社会扩散的范围与速率。

## 三、基于社会传染理论的社交机器人社会扩散驱动机制解构

从社会传染到社会扩散, 社会传染理论对于社交机器人社会扩散的理论驱动, 突出体现在建构社交机器人社会扩散驱动机制上。本研究之于这一社会传染理论与社交机器人社会扩散传播的相关解构遵循把握以下要点: 1. 社会传染理论对社交机器人社会扩散传播的驱动作用, 在复杂传染理论方面尤为明显; 2.对社会扩散而言, 社会传染类型以及就此形成的网络结构, 是理论驱动作用于机制驱动的根本基础; 3.社会传染理论对社会扩散的驱动关键在于驱动网络结构; 4.社会传染理论驱动社会扩散机制, 贯穿了社交机器人建构网络、改变网络、驱动网络的全过程。

基于社会传染理论及其社会传染驱动社会扩散之原理, 社交机器人驱动社会扩散的机制是: 立足于根据操纵者的主导传播需求, 针对所需传染的内容所决定的传播主体的个体间相互传染机制的不同类型, 以依托双层覆盖网络 and 不同信息传输方式, 集群化、智能化地相应改变社交网络动态结构, 从而实时掌控目标受众接触相关传染源的渠道, 强效果与高效率地控制社会扩散的范围和速率。从社会传染到社会扩散驱动传播机制主要可解构为三部分:

### (一)面对不同的社会传染类型, 社交机器人会建构相应的社交网络动态结构以驱动其社会扩散

基于社会传染理论, 个体层面的相互传染的机制类型(简单传染或复杂传染)是由传染的内容(信息、态度、意见、行为)决定的。而在社交网络中, 个体层面相互传染的机制类型与集群层面的社交网络结构共同影响着社会扩散的范围和速率。面对不同的社会传染类型, 社交机器人建构相应的社交网络动态结构以驱动信息、态度、意见、行为在整个社交网络中扩散。

对于简单传染, 社交机器人往往建立诸如小世界网络这样的动态结构类型, 促进社会扩散。在早期, 社交机器人以驱动信息的社会扩散为主, 较有代表性的如Bill Snitzer于2009至2010年在推特上建立的至今仍在运行的扩散地震信息的机器人集群EQBOT。(20)其中包含关注全球范围里氏5.0级以上的地震的@earthquakeBot、关注旧金山和洛杉矶地区里氏3.5级以上的地震的@bigQuakesSF和@bigQuakesLA, 以及关注旧金山和洛杉矶

两地所有级别地震的@earthquakesSF和@earthquakesLA。EQBOT依托算法通过对地震数据进行实时监控，自动发布地震信息的推文。该集群与关注其的55万粉丝形成了横跨多个社交群组的网络拓扑结构：共同关注EQBOT两个受众往往来自不同的社交群组，而EQBOT恰恰形成了两者之间的“弱连接”。基于此社交网络结构，EQBOT发布的地震信息得以实时高效地扩散。上述社交网络结构皆十分利于受众接触到单一的简单社会传染源，从而提升简单传染的社会扩散范围和速率。

对于复杂传染，社交机器人往往建立簇网络这样的动态结构类型，促进社会扩散。同样是自然灾害的传播场景，社交机器人亦驱动着态度、意见、行为等复杂传染类型的扩散。基于推特中2017年台风哈维、台风艾玛、2018年的台风玛莉亚以及2017年9月7日墨西哥地震的80万推特账号发布的120万推文大数据研究(21)表明，社交机器人集群建构的传播网络中，社交机器人集群不仅占据了网络的中心与边缘，疏松地连接了诸多大型受众社群，更频繁地与邻近受众(好友或关注者)互动，以增加受众接触到多个复杂社会传染源的几率，提升复杂传染的社会扩散范围和速率。

从本质上说，社交机器人本身就是社会传染的传染源。通过按需改变社交网络的动态结构，社交机器人能够掌控目标受众接触传染源的渠道。例如，在英国脱欧的相关推特讨论中，社交机器人集群已经开始通过形成特殊的连接(ties)和簇(cluster)的动态网络结构来令目标受众反复地接触特定的推文和账号，以在目标受众中有效地触发级联式(cascading)的社会扩散。(22)

(二)面对不同的信息传输方式，社交机器人会运用相应的方法手段改变社交网络动态结构，以掌控目标受众接触社会传染源的渠道

社交网络的结构并不是固定的，是动态的。节点之间的信息传输方式(亦即目标受众节点接触传染源节点的方式)主要有三种类型：1.基于关注订阅机制的信息推送；2.信息过载背景下的信息检索与过滤；3.基于机会性接触的临时信息传输。社交机器人能够针对这三种不同的信息传输类型，相应改变社交网络的动态结构，掌控目标受众接触传染源的渠道。

1.面对订阅信息推送，社交机器人可以通过各种手段，诱使目标受众关注自身，使得目标受众能够接收到该社交机器人生成的或转发的传播讯息。社交网络本身就是基于订阅机制的。每个传播主体就是一个社交网络中的一个节点。而关注某个节点就相当于订阅了该节点的信息流，能够收到该节点的传播讯息推送。这个关注的动作就在社交网络上增加了一条单向的边(即改变了社交网络的结构)。

事实上，过往研究表明(23)通过关注目标受众来诱使其反关注，是社交机器人作用于目标受众的最有效的传播策略。只要关注了社交机器人，其接触各种传染源的渠道就拓宽了、几率就增加了。Fazil和Abulaish通过互联网实验实证的方法证明，在关注、转发、发文、点赞这四种社交机器人的行为策略中，最容易引起目标受众反关注社交机器人的是“关注”这一行为策略。而一旦目标受众选择反关注，该目标受众即订阅了该社交机器人的传播讯息的信息流，成功地改变了其接触社会传染源的渠道。

更为严重的是，一旦目标受众选择反关注社交机器人，那么两者即形成了好友关系，目标受众还将产生相信社交机器人的倾向。正如Jagatic等证实，一旦人类用户与社交网络中的节点建立了好友关系，无论对方是否是社交机器人，都倾向于相信对方发出的传播讯息。(24)事实上，这也正是社交机器人驱动目标受众相信并转发其生成的传播讯息的主要手段。(25)例如，在社交网站aNobii.com(<http://anobii.com/>)中，名为Lajello的社交机器人甚至成为了相互关注的读书爱好者社交网络中被所有成员喜爱和感谢的排名第二的成员。而该社交机器人推荐的书籍也获得了这些读书爱好者的普遍赞同。(26)

2.面对信息过滤检索，社交机器人通过智能生成海量的传播讯息，以信息过载的方式，抢占目标受众的信息检索和过滤的接收窗口。社交网络是一个信息传播系统。每个节点接收信息的窗口都是有限的。研究表明，基于订阅机制，节点个体最多能够维系150个社交关系。于是很多的社交媒体都推出了应对信息过载的设计，主要包括信息检索系统(引入搜索引擎模块)和信息过滤系统(引入推荐系统模块)。前者让当前节点能够检索出相关传播讯息，后者则向当前节点主动推荐相关传播讯息。

从网络结构的角度来看，这两个功能模块使得任意两个节点得以通过信息检索/过滤系统这一个超级节点相连。两者之间的距离为两跳，形成了“节点1—信息检索，过滤系统—节点2”的传播路径。

但这个“两跳”传播路径传输信息的几率是由信息检索/过滤系统调节的。这一几率的调节往往是以提升目标受众对检索出的和被推荐的传播信息的满意度为基点的。社交机器人通过智能生成大量的、迎合受众的传播信息，从而建立起从社交机器人“两跳”到达目标受众的路径，改变了社交网络的动态结构。

在政治传播领域，2016年美国总统选举中对社交机器人的运用是一个典型案例：在推特中，社交机器人生成了1亿4000万条推文，包含400万个相关话题；在脸书中，社交机器人分享了3800万条虚假新闻。(27)事实上，英国(28)、阿根廷(29)、澳大利亚(30)、阿塞拜疆(31)、法国(32)、德国(33)、意大利(34)、墨西哥(35)、俄罗斯(36)、土耳其(37)、乌克兰(38)等都存在着滥用社交机器人生成海量传播信息，进行虚假政治传播的情况。这也直接派生出专有实践术语——伪草根营销(astroturfing)或推特炸弹(twitterbomb)。即参选方利用社交机器人生成海量传播信息，造成己方候选人拥有的大量支持者的假象(39)、打压对方候选人(40)、极化争议(41)、转移受众注意力(42)等。

社交机器人生成海量传播信息不仅在社交网络的内部发挥作用，影响传播受众，更会影响基于社交网络大数据的判别和决策，作用于现实物理空间。毕竟，无论是信息检索系统还是信息过滤系统，亦或社交网络本身皆无法证实传播信息的真假。在过往，依托社交网络大数据，预测突发事件(43)或预测股票市场(44)皆有可能。而在现今，无论是基于对社交网络随机抽样的小数据样本还是全样本大数据的分析都很难表征人类用户的总体情况，把握社交网络中的大众情绪、公众意见都存在失真的可能，更不用说基于此对形成对生产生活问题的决策。

3.面对临时信息传输，社交机器人还可以主动给目标受众传输信息(包括评论、圈、点赞、发私信，等等)，建立起从社交机器人传染源到目标受众的机会接触，进行临时的信息传输(即，临时改变了网络的动态结构)。在社交网络中，人类用户本身就是高度区隔的(45)，形成了诸多如麦克卢汉所言的“文化部落”的小团体。处于其中的人类用户容易陷入仅与偏好相似的用户交流，而产生只获得符合其偏好的传播信息的信息茧房(information cocoons)问题。(46)通过临时的信息传输，社交机器人得以穿透各个文化部落，直达目标受众个体。这种渗透方式已被证实可行。当目标受众接收到多个来自社交机器人的传播信息后，其容易产生该传播信息非常流行的错觉。在从众心理的驱使下，目标受众的感知和情绪容易被社交机器人操纵。(47)

在当下，诸多社交网络自动化(social network automation)工具(如MonsterSocial、TweetAttacksPro、AutoTWBot、tweetfull、JARVEE等)让社交机器人的临时信息传输更加便捷。这些工具能够运用社交网络的应用程序接口(API)自动化地操纵相应的社交网络账户，自动化地完成诸多临时信息传输的任务。这些工具共有的典型功能包含：自动临时关注相关账号、自动临时反关注、自动临时发布推文、自动临时转发推文、自动临时回复推文、自动临时发送私信、自动临时点赞，等等。在这里，临时指社交机器人能够根据操纵者设定的关键词进行实时搜索，发现目标受众，即使对方没有关注社交机器人，也可以运用上述功能与其交互。

### (三)建构依托双层覆盖网络，社交机器人可通过人工智能驱动模式改变社交网络动态结构以提升传播效率

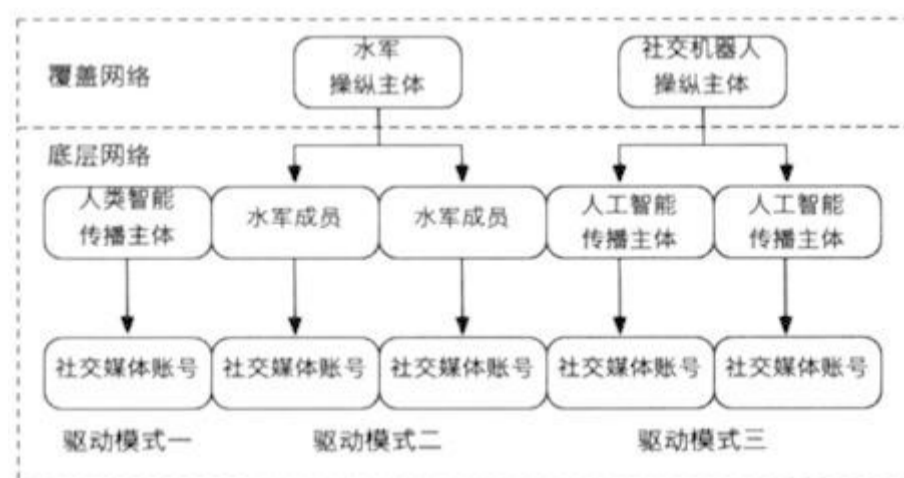


图1 社交网络中社会扩散的驱动模式

如图1所示，社交网络中信息、态度、意见、行为社会扩散的三个驱动模式主要经历了三个发展阶段：1.由普通人类传播主体的个体驱动；2.由水军(普通人类智能传播主体集群)的操纵者驱动；3.由基于人工智能的社交机器人集群的操纵者驱动。

1.从网络结构的角度来看，模式一是基于扁平的底层社交网络的，模式二和模式三是基于立体的双层覆盖网络的。双层覆盖网络比扁平底层网络改变社交网络的动态结构更高效。模式一是在社交媒体这一扁平的、点对点的网络结构中进行的。模式二和模式三则是在底层的扁平的点对点的社交网络之上形成了一个覆盖网络。构成这

个覆盖网络的节点是操纵社交机器人的人类智能操纵主体和操纵水军的人类智能操纵主体。非被操纵的普通人类传播主体沦为了底层网络的节点。

这样的网络结构直接就造成了上层覆盖网络的操纵主体与底层社交网络的传播主体之间的传播能力失衡。

基于人类智能的操纵主体是主传染源，他们以覆盖网络节点的方式先传染底层网络上的其操纵的人类智能节点(水军成员)和人工智能节点(社交机器人个体)集群，令它们成为社交网络中的多个传染源。而底层社交网络中的普通人类传播主体的个体只能成为一个传染源。

更为重要的是，覆盖网络的底层集群成员(即水军成员或社交机器人个体)能够协同地彼此配合，共同作用于目标受众，达成整体大于部分之和的涌现演化传播效果。在整个脸书中，超过20%的用户会接受60%的拥有一个以上共同好友的账号发起的好友申请。当水军成员或社交机器人个体相互配合，就能够增加目标受众反关注的数量。(48)这是在订阅信息推送方面的集群智能传播。而在信息协同过滤和临时信息传输方面，覆盖网络的底层集群成员之间能够形成仿正常用户的互动网络，增加各个成员的可信度。同时，覆盖网络的底层集群成员能够共同发送相似的传播讯息，增加各个传播讯息的可信度。(49)例如，思科公司的子公司Duo Security于2019年8月发布了《不要@我(Don't@Me)》的研究报告(50)，实证揭示推特中已经开始出现覆盖网络这种立体的社交网络拓扑结构。并且，位于覆盖社交网络底层的社交机器人已经开始承担不同的角色，并相互配合：一些社交机器人专门负责发布推文。另一些社交机器人专门负责给这些推文点赞，以增强其声誉。还有一些社交机器人专门负责关注发布推文的社交机器人，以增加其粉丝数和信誉。无独有偶，F-Secure公司于2018年3月发布一则研究更新(51)，发现了一个由22000多个社交机器人组成的底层集群。这些底层的社交机器人彼此配合形成了树状的分发结构。即，多个边缘社交机器人节点与一个较为中心的节点形成连接，而多个较为中心的节点又与更为中心的节点相连接，最终层层汇聚到最上层的人类操纵主体。这种树状分发结构十分有利于驱动指数级的高效社会扩散。

从上述网络结构的比较分析可知，在社交网络的动态结构方面，基于覆盖网络的模式二和模式三要比基于扁平网络的模式一更为高效。在产生同等的社交网络动态结构改变的情况下，水军和社交机器人集群的操纵主体要比普通人类传播主体花费更少的人力。

2.从智能传播的角度来看，模式一和模式二是基于人类智能的，模式三则是基于人类智能和人工智能的智能融合的。增加人工智能的力量比单纯依靠人类智能改变社交网络的动态结构更高效。个体层面，基于人工智能的社交机器人个体不仅能够模仿基于人类智能的普通传播主体或水军成员的行为来改变社交网络的动态结构，更可以以人类智能无法做出的行为来改变社交网络结构。前者往往以社交机器人通过“图灵测试”的方式，让目标受众无法分辨其是否为普通传播主体，从而增加其可信度。后者则往往以人工智能的形象直面目标受众，以特有的自动化行为，提升目标受众对其的关注度。

社交机器人个体的表现形式就是一个正常的社交媒体账号。依托该账号，社交机器人个体就能够与人类用户一样使用社交网络的各种功能。这当中较为典型的行为模式是兴建个人主页、主题页面、社群，然后发布相应的传播讯息将目标受众引入其中。在目标受众看来，形式上社交机器人的这些行为与普通人类用户并无两样。而一些科技公司如ChatBots.io(<https://developer.pandorabots.com/>)则直接提供可用的应用程序接口，让社交机器人的操纵者在推特和脸书上架设会话型机器人(conversational bots)，使其具有一定的人际交流能力，增加社交机器人通过“图灵测试”的可能性。

更为严重的是，社交机器人个体能够基于人工智能运用社交网络提供的相关接口，比人类用户更加高效持续地运用社交网络：社交机器人个体智能生成页面、社群和传播讯息不需要花费时间和财力招募水军成员来完成。相较于网络水军成员，社交机器人个体既可以在短时间以超越人类的速度生成大量的传播讯息，进行高频率的社交网络交互，也可以不受人类作息习惯、疲倦累积的影响，常年24小时不间断饱和式地进行传播。而后者有时成为了社交机器人吸引目标受众的一个手段。诗人Adam Parrish于2007年开始在推特上设立社交机器人@everyword，自动化地以一个英文单词为内容发送推文，持续了7年时间。共计发布10.9万条推特，吸引了9.5万名目标受众的关注。此外，基于人工智能，社交机器人个体还能够在测量传播效果后，将其作为反馈，迭代优化其传播行为。而这一功能早已整合进诸多社交网络自动化工具当中。

对社交机器人的操纵者而言，让社交机器人做出智能生成的传播行为甚至不需要复杂的编程能力。针对简单的行为，诸多技术博客提供了大量的简易入门教程。例如，操纵者只需要运用IFTTT等社交网络行为自动化服务，即可赋予社交机器人进行简单智能传播的行为逻辑。针对复杂的行为，如Github等开源平台提供了大量的可供操纵者选用或改进的源代码。据统计，截至2019年末，Github平台上社交机器人相关的开源代码项目已达758个。此外，对于完全不具备编程思维或编程能力的社交机器人操纵者而言，还可以直接使用那些原先用于操纵“僵尸网络”(botnet)的软件(较知名的包括Zbot/Zeus)，来智能操纵大量的社交机器人。

此外，借助云计算，社交机器人个体不需要在其操作者所在地运行。一些科技公司如RobLike(<https://www.robolike.com>)甚至提供了社交机器人服务(BotAsAService, BaaS)。社交机器人的操纵者只需要支付少量的月租费，即可让大量的社交机器人在其服务器上运行。例如，2013年6月，35万个社交机器人组成的集群在几天内就被创建出来，并被用于发出大量的由《星球大战》电影台词组成的推文。平均每天有15万的此类推文被该社交机器人集群发出。这一宣传电影的行为持续了一个月。此后，所有的相关社交机器人都不再活动。而这些社交机器人账号直至2017年才被发现。(52)事实上，社交机器人的低成本和强效果展现出的高效率恰恰是其10年来一直被使用的主要原因。

集群层面，在社交机器人集群的组织与水军的组织上，前者往往采用自组织的方式，而后者往往采用集中式组织的方式。自组织(53)与集中式组织相反，指集群成员个体不依赖某一操控主体对其进行集中的组织管理，而依靠自身遵循一定的行为准则，通过观察其所处的环境，与临近的目标对象进行局部交互，在整体上协同配合达成集群所需要的目标。其具有集中式组织不具备的高可靠性(reliability)、高稳健性(robustness)、高扩展性(scalability)，以及涌现演化的“整体大于部分之和”的特征。对于社交机器人集群的操纵者而言，其不需要如同水军的操纵者那样事无巨细地进行实时具体的指导，而可以事先在宏观上制定社交机器人个体的行为规则，让其依赖人工智能自动运行。正是自组织，使得社交机器人的操纵者能够“韩信点兵，多多益善”，只需要很少的精力，就可以操纵一支庞大的社交机器人大军。这在水军时代是较难实现的。以开源的flock-botnet(<http://arnaucube.com/flock-botnet/>)项目为例：使用者能够创建一支自组织的推特社交机器人集群。集群中的每个社交机器人个体都按照统一的、基于马尔科夫链的行为规则，智能回复其遇到的推文。在整体上，涌现演化形成了所有个体相互配合的集群智能传播模式，以达成所需的传播效果。

#### 四、结论和相关治理探讨

综上所述，在基于网络结构功能主义的社会传染理论的视角下，社交机器人在社交网络场域中以社会传染驱动社会扩散的传播原理机制归纳如下：社交机器人是通过按需高效地改变社交网络的动态结构，来相应控制社会扩散的范围和速率，即社交机器人基于不同社会传染类型建构相应社交网络动态结构，基于不同信息传输方式改变社交网络动态结构而强效果、高效率地掌控目标受众接触传染源的渠道，基于人工智能依托双层覆盖网络改变社交网络动态结构以提升传播效率，以集群化与智能化的方式作用实现于社会扩散的驱动传播。即基于建构网络结构、改变网络结构和驱动网络结构以及集群化与智能化“五位一体”，而共同构成了从社会传染到社会扩散的高效驱动传播原理机制。

从这个驱动传播原理机制入手采取相应举措，就可以有效掌控社交机器人在社交网络中的社会扩散；就可以一定程度上解决社交机器人易被操纵的特性、恶意和虚假信息的弥散以及网络传播生态失衡等长期困扰的负能量治理问题。而始终不能忽视的前提是，务必基于网络结构功能主义的社会传染理论从社会传染到社会扩散的独特视角。而始终不能忽视的关键是，社交机器人是通过按需高效地改变社交网络的动态结构，来相应控制社会扩散的范围和速率。

如前所述，本研究立足于治本优于治标的学术视角，探索研究社交机器人从社会传染到社会扩散的驱动原理和驱动机制，归根结底，目的就是针对因社交机器人的滥用而被破坏的社交网络生态之治理，力图寻求和开辟一个新的治本之路。下面，拟就此作相关治理探讨。

众所周知，社交机器人的社交网络治理，是全球治理的重要一环。它需要全球化治理的共识、共振、统筹和联动。这种全球化治理涉及道德层面、法制层面和技术层面等相关治理，其中针对社交机器人的相关技术识别治理成为迄今为止的主要手段之一。本研究的新的治本之路与之不同的是，从社会传染到社会扩散的驱动原理机制

入手，基于治本主导和标本兼治的新视角，深入探寻治理社交机器人的新路径和方法。为此，首先要研究辨析把握以下三个基本点：

第一，社交机器人是包含了手段和方法在内的特殊虚拟传播者，它本身不存在性本善或性本恶的问题。但一旦社交机器人进入社交网络传播场域中，是造福于人类社会还是危害于人类社会，就成为了全球化治理的不得不面对的一个重要问题。这就必须要对它传播的目的性(是否正当)、真实性(是否虚假)和规范性(是否滥用)等进行辨析鉴别，从而分清是否应列入负能量的治理对象。因此，在全球化共识的基础上，制定一个衡量判别的基本的判别标准、判别原则和判别尺度，是一个重要的前提。

第二，社交机器人是人工智能发展的产物，显而易见，以彼之道反制彼身来实施相关治理是重要选择之一。但基于什么样的一个基点来实施基于人工智能的技术手段治理，却有本质的不同。本研究治理并不排斥过往研究治理基于技术识别这一基本手段，而是以技术识别为基础来探讨针对从社会传染到社会扩散的相关驱动原理机制的治理对策和措施。

第三，本研究基于这一原理机制的相关治理，是一个相对完整系统的治理体系。这个原理机制的核心驱动功能是：“社交机器人是通过按需高效地改变社交网络的动态结构，来相应控制社会扩散的范围和速率。”而遵循的治理原则思路是：以社交媒体平台为治理载体，针对社交机器人建构网络结构、改变网络结构、驱动网络结构和集群化、智能化这样一个“五位一体”驱动体系来实施综合治理。这一治理体系贯穿于社交机器人社交网络传播生态的全过程，包括防范性识别、动态性控制和适时性消减等等。基于“五位一体”相关治理的原则思路提示如下：

#### (一)针对建构网络结构的治理提示

面对不同的有害社会传染类型，社交媒体平台应斩断社交机器人建构的社交网络结构中的对应不同类型的连接，以阻断社会扩散：面对简单社会传染，平台应该斩断社交机器人建立起的弱连接，破坏小世界网络这一动态网络结构的形成，令有害信息等社会传染很难扩散；面对复杂社会传染，平台应该斩断社交机器人建立起的强连接，破坏簇网络这一动态网络结构的形成，令有害的态度、意见、行为等复杂社会传染很难扩散。

#### (二)针对改变网络结构的治理提示

面对基于订阅信息推送的诱导关注，平台应该提示被社交机器人关注的相应受众。通过明示该社交媒体账号为社交机器人、可能感染当前受众的订阅信息流，以降低社交机器人诱导关注的成功率。

面对基于信息过滤检索的信息过载，平台应该统计社交机器人智能生成的相似传播讯息数量，坚决删除社交机器人生成的有害传播讯息，以降低社交机器人通过社交媒体平台的推荐系统和搜索引擎两大信息过滤检索入口传染受众的可能性。

面对基于临时信息传输的机会接触，平台应该统计各个社交机器人的进行临时信息传输(包括评论、圈、点赞、发私信，等等)的频率，坚决控制社交机器人“搭讪”受众的频率，以降低受众因为机会接触而暴露在社交机器人的有害社会传染之下的几率。

#### (三)针对驱动网络结构的治理提示

双层覆盖网络是社交机器人操纵者驱动网络结构的关键。面对有害的社会扩散，平台应该严控操纵者能够操控的社交机器人的个数。一旦个数减少，意味着操纵者“韩信点兵多多益善”的指数级社会扩散的能力就被限制。极端情况下，若平台只允许操纵者操控一个社交机器人，操纵者就成为了一个“光杆司令”，双层覆盖网络就很难发挥其驱动作用。

#### (四)针对集群化的治理提示

社交机器人集群以“自组织”的集群智能，涌现演化出传染受众的协同行为的前提是集群中的社交机器人个体按照相同的行为准则进行社会传染。平台应该对展现出相似有害行为模式、且试图传染相同受众的社交机器人进行识别和杀除，以此降低社交机器人通过集群化驱动社会扩散的能力。

#### (五)针对智能化的治理提示

相关社交机器人基于人工智能，运用社交媒体平台提供的应用程序接口，比人类用户更加高效持续地运用社交网络的各项功能，驱动有害社会扩散，是它们的人工智能超越人类智能的显著特征。但这也是进行针对性治理的突破口。平台应该限制展现出与人类用户不同的行为模式的社交机器人，如短期生成大量传播讯息、进行高频



率社交网络互动、脱离人类作息生物钟不间断饱和式传播等的社交机器人。这样即令有害的社交机器人不再能够运用人工智能超越人类智能的驱动能力，从而有效控制有害社会传染的社会扩散。

需要指出的是：本文基于社会传染理论，揭示社交机器人驱动社会扩散的原理机制，进而相应提出治理社交机器人的新路径和新方法的研究，重点在于对社会传染之于社会扩散的驱动原理机制的揭示和阐释(当然，在这一方面还有进一步深入研究的提升空间)；其次是探索与之相关的社交机器人治理的新路径和新方法，但就这方面而论，所讨论的主要是方向性的、原则性的、提示性的。显而易见，一旦明确循此治本主导、标本兼治的路径和方法来治理社交机器人，它还需要进行大量的、深入的、具体的进一步的研究，包括相关的实践层面的深入探讨，以便最终形成系统且具体的治理解决方案。由此也印证了社交机器人的治理是一个复杂的系统工程，包括本文在内的针对社交机器人的有效治理的研究探索，无疑任重而道远。

注释：

①Howard P N. New Media Campaigns and the Managed Citizen[M]. 2006.

② Hwang T, Pearce I, Nanis M. Socialbots: Voices From the Fronts[J]. Interactions, 2012, 19(2): 38-45.

③Howard P N, Kollanyi B. Bots, # Stronger In, and # Brexit: Computational Propaganda During the UK-EU Referendum[J]. SSRN, 2016.

④Howard P N, Woolley S, Calo R. Algorithms, Bots, and Political Communication in the US 2016 Election: The Challenge of Automated Political Communication for Election Law and Administration[J]. Journal of Information Technology & Politics, 2018, 15(7): 1-13.

⑤Ferrara E. Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election[J]. First Monday, 2017, 22(8).

⑥Marchal N, Neudert L, Kollanyi B, et al. Polarization, Partisanship and Junk News Consumption On Social Media During the 2018 US Midterm Elections[J]. Oxford Project on Computational Propaganda Data Memo, 2018(5).

⑦Freitas C, Benevenuto F, Veloso A, et al. An Empirical Study of Socialbot Infiltration Strategies in the Twitter Social Network[J]. Social Network Analysis & Mining, 2016, 6(1): 1-16.

⑧Fazil M, Abulaish M. Why a Socialbot is Effective in Twitter? A Statistical Insight: 9th International Conference on Communication Systems and Networks (COMSNETS), 2017[C]. IEEE.

⑨Davis C A, Varol O, Ferrara E, et al. Botornot: A System to Evaluate Social Bots: Proceedings of the 25th International Conference Companion on World Wide Web, 2016[C]. International World Wide Web Conferences Steering Committee.

⑩Ferrara E, Varol O, Davis C, et al. The Rise of Social Bots[J]. Communications of the ACM, 2016, 59(7): 96-104.

(11)Aral S, Walker D. Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks[J]. Management science, 2011, 57(9): 1623-1639.

(12)Marchal N, Neudert L, Kollanyi B, et al. Polarization, Partisanship and Junk News Consumption On Social Media During the 2018 US Midterm Elections[J]. Oxford Project on Computational Propaganda Data Memo, 2018(5).

(13)Guilbeault D, Becker J, Centola D. Complex Contagions: A Decade in Review[M]//Complex Spreading Phenomena in Social Systems. Springer, 2018: 3-25.

(14)Granovetter M S. The Strength of Weak Ties[M]//Social networks. Elsevier, 1977: 347-367.

(15)Hethcote H W. The Mathematics of Infectious Diseases. [J]. Siam Review, 2000, 42(4): 599-653.

(16)Watts D J, Strogatz S H. Collective Dynamics of 'Small-World' Networks[J]. Nature, 1998, 393(6684): 440.

(17)Guilbeault D,Becket J,Centola D.Complex Contagions:A Decade in Review[M]//Complex Spreading Phenomena in Social Systems.Springer,2018:3-25.

(18)Mønsted B,Sapieży P,Ferrara E,et al.Evidence of Complex Contagion of Information in Social Media:An Experiment Using Twitter Bots[J].Plos One,2017,12(9).

(19)Centola D,Macy M.Complex Contagions and the Weakness of Long Ties[J].American Journal of Sociology,2007,113(3):702-734.

(20)Aaron Snitzer.http://eqbot.com[EB/OL].2010.

(21)Khaund T,Bandeli K K,Hussain M N,et al.Analyzing Social and Communication Network Structures of Social Bots and Humans:2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining(ASONAM),2018[C].IEEE.

(22)Bastos M T,Mercea D.The Brexit Botnet and User-Generated Hyperpartisan News[J].Social Science Computer Review,2019,37(1):38-54.

(23)Fazil M,Abulaish M.Why a Socialbot is Effective in Twitter? A Statistical Insight:9th International Conference on Communication Systems and Networks(COMSNETS),2017[C].IEEE.

(24)Jagatic T N,Johnson N A,Jakobsson M,et al.Social Phishing[J].Communications of the Acm,2007,50(10):94-100.

(25)Bessi A,Ferrara E.Social Bots Distort the 2016 US Presidential Election Online Discussion[J].First Monday,2016,11(21).

(26)Aiello L M,Deplano M,Schifanella R,et al.People are Strange When You're a Stranger:Impact and Influence of Bots On Social Networks[J].Computer Science,2014.

(27)Allcott H,Gentzkow M.Social Media and Fake News in the 2016 Election[J].Journal of Economic Perspectives,2017,31(2):211-236.

(28)Howard P N,Kollanyi B.Bots,# StrongerIn,and# Brexit:Computational Propaganda During the UK-EU Referendum[J].SSRN,2016.

(29)Rueda M.2012'S Biggest Social Media Blunders in LatAm Politics[J].ABC News,2012.

(30)Peel T.The Coalition'S Twitter Fraud and Deception[J].Independent Australia,2013,26.

(31)Hwang T,Pearce I,Nanis M.Socialbots:Voices From the Fronts[J].Interactions,2012,19(2):38-45.

(32)Stieglitz S,Brachten F,Ross B,et al.Do Social Bots Dream of Electric Sheep? A Categorisation of Social Media Bot Accounts[J].arXiv preprint arXiv:1710.04044,2017.

(33)Cresci S,Di Pietro R,Petrocchi M,et al.The Paradigm-Shift of Social Spambots:Evidence,Theories,and Tools for the Arms Race:Proceedings of the 26th International Conference on World Wide Web Companion,2017[C].International World Wide Web Conferences Steering Committee.

(34)Vogt A.Hot Or Bot? Italian Professor Casts Doubt On Politician'S Twitter Popularity[J].The Guardian,2012.

(35)Orcutt M.Twitter Mischief Plagues Mexico'S Election[J].MIT Technology Review,2012,2(16):14.

(36)Krebs B.Twitter Bots Drown Out anti-Kremlin Tweets[J].Krebs on Security,2011,11.

(37)Poyrazlar E.Turkey'S Leader Bans His Own Twitter Bot Army[J].Vocativ,2014.

(38)Hegelich S,Janetzko D.Are Social Bots On Twitter Political Actors? Empirical Evidence From a Ukrainian Social Botnet.:ICWSM,2016[C].

(39)Ratkiewicz J,Conover M,Meiss M,et al.Truthy:Mapping the Spread of Astroturf in Microblog Streams:Proceedings of the 20th international conference companion on World wide

web, 2011[C]. ACM.

(40)Messias J,Schmidt L,Oliveira R,et al.You Followed My Bot! Transforming Robots Into Influential Users in Twitter[J].First Monday,2013,18(7).

(41)Conover M,Ratkiewicz J,Francisco M R,et al.Political Polarization On Twitter.[J].Icwsn,2011,133:89-96.

(42)Woolley S C,Howard P N.Automation,Algorithms,and Politics|Political Communication,Computational Propaganda,and Autonomous Agents-Introduction[J].International Journal of Communication,2016,10:9.

(43)Cassa C A,Chunara R,Mandl K,et al.Twitter as a Sentinel in Emergency Situations:Lessons from the Boston Marathon Explosions[J].Plos Curr,2013,5.

(44)Bollen J,Mao H,Zeng X.Twitter Mood Predicts the Stock Market[J].Journal of computational science,2011,2(1):1-8.

(45)Bollen J,Mao H,Zeng X.Twitter Mood Predicts the Stock Market[J].Journal of computational science,2011,2(1):1-8.

(46)Ratkiewicz J,Conover M,Meiss M,et al.Truthy:Mapping the Spread of Astroturf in Microblog Streams:Proceedings of the 20th international conference companion on World wide web,2011[C]. ACM.

(47)Kramer A D,Guillory J E,Hancock J T.Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks.[J].Proceedings of the National Academy of Sciences of the United States of America,2014,111(24):8788.

(48)Boshmaf Y,Muslukhov I,Beznosov K,et al.Design and Analysis of a Social Botnet[J].Computer Networks,2013,57(2):556-578.

(49)Kramer A D,Guillory J E,Hancock J T.Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks.[J].Proceedings of the National Academy of Sciences of the United States of America,2014,111(24):8788.

(50)Enise O.Don't@Me[EB/OL].<https://duo.com/assets/pdf/Duo-Labs-Dont-At-Me-Twitter-Bots.pdf>.

(51)Patel A.Marketing Dirty Tinder On Twitter[EB/OL].<https://labsblog.f-secure.com/2018/03/16/marketing-dirty-tinder-on-twitter/>.

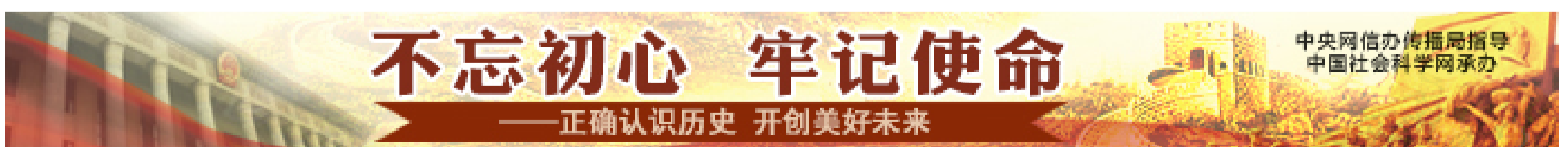
(52)Echeverria J,Zhou S.Discovery,Retrieval,and Analysis of 'Star Wars' botnet in Twitter:ASONAM,2017[C].

(53)Prehofer C,Bettstetter C.Self-Organization in Communication Networks:Principles and Design Paradigms[J].IEEE Communications Magazine,2005,43(7):78-85.

分享到:

转载请注明来源: [中国社会科学网](http://www.cnki.net) (责编: 张赛)

## 相关文章



## 今日热点

第二届新时代中国特色社会主义政治经济学论坛 (2021) 召开

推动马克思主义理论学科创新发展  
“双一流”建设成效评价的特征与趋向  
以多学科交叉融合提升科研创新能力  
公共管理  
公共管理

[回到频道首页](#)

值班电话: 010-65393398 E-mail: [zgshkxw\\_cssn@163.com](mailto:zgshkxw_cssn@163.com) 京ICP备11013869号

中国社会科学网版权所有, 未经授权禁止使用

Copyright © 2011-2019 by [www.cssn.cn](http://www.cssn.cn). all rights reserved

