

基于遗传算法和贪心算法的网络舆情传播 优化模型构建研究

杨文阳

(西安石油大学 陕西 西安 710065)

摘 要: 当前众多的社交网络可以在用户之间共享和传播信息,为了有效识别社交网络舆情信息源的特征,网络舆情信息源在启动和支持舆情传播过程中发挥至关重要的作用。但调整网络舆情信息源具有一定的成本,因此,如果舆情信息传播不能优化配置会导致在调整过程中浪费大量的资源,甚至会导致网络舆情信息源性能的下降。本文以微信为例讨论了网络舆情信息源设置和优化问题,并提出了解决该问题的模型。该模型结合了信息源模型和信息传播模型,该模型把信息源的个体特征作为其定位特征。该模型可以处理微信舆情传播优化所需的所有步骤:网络监控和数据收集、参数和识别以及信息源布局优化等。基于便利抽样对微信平台收集的数据集进行算法分析,验证了该模型采用遗传算法生成的解决方案可有效应用于网络舆情信息源的布局优化。

关键词: 遗传算法; 贪心算法; 网络舆情传播; 优化模型

中图分类号: TP391 **文献标志码:** A **文章编号:** 1673-5692(2020)11-0057-08

Research on the Construction of Optimization Model of Network Public Opinion Propagation Based on Genetic Algorithm and Greedy Algorithm

YANG Wen-yang

(Xi'an Shiyou University, Xi'an 710065, China)

Abstract: At present, many social networks can share and disseminate information among users. In order to effectively identify the characteristics of social network public opinion information sources in the network, information sources of network public opinion play a vital role in the process of initiating and supporting public opinion dissemination. However, they have a certain cost when we information sources of network public opinion are adjusted, if their dissemination cannot be optimally configured, a large amount of resources may be wasted in the adjustment process, and even the performance of information sources in network public opinion may be reduced. In this article, WeChat is taken as an example to discuss the problem about information sources of network public opinion in setting and optimization, and propose a model to solve it. This model combines the information source model and the information dissemination model, and the model takes the individual characteristics of the information source as its feature. This model can handle all the steps required for the optimization of WeChat public opinion dissemination: network monitoring and data collection, parameter and setting identification, and information source layout optimization. Through algorithmic analysis of the data set collected by the WeChat platform based on convenience sampling, it is verified that the solution generated by the model using genetic algorithm can

收稿日期: 2020-09-02 修订日期: 2020-11-02

基金项目: 陕西省社会科学基金资助项目(2019N017)。

be effectively applied to the optimization of the layout of information sources of network public opinion.

Key words: genetic algorithm; social network; wechat; network public opinion dissemination; optimization model

0 引言

社交网络已经成为我们日常生活不可或缺的重要组成部分,它方便人们基于相关主题快捷交流、共享和传播信息、知识以及个人观点。社交网络舆情传播的爆发性和扩散性强,社会影响力大^[1]。网络舆情信息通常是通过公共传播机制进行传播的:转发、关注和交流,其传播过程具有不同的动力机制,这取决于其初始条件、传播阶段、用户对舆情信息的反应程度、参与者以及网络舆情信息源如何传播信息。网络舆情的主体包括公职部门、网络强势媒体、意见领袖和普通网民等,不同的利益相关者在社交网络的帮助下快速推送自己的意见、情绪、利益诉求和兴趣。在不同条件和阶段下,网络舆情信息传播结果会发生明显的变化。当网络舆情大规模爆发时会形成舆论,会对社会产生极大的影响。

信息源特征直接影响网络舆情传播效果,信息源特征主要有信息源的可信度、信息源的专业性和信息源的吸引力^[2]。在信息内容相同时,信息源的可信度越高,信息越容易被网民认同,反之,被网民接受的概率就会降低。但信息源的可信度来源于信息源的专业性和吸引力,网络用户的身份认证和粉丝数是信息源特征的直接影响因素^[3]。舆情信息的传播主要依赖于信息的评论和转发来实现^[4]。

因此,信息的评论数和转发数是网络舆情信息源特性和传播效果的最重要评价指标。为了实现网络舆情信息传播的预期效果,比如在有限的时间范围内用户可以一次或多次重新发布信息,应该配置网络舆情信息源在网络中的设置和定位。舆情信息评论者数量和转发数量决定了特定信息源对信息传播的可用性及成本,成本是将信息源运作过程中必须花费的支出。舆情传播过程中,易接受网民占比会随着时间演化而不断减小^[5]。不同的网络用户在信息传播中可以达到相同的传播效果,这些用户为了达到期望的传播效果既可以选取成本较高的信息源,也可以选取成本较低的信息源^[6-7]。

为了启动和执行信息传播,必须将不同特征的信息源引入网络中。但是要改善信息源性能和降低运作成本,必须对网络舆情传播过程进行优化布局。

为了解决这个问题,需要对具体的解决方案进行评估,常见的用于描述信息传播过程模型有:线性阈值模型、流行病学模型以及 Daley-Kendal 模型等。但这些模型没有考虑网络舆情信息源的个体特征,这对于网络舆情能否成功传播至关重要。解决网络舆情传播优化任务还需要具备以下要素:用户身份特征、信息转发数、信息关注数以及可能对预期传播产生正面或负面影响的信息传播过程等。声誉高的用户在社交网络中的影响力较大,其网络行为在舆情网络信息传播过程中起着重要的作用^[8]。因此,需要对舆情信息传播过程涉及的用户身份、信息的评论数和转发数进行识别,并基于这些指标获得用于优化模型的参数。

总之,在信息传播过程中用户数据、信息的评论数和转发数等关键指标都需要一种工具来解决识别和优化任务。本文提出的框架模型旨在为网络舆情传播过程的优化提供完整的解决方案,其作用主要体现在以下三个方面:1) 该模型主要包括三个模块:监测和数据收集模块、识别模块以及信息源布局优化模块;2) 修改遗传算法以解决网络舆情传播过程的优化问题,同时能调整信息源的特征,基于可扩展的可疑感染模型对信息源传播过程的个别特征进行模拟;3) 对现有的解决方案进行实验研究验证其是否合理。

1 研究现状及问题的提出

1.1 研究现状概述

目前,社交网络舆情传播研究正呈逐年增长趋势,从研究热点看,国内外研究主要聚焦社交媒体环境下的网络舆情用户行为、不同社交媒体的舆情传播特征等^[9]。有很多研究主要聚焦于分析网络舆情传播过程或网络舆情在社交网络中的优化方法,通过梳理发现该研究主要集中于三个方面:优化传播方法、信息源布局的性能评估以及识别与爬网的优化解决方案等。

网络舆情传播有以下三种优化方法:1) 利用启发式方法进行优化,如利用爬网算法优化舆情传播解决方案,之前的研究只关注信息源的位置并没有

关注其特征^[10-11]; 2) 利用贪婪算法来挖掘最具影响力的网络用户, 其目标不是优化舆情传播解决方案, 因此不会构建不同的信息源来表征其特征^[12]; 3) 如果网络舆情在传播过程中存在多个传播通道和节点阻力, 应利用算法优化网络舆情信息源的传播路径^[13], 但该研究没有考虑网络舆情信息源的个体特征。网络舆情信息源的布局性能可以采用以下四种方法进行评价: 1) 基于拓扑的评价模型, 例如按照程度和中间度进行评价^[14], 这种评价方法通常比其他方法更快, 但在近似值评估方面相对较差; 2) 基于仿真模拟的模型, 比如流行病学的分析模型^[15], 但这些模型不允许根据网络舆情信息源的单个特性来评价信息源传播的解决方案; 3) 基于多元代理人的模型, 该模型可模拟单个消息重新发布信息^[15]; 4) 基于网络传播和区块链理论提出社交网络舆情传播模型, 该模型考虑该类社交网络中由合理量化价值贡献而产生的激励机制对于用户传播信息产生的影响^[16]。但这些研究没有将这些模型用于优化网络舆情信息源布局, 也没有充分挖掘利用网络舆情信息源的个别特征, 比如生成消息的频率和质量。

1.2 研究问题的提出

综上所述, 目前大多数相关研究关注单个网络

舆情信息特征的识别、分布式爬行算法、克服网速率限制、网络事件监测等, 在不考虑成本的情况下对信息源特征进行识别, 没有基于网络舆情特征对网络舆情传播过程中的用户数据、信息的评论数和转发数等关键指标进行评价。基于此, 本文提出以下两个问题:

1) 在传播范围广、信息传播数据较大的社交网络中, 如何低成本对网络舆情信息源的个体特征集进行有效识别?

2) 基于网络舆情的个体特征如何为网络舆情传播过程中涉及的用户数据、信息的评论数和转发数等关键指标提供一个有效的布局方案? 如何有效评价网络舆情信息传播的效果?

2 网络舆情传播优化模型的构建

网络舆情信息源优化模型主要由搜寻器模块、识别模块和优化模块三个部分组成, 它可以应用于在线(或流媒体)和离线模式, 如图1所示。在线模式可对网络和数据收集进行持续监视, 并依据网络舆情变化及时识别、校正或调整网络舆情信息源的布局, 其总体处理流程包括以下五个方面: 1) 用户选择一个自己希望完成舆情传播的网络平台, 或定

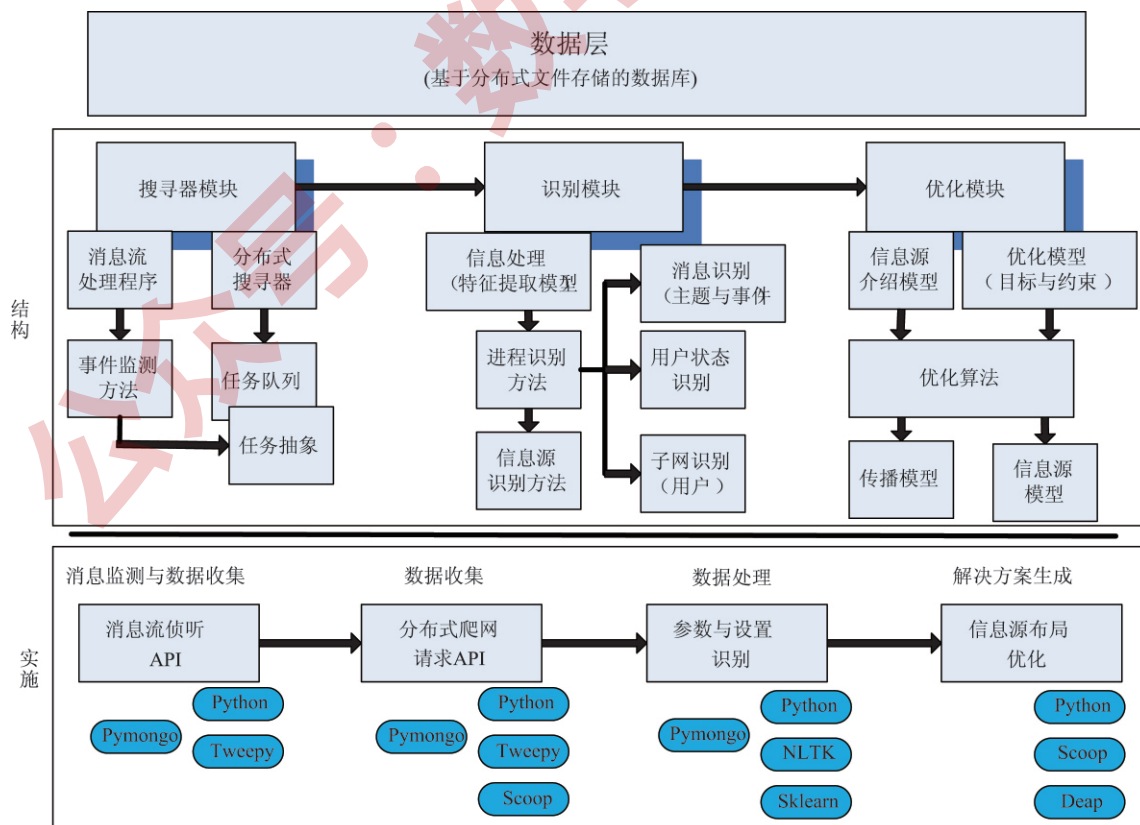


图1 网络舆情信息源优化模型

义允许模型修复网络的关键字; 2) 监视模块收集所有必须的数据; 3) 识别模块对现有主题、进程、用户状态以及由特定主题或进程形成的子网络进行识别, 并构建由优化模块使用的描述, 用户可以限制或指定模块必须处理的数据, 并明确执行此操作必须使用的模型; 4) 调用优化模块, 该模块采用识别模块生成的数据、目标描述和约束条件描述, 并执行对网络舆情信息源的布局优化; 5) 用户对网络舆情信息源的布局的前面或结尾进行描述, 该描述包括定位网络舆情信息源以及明确每个网络舆情信息源的参数。

搜寻器模块的目标是监视网络舆情并收集数据、消息、用户信息、转发或关注的事件。如果数据层关注并识别其包含的有关用户信息或信息不完整, 它会监测出现的消息流并转发、提取诸如用户记录和表单任务之类的舆情信息, 然后将这些任务分派到分布式搜寻器, 分布式搜寻器将根据可用的网络速率来收集所需的数据。利用这些收集的数据可以恢复网络拓扑、进程和用户特征。识别模块的主要目标是恢复和识别所有参数、设置和实体, 比如进程和网络拓扑等, 而进程和网络拓扑是对监视模块收集的网络舆情数据进行优化必不可少的。用户可以决定模块必须执行的进程以及必须使用的模型, 用户也可以插入自己设计的应用模型, 并用其替换现有的模型, 比如当前用户可以使用增量聚类方法来完成主题和进程识别。社交网络用户对应着现实中的人, 因此耦合网络节点需要采用双向连接模式^[17]。优化模块获取的数据由识别模块生成, 可通过约束条件与优化目标为网络舆情信息源布局提供一系列解决方案, 该模块基于 NSGA-II 遗传算法进行改进。NSGA-II 遗传算法可对网络舆情信息源在网络中的位置和特征集进行识别, 并使用单点交叉操作和多类型突变, 如添加、删除或更改配对信息等。为了评估该模块形成解决方案的质量, 该模块在对微信上的信息传播过程进行仿真。

该网络舆情信息源优化模型主要有以下优点:

1) 具有完整的解决方案, 可以处理网络舆情信息源布局优化所需的所有步骤, 主要包括收集所需舆情数据、进程、状态标识以及信息源布局优化等; 2) 具有灵活的模块化结构, 允许应用和替换不同的模型和方法(如进程、信息源模型、聚类和优化方法等); 3) 可有效评价解决方案的质量, 该模型将网络舆情信息源的拓扑特性作为信息源的单个特性来评价模型形成解决方案的质量。

3 网络舆情信息源布局模型优化方法

网络舆情信息源布局优化问题可以理解为多目标优化问题, 因为它可以优化网络舆情信息分发过程的许多不同目标, 如对微信转发数量进行计数的转发器, 转发器的增长速度、成本以及网络带宽等。为了优化网络舆情信息分发过程的一般性目标, 本研究主要针对转发器数量和成本进行分析。

假设让 $N = \{ (V, E) \}$ 为社交网络的有向图, V 代表网络节点集合, E 代表边的集合。假设让 $G = \{ G_1, G_2, \dots, G_n \}$ 作为目标函数的集合: $\forall 1 \leq i \leq n, F_i$ 为获取目标 i 的值和理想目标值 r_i , K_i 为平滑系数。需要注意的是每个 F_i , 一般情况下, F_i 值取决于网络拓扑、网络舆情信息源布局和传播的时间长短: $F_i = f_i(N, S, T)$ 。通过仿真可以得出特定目标值, 如微信转发数量进行平均计数。网络舆情信息源布局优化的目标是找到信息源 S' 的配置以达到其功能的

最大化: $S' = \argmax \sum_{i=1}^n w_i G_i(2)$, $S' = \{ (v_i, c_j, m_k) \mid \forall i, j, v_i \in V, c_j \in C, m_k \in R \}$, 这表示网络舆情信息源布局将信息源的特征映射到网络节点中; V 表示位置顶点集; C 表示信息源配置集; m_k 表示配对成本; w_i 表示权重系数, 它可表明单个优化目标的重要性。每个信息源由其在网络中的位置(v_i)和配置(c_j)所决定, 它们可以描述信息源的各个特征(如微信消息发布的频率和消息内容的质量等)。需要指出的是, 与特定信息源相关的优化目标值可由微信用户自主设定, 但这在一定范围内会使系统优化模型失去作用, 因为有些信息源目标会引起争议, 这就需要目标进行优化, 但在优化目标之前建立有效的解决方案并从中选择关键节点。为了处理目标优化任务, 遗传算法可用于有效识别单个舆情信息传播源的位置及其配置。

综上所述, 现有的网络舆情传播模型认为信息源只是一个促使网络信息传播的初始节点, 此类信息源通过其网络拓扑特性(如关注度)进行区分, 并且通过这些特性评估其潜在成本。因此, 网络舆情信息源的个体特性对其布局优化非常重要, 这主要体现在信息源成本、消息频率和产生消息的质量这三个方面。消息频率有助于调节信息源对网络舆情的影响强度^[18]。产生消息的质量是影响舆情信息重新发布概率的重要因素之一。相对于微信主题和内容模糊舆情信息来说, 格式规范且内容丰富的消

息更容易被频繁转发,这对舆情信息传播会产生明显的影响^[19]。但网络舆情信息的质量取决于信息源作者的编写能力和转发信息的微信用户。在真实的网络环境里,可以通过分析微信消息的传播级联来评估微信用户的质量。近似值能基于网络舆情信息源的个体特性对微信消息位置和配置进行优化,这是解决这一问题的最理想的方案。网络舆情信息源成本通常是根据微信消息关注者的数量和转发的平均数进行估算。

现有的网络舆情信息源 SIS 模型通常以 α 概率来描述信息源从“S”型状态到“I”型状态的转换^[20]。但在微信中有多个信息源的情况下,应有多个 α (如 α_1 、 α_2 、 α_3 等) 来表示特定信息源或转发信息源的网络节点从“S”型状态到“I”型状态的转换。但在具体实践中很难精确表明一个网络节点应该使用 α 值的大小,扩展 SI 模型可以很好解决这个问题。与 SIS 模型相比,扩展 SIS 模型使用单个的消息进行传播,消息是网络舆情信息传播的基本载体,转发是一种通过微信平台大规模传播消息的通用机制。因此,转发消息的微信用户支持信息传播,并且也会表达自己对消息内容感兴趣。在扩展 SIS 模型中,不同信息源的微信消息会有不同的转发概率,每个信息源都根据其频率生成消息,每条消息均按照常规 SI 模型进行传播,并具备从其信息源获得的 α

值。如果微信消息处于敏感状态,这意味着微信用户不能转发任何该方面的内容(或根据期望值未进行大规模转发)。当有新的消息发送给微信用户时,微信用户可以将其转发,但消息在微信用户发布或转发后存活时间也是有限的^[23]。

图 2 显示了两个不同信息源的微信消息传播过程,每一条消息都将可能推送到微信用户的信息源中,微信用户通过订阅该消息就可以得到并经常阅读或推送该消息。信息源产生的微信消息越多,其传播消息的范围和受众就越多。网络舆情信息传播模型在实施过程中需要考虑以下四个因素: 1) 信息源的个别特性可能会严重影响整个传播过程和寻找最佳解决方案。2) 微信用户的实际变化,如要转变微信用户为消息受众状态的转发数。为了更精确地评估微信用户传播消息呈现的关注程度,可以用一个以上的转发数来表征微信用户的受众状态。网络节点的大小和颜色反映了它们的关注程度,网络节点越大和越暗,其关注度越高。3) 该模型基于消息频率和内容质量变化控制消息传播,这有助于潜在的调节整个舆情消息传播的关注程度(如微信消息评论的形成)。4) 模拟不同进程的网络舆情信息传播,特定的微信用户可能会同时受到多个进程的影响。因此,处于多个受众状态或进程会对微信用户传播网络舆情消息造成一定程度的干扰。

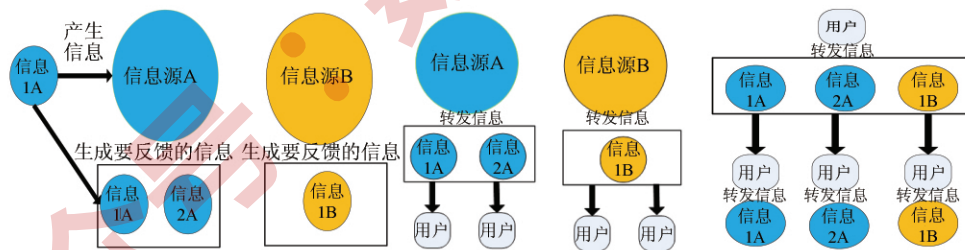


图 2 两个不同信息源的微信消息传播过程

网络舆情信息源优化模型框架最后一个环节就是源代码引入模型,该模型定义了信息源如何与网络中的具体节点进行关联。具体来说,可以通过以下方法进行关联: 1) 雇用现有微信用户作为信息源作者来编写具有其特性的消息。2) 雇用微信用户推送专门撰写的消息或抓取现有微信用户来推送消息。在这种情况下,生成消息的质量不是由微信用户自己决定的,而是可以推送到具有可用特性的网络节点上(如可生成人工节点并用它替换现有网络节点)。3) 创建新的网络节点并将其链接到网络中。在这种情况下,所产生消息的质量具有可用的

特征,并可以链接到一组微信用户,甚至可以通过额外的费用购买他们的转发的消息。

4 实验设计与验证

4.1 实验设计

为了验证本研究提出的网络舆情信息源优化模型,本文将应用于优化从微信社交网络获取关于新型冠状病毒疫情的信息源评论数和转发数数据。该模型选取从微信公众服务号和朋友圈获取的新型

冠状病毒疫情相关主题消息,总体抽样数据收集时间为 2020 年 3 月 1 日到 7 月 31 日,基于便利抽样分类标识这些变量数据并建立数据表,并利用 Matlab 对模型进一步求解验证,并对上述理论分析进行数值仿真。该实验基于疫情防控措施下的仿真验证信息源特征优化模型对社交网络舆情传播的防控效果。研究变量为信息源个体特征,主要分为疫情防控常识、政府发布信息、虚假信息等信息类型。本研究选取信息源规模为 $N = 100$ 个监测节点,信息源特征因素的取值范围均为 $[0, 1]$ 。抽样框的选择符合总体组成的特征。为了降低微信其他信息对实验的干扰,需要基于以下原则对数据进行清洗: 1) 去除未获得信息源身份认证的数据; 2) 去除因发布者、内容、转发数和评论数完全相同的数据; 3) 去除不能识别到的微信推送者粉丝信息的数据。

本研究通过聚类算法实现对不同特性的舆情消息进行快速分类识别,该实验选择了几个疫情防控子主题及其网络拓扑来对其进行优化,这主要反映在微信用户对疫情防控传播以朋友圈转发数量与所关注此类消息的数量。疫情防控子主题扩展性信息应该出现在已经对这种信息表现出兴趣的微信用户朋友圈。疫情防控第一个子主题主要围绕以下最受关注的关键词收集舆情信息数据: 抗击疫情、预防疫情、疫情健康生活等,其中核心的关键词是“抗击疫情”,该舆情信息由 6 508 个网络节点数据组成。疫情防控第二个子主题是疫情数据通报,其核心信息是最新疫情信息通报“国内外最新确诊病例和死亡病例”、“国内中高风险地区”。该舆情信息由 2 621 个网络节点数据组成,所有数据均在微信平台上进行收集。该系统监控模块收集 2020 年 3 月 1 日到 7 月 31 日的数据集进行标签。由于微信网络节点具有多个信息传播方式,因此高关注度的网络节点往往在舆情信息传播效果上具有较高的影响力。基于此,该实验利用近似值进行舆情信息传播成本估算: 成本 = 舆情信息 × 转发数。该公式表明,舆情信息影响程度越高,此信息源的微信推文转发数就越多。应该注意的是,舆情信息关注度较小的网络节点也可能对受众具有较高的影响力,这主要是由于微信内容质量好以及可信度高。如果舆情信息关注人数相同,这表明这些信息具有相同的转发量和潜在影响力。

4.2 实验验证及结果

第一个验证实验基于第一个子主题的网络节点

信息利用多组转发器检验该优化模型解决方案的质量。该实验利用遗传算法和贪心算法比较优化模型解决方案,该算法通过减少网络关注度来对网络中的所有节点进行排序,并选择关注度排名为第一个节点,从而形成优化模型解决方案。本实验中信息传播仿真设置如下: 每天在 8 时、12 时、16 时、20 时 4 次集中监测微信公众服务号和朋友圈信息,每次检测微信用户在微信中查看其消息并可以进行转发的时间,微信发文或转发后的可见性持续时间为 4 个小时,微信用户必须拥有至少 1 条所生成信息的转发。算法为信息源生成的每个解决方案均运行 30 次,最终采用其运行平均值。系统将信息关注度最高的 45 个网络节点和采用概率最高的信息源解决方案的成本视为舆情信息传播成本的上限。基于此可以推断,如果选择任何网络节点作为信息源,系统可以根据需要使用该节点来发布所编写或转发的消息。

图 3 ~ 图 6 描述了遗传算法和贪心算法为信息源生成的优化模型解决方案成本测试结果,两种方案分别用不同颜色进行区别,该算法按照网络节点的关注度从大到小对节点进行排序并进行仿真实验。通过比较图 3 ~ 图 6 中的正面曲线可以看出,遗传算法产生的解决方案要比贪心算法好一些。在扩频过程中,遗传算法仅用了 6 个单元数据获得 225 个微信用户,而贪心算法需要用 13 个单元数据实现相同的结果。遗传算法可以在成本上获得以下解决方案改进效果: 图 3 为 25.26%,图 4 为 22.65%,图 5 为 33.98%,图 6 为 31.82%,该结果通过设置的最大可用成本进行了标准化规范。

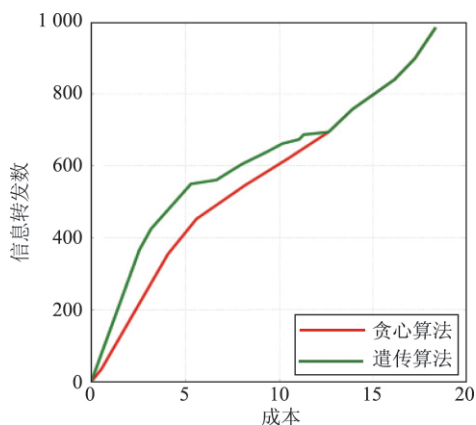


图 3 不同信息源的布局优化效果

图 4 和图 5 显示了微信信息发文在 1 天和 2 天后的转发信息量与成本的关系趋向,其舆情信息持续可见时间会逐渐减少。但精选的可信度高信息源

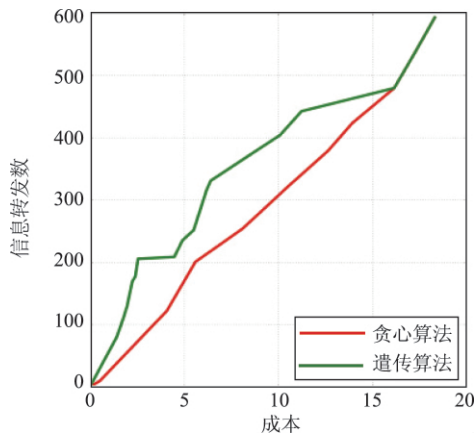


图4 微信发文1天后的可见性效果

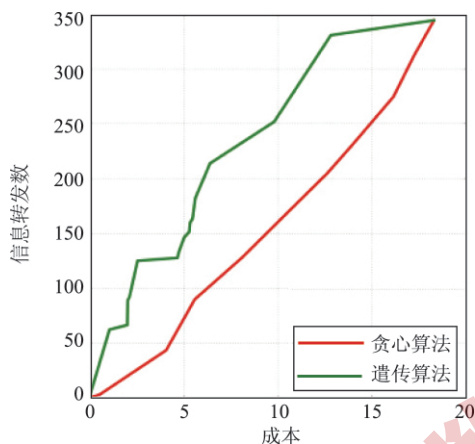


图5 微信发文2天后的可见性效果

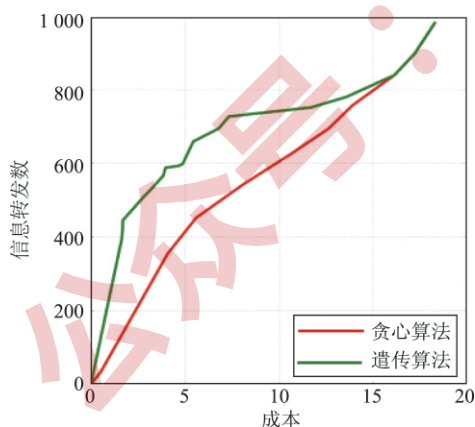


图6 信息源数量最大值的优化效果

推送到微信中会提升其关注度,被微信用户转发的可能性增加,这会导致非线性转发关联信息量的增长。使用可信度高的信息源可以获得更多的信息转发数,贪心算法提出的优化模型解决方案在这种情况下可以发挥其信息源特征快速识别的优势。在成本较小的情况下,最好使用遗传算法设计的优化模

型解决方案,因为该算法可有效识别不同特征的网络舆情信息源改善传播效果。

第二个验证实验基于第二个子主题的网络节点信息利用多组转发器检验该优化模型解决方案的质量,该主题信息是最新疫情数据通报,其中转发了“国内疫情最新进展”的新闻信息,通过该实验验证该优化模型框架在不同条件下具有较好的适用性。实验目标是该解决方案通过信息转发计数确定信息关注度,其中18个网络节点的关注度最高,由这些关注度高的网络节点组成优化模型解决方案。这种方法有助于突破海量信息的限制,可以应用于优化网络舆情信息传播过程。通过图7可以看出,遗传算法能够找到另一种优化模型解决方案,虽然其信息转发速度较慢,但网络舆情传播过程中涉及的用户数据、信息的评论数和转发数等关键指标却大致相同,该优化模型解决方案在网络舆情信息传播的效果具有较好的适用性。

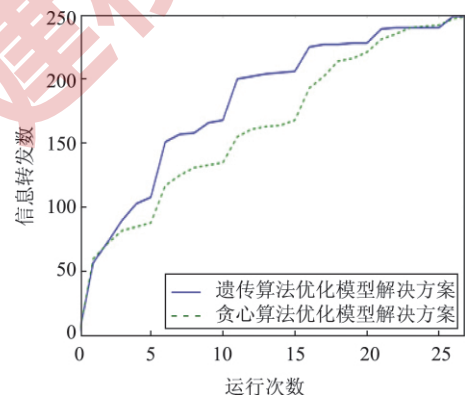


图7 优化调整的最终结果

5 结 语

本文基于信息源的个体特征提出了一个通用的网络舆情信息源布局优化模型,该模型基于微信平台利用遗传算法和贪心算法设计网络舆情传播优化模型解决方案,使用“SIS”模型评价信息源个体特征及传播效果,并针对新冠疫情防控舆情进行仿真,验证该优化模型的可靠性。实验研究表明,该优化模型可有效评价用户数据、信息的评论数和转发数等关键指标,该优化模型提供的解决方案最高可以提升33.96%。研究结果还表明,使用不同特征的网络舆情信息源可以降低相同特征识别的成本,从而为用户提供更好的网络舆情的传播路径。因此,在网络舆情信息成本条件固定的情况下,使用不同

特征的信息源会获得最理想的传播效果。

由于该模型框架采用模块化设计和灵活性架构,系统扩展性能好,在未来的研究工作中针对更复杂的网络舆情传播模型进行深入分析,以设计与开发更有效的网络舆情传播优化模型解决方案。在舆情信息传播优化模型具体实施中,由于微信用户两极分化和其他影响因素而导致的信息转发、修改和删除概率会随时变化,对这方面的信息进行及时追踪和优化对舆情信息传播效果具有重要价值。虽然这会导致微信舆情信息收集过程中数据量的海量增长,但运用目前成熟的大数据技术可以对这些数据进行分类和特征识别,不断提高对舆情信息的识别能力和优化效果。长远来看,社交网络应从公众情绪、媒体责任、政府治理提高对网络舆情传播影响力的可控性^[21]。

参考文献:

- [1] 林芹,郭东强. 优化 SIS 模型的社交网络舆情传播研究—基于用户心理特征[J]. 情报科学, 2017, 35(3): 53-56.
- [2] LIU Z, LIU L, LI H. Determinants of information retweeting in microblogging [J]. Internet Research, 2012, 22(4): 443-466.
- [3] 张玥,孙霄凌,浦正宁,等. 微博舆情传播影响因素研究[J]. 情报资料工作, 2014, (3): 59-64.
- [4] 胡媛. 微博客中基于时序的非正式信息流机制研究—以 sina 微博为例[J]. 图书情报知识, 2011, 142(4): 111-117.
- [5] 彭程,祁凯,黎冰雪. 基于 SIR-EGM 模型的复杂网络舆情传播与预警机制研究[J]. 情报科学, 2020, 38(3): 145-153.
- [6] CHA M, HADDADI H, BENEVENUTO F, et al. Measuring User Influence in Twitter: The Million Follower Fallacy [J]. ICWSM, 2018, 10(30): 10-17.
- [7] TANG X, QUAN Y, MIAO Q, et al. Intelligent Computation in Big Data Era [M]. Berlin: Springer Berlin Heidelberg, 2019.
- [8] 贾盼斗,尹春华. 区块链社交网络信息传播特征及规律研究[J]. 情报科学, 2020, (7): 104-109.
- [9] 王晰巍,邢云菲,张柳,等. 社交媒体环境下的网络舆情国内外发展动态及趋势研究[J]. 情报资料工作, 2017(4): 6-14.
- [10] LI S, ZHU Y, LI D, et al. Rumor restriction in online social networks [C]//2013 IEEE 32nd International Performance Computing and Communications Conference (IPCCC). New York: IEEE Press, 2018, (12): 1-10.
- [11] CHEN W, COLLINS A, CUMMINGS R, et al. Influence maximization in social networks when negative opinions may emerge and propagate [C]//The 2011 SIAM International Conference on Data Mining. Mesa: ACM, 2019: 379-390.
- [12] LESKOVEC J, KRAUSE A, CARLOS G, et al. Cost-effective outbreak detection in networks [C]//Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Boston: ACM, 2017: 420-429.
- [13] HE X, SONG G, CHEN W, et al. Influence Blocking Maximization in Social Networks under the Competitive Linear Threshold Model [J]. SDM, 2018, (1): 463-474.
- [14] CHEN W, WANG Y, YANG S. Efficient influence maximization in social networks [C]//Proceedings of the 15th ACM SIGKDD International Conference on Knowledge discovery and data mining. Beijing: ACM, 2019: 199-208.
- [15] PEZZONI F, JISUN A, PASSARELLA A, et al. Why do I retweet it? An Information Propagation Model for Microblogs [M]. New York: Springer International Publishing, 2019.
- [16] 宾晟,孙更新,周双. 基于区块链技术的社交网络中舆情传播模型 [J]. 应用科学学报, 2019, 37(2): 191-202.
- [17] 魏静,丁乐蓉,朱恒民,等. 全媒体时代基于情感和耦合网络的舆情传播模型研究 [J]. 情报杂志, 2020, 39(8): 110-116.
- [18] ROMERO D, GALUBA M, ASUR W, et al. Machine Learning and Knowledge Discovery in Databases [M]. Berlin: Springer Berlin Heidelberg, 2018.
- [19] CHOI J, CROFT W B, KIM J Y. Quality models for microblog retrieval [C]//In Proceedings of the 21st ACM International Conference on Information and Knowledge Management. Maui Hawaii: ACM, 2019: 1834-1838.
- [20] 邓春林,何振,杨柳. 基于 SIS 模型的网络群体性事件传播及防控研究 [J]. 情报杂志, 2016, 35(5): 79-85.
- [21] 汪明艳,陈梅. 社交媒体网络舆情传播影响力研究综述 [J]. 情报科学, 2017, 35(5): 171-176.

作者简介



杨文阳(1980—),男,河南人,博士,副教授,研究方向为数字化学习、网络舆情传播;

E-mail: ywy80910@163.com