[](http://crossmark.crossref.org/dialog/?doi=10.1145%2F3428690.3429156%26domain=pdf%26date_stamp=2021-01-19)**Predicting Emotion Dynamics Sequence on Twitter via Deep Learning Approach**

Debashis Naskar

Computer Technology Group Artificial Intelligence

(GTI-IA), Universitat Politècnica de València, Spain

debashis@drtc.isibang.ac.in

Eva Onaindia

Valencian Research Institute for Artificial Intelligence

(VRAIN), Universitat Politècnica de València, Spain

onaindia@dsic.upv.es

Miguel Rebollo

Valencian Research Institute for Artificial Intelligence

(VRAIN), Universitat Politècnica de València, Spain

mrebollo@upv.es

**ABSTRACT**

当前诸多研究，大多聚焦对网络舆情传播的特点归纳、模型构建以及影响力用户的识别，本文通过构建高影响力人物情感演化仿真模型，就高影响力人物相关事件传播过程、传播特征以及信息传播周期进行分析。探究高影响力人物作为舆情信息传播主体的行为特征和情感变化，旨在预测他们的行为或情感，理解他们在特定事件中的态度和立场，从而为政府、媒体等机构提供有价值的信息参考，帮助他们制定更加合理的舆论引导策略和公共关系管理方案。

**CCS CONCEPTS**

• **Information systems** → **Sentiment analysis**; Emotion dynam- ics; • **Computing methodologies** → Model verification and vali- dation;.

**KEYWORDS**

高影响力人物 网络舆情 舆情仿真 情感分析s

**ACM Reference Format:**

DebashisNaskar, Eva Onaindia, Miguel Rebollo, and Sanasam Ranbir Singh. 2020. Predicting Emotion Dynamics Sequence on Twitter via Deep Learn- ing Approach. In The 18th International Conference on Advances in Mo- bile Computing and Multimedia (MoMM ’20), November 30-December 2, 2020, Chiang Mai, Thailand. ACM, New York, NY, USA, [5](#bookmark1) pages. [https:](https://doi.org/10.1145/3428690.3429156) [//doi.org/10.1145/3428690.3429156](https://doi.org/10.1145/3428690.3429156)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

MoMM ’20, November 30-December 2, 2020, Chiang Mai, Thailand © 2020 Association for Computing Machinery.

ACM ISBN 978-1-4503-8924-2/20/11. . . $15.00 <https://doi.org/10.1145/3428690.3429156>

Sanasam Ranbir Singh

Indian Institute of Technology Guwahati, India

ranbir@iitg.ernet.in

**1 INTRODUCTION**

在全球化背景下，国际舆论场日益复杂多变，任何一起负面舆情都可能对国家形象和国家利益造成严重影响。而越来越多的人习惯借助互联网通过twitter、YouTube等社交媒体平台来关注热点事件、发表评论、进行价值评判。因此，及时掌握社会舆情动态，探索舆情产生特点和变化规律，识别出社交网络中具有高传播影响力的用户借以有效引导，针对性地提出相应导控策略，对于营造风清气正的国际舆论环境，维护国际社会稳定稳定，推动文明发展具有重要的理论意义和应用价值。

近几年，国内外学者在网络舆情方向展开了大量研究，何英等使用Matlab 构建网络舆情扩散方向估计模型，其提出的算法可以提高网络舆情突变扩散方向的判断精确度；赵磊等采用BP神经网络的理论构建模型使得模型能够在对舆情事件热度变化趋势有更精确的仿真效果；赵凯构造了两类媒体作用下舆情传播的传染病模型，进而利用微分方程的稳定性理论对模型进行了动力学分析。Yang等人基于国外的信息传播媒体twitter为研究对象，通过对该信息传播网络的用户进行传播行为的定量分析，提出了一种新的信息传播模型，该模型通过对网络对象的节点影响进行评估，从而预测该信息传播网络进行发散状态的趋势，从而对其进行预测。魏静等将情绪感染理论融入SIR模型，构建了考虑亲密度及情感倾向的社交网络舆情传播模型，并基于无标度网络对舆情传播影响因素进行了分析。突发事件网络舆情传播过程中，关键用户（影响力用户、意见领袖）或网络“大V”等往往扮演着重要的角色，影响着事件的发展和舆论的走向，挖掘网络中影响力节点具有重要的现实意义。吕琳媛和任晓龙等人对这方面工作做了较为详尽的综述，介绍了复杂网络领域具有代表性的多种重要节点挖掘方法，详细比较了各种方法的计算思路、应用场景和优缺点。谢英香等则通过对社会网络分析法中的中心度的分析，利用MDS等方法，分析了虚拟社区中的用户的影响力，并进一步揭示该社区存在意见领袖现象。S.Jonnalagadda等则综合分析了点度中心性、点度中介性以及点度紧密性等反映中心的指标，从而发现了医学在线社区中具有较大影响的意见领袖。在舆情应对方面，曾润喜着重探索了网络舆情对群体性突发事件的影响与作用，谢耘耕等人结合微博舆论的传播特点，探索出微博舆论的生成演变总体过程，并针对影响舆论演变的五大要素，提出了政府舆论引导的四点建议和对策。

目前国内学者专门进行事件网络舆情演变建模和仿真的研究相对较少，因此，本研究拟以高影响力人物发表事件为例，对其事件中涉及人物的情感进行仿真模拟，以模拟真实舆情传播过程，揭示网络事件传播所具有的一定规律。首先，通过收集指定事件发生过程中的相关话题数据，研究其中网络舆情演化规律，分析人物主体在舆情信息传播中的属性特征。然后，构建网络舆情演化的仿真模型，从用户比例、演化次数、响应时长、影响力系数出发，探究舆情演化的影响因素。最后，基于上述研究结果，提出网络舆情演化的干预措施，为政府舆情治理提供决策参考。

The existing works [[4,](#bookmark7)[11] mostly concern about influence mod](#bookmark2)- els based on neighbors’opinions and their studies only focus on discrete categories of sentiment polarities. But it is also important to know which significant emotion plays a major role to influence

the users. Therefore, we focus on the influential role of neighbour emotions’and predict how emotions are dynamically transmitting from one state to another across different topics. To address the gap, we proposed a prediction model called Emotion-based User Sequential Influence Model (E-USIM) which capture neighbors’par- ticular emotional information from their long past history. In this model, we adopt Recurrent Neural Network (RNN) architecture and its variant Gated Recurrent Units (GRUs) [[5] which has an](#bookmark8) able to integrate the historical information with the new coming information for prediction. Our proposed model contains a set of primary emotions which have 90% accuracy to classify the emotions [[8]](#bookmark9). To exhibit the efficacy of the proposed model, we perform the experiments on the collected Twitter datasets on different topics.

**Table 1: Size of the real datasets**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Event/Topic** | **Categories** | **# of Tweets** | **# of Users** | **Avg # of Sent Tweets** |
| #COVID19 | Policy | 616343 | 1260 | 4 |
| #Brexit | Policy | 686434 | 2688 | 6 |
| #GrenfellTower | Accident | 136821 | 2297 | 3 |
| #SyriaGasAttack | Terrorism | 10823 | 557 | 3 |

**2 DATA PREPARATION AND EMOTION**

**STATE LABELING**

为分析国际对于涉华事件地情感，通过爬取twitter上关于四个事件地推文并创建了数据集，如图1所示。通过Bert模型将用户情感识别为正面/中性/负面三类。

BERT预训练模型先在一个大型的数据集上进行大量的预训练，然后在具体的下游任务进行Fine-Tuning，一举刷新了11项NLP任务的最新结果。该模型同样使用了Transformer特征抽取器进行特征抽取。BERT在预训练中使用“MLM”（MaskLanguage Model）和“NSP”（Next Sentence Prediction）将上下文语义融入到词向量中，MLM的具体做法是从语料中随机选择15%的单词抠掉，用[Mask]进行替换，然后根据上下文去预测被Mask的单词，这样就将单词前后的上下文都嵌入到了单词中，又避免了间接“看到自己”。NSP是指在语言模型预训练中，可以选择两种方式来构建训练样本：一种是按照语料库中句子的真实顺序连接两个句子，另一种是随机从语料库中选择一个句子并将其拼接到第一个句子的后面。除了进行Masked语言模型任务外，还要求模型进行句子关系预测任务，即判断第二个句子是否是第一个句子的后续句子。这样做的原因是很多NLP任务都需要进行句子关系的判断，而单词级别的预测无法涵盖这种句子层面的关系，因此增加这个任务有助于提高下游句子关系判断任务的表现。

(X, μ, σ) (1)

i1 

X =



i=1

where, X is the mean value of valence (similarly, mean value of arousal), μ and sigma are the word’s mean and standard deviation value of valence (equivalently for arousal) and the total number of emotional words is N.

**3 EMOTION-BASED USER SEQUENTIAL INFLUENCE MODEL**

在本节中，我们运用元胞自动机模拟了社交平台对于国际事件舆情演化仿真模型，其中我们加入了高影响力人物地影响力计算以匹配实际地影响力，从而预测高影响力人群的情感状态。

**3.1 元胞自动机**

基于舆论动力学理论，本研究采用元胞自动机模型来对“7·20 郑州特大暴雨”突发公共事件进行舆情演化分析。元胞自动机由具有元胞状态的元胞组合而成，在规 定的空间内与元胞周边的邻居按照一定规则进行演化的一组函数。

元胞作为最基本组成单元，通常是有限的，将它们分布在二维空间的离散网格上。每个元胞都具有各自的状态，不同建模情况下元胞状态也各有不同，元胞状态一般根据建模情况分为 0，1 的二进制类型或者是{s1，s2，…，sn}的整数型离散集，本文定义元胞状态为情感倾向。元胞空间是指元胞分布的空间网格集合。一般来说，元胞空间需注意其几何划分、边界类型和构型。定义元胞自动机的邻居模型为摩尔型，元胞的邻居则在之前的基础上扩充到八个相邻元胞。规则也被称为演化规则，是一种用于确定元胞空间中的元胞与其邻居当前状态 及下一时刻状态的动力学函数。规则是元胞自动机模型演化的关键所在，整个系统 基于演化规则在某一特定空间和时间上进行不断演变。规则的具体函数见公式

𝑆𝑖 𝑡+1 = 𝑓(𝑆𝑖 𝑡 , 𝑆𝑁 𝑡 )

（1.1）

其中，𝑆𝑖 𝑡+1代表元胞 i 在 t+1 时刻的状态；f 代表元胞自动机的更新规则；𝑆𝑖 𝑡代 表元胞 i 在 t 时刻的状态；𝑆𝑁 𝑡 代表 t 时元胞的状态合集。为了更加符合实际演化情况， 规则一般分为可逆规则、总和规则等。可逆规则是用于反映现实中某些具有各向同性 的物质，其好处在于使得规则总数减少一半；总和规则是指某元胞各邻域内元胞状态 总和为特定整数值，其优势在于极大减少了规则数。

对于一个元胞，它受到周围八个元胞的影响，计算公式为：

对于影响力系数来说，使用𝐸𝑖𝑗表示元胞坐标为（i，j）的用户对其他元胞的影响

力。其中，𝐸𝑖𝑗的值越大代表该用户越容易影响到周围的其他用户的元胞状态。

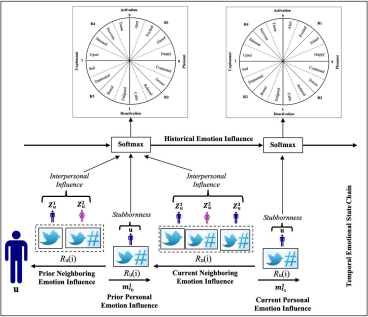
对于坚定性系数来说，除政府认证用户之外，其他三种类型的用户在不同演化次

数中的坚定性都会随机变化，使用𝑊𝑖𝑗表示元胞坐标为（i，j）的用户坚定性。因此，

该坐标的元胞在演化次数为 t 次时的情感倾向值见公式（4.3）。

𝑌𝑖 𝑗 (𝑡) = 𝑊𝑖 𝑗 × 𝑌𝑖 𝑗 (𝑡 − 1) + (1 − 𝑊𝑖 𝑗) × ∑ 𝐻𝑖 𝑗 (𝑡 − 1) / ∑ 𝐸𝑖 𝑗

Predicting Emotion Dynamics Sequence on Twitter via Deep Learning Approach MoMM ’20, November 30-December 2, 2020, Chiang Mai, Thailand



**Figure 1: Emotion-based User Sequential Influence Model**

根据模型参数设置情况，进一步明确各主体的模型演化规则。使用)表示演化次数为 t 次时该元胞的情感倾向值。的数值为各类用户的初始状态值。

对于政府认证用户，由于其具有坚定的政治立场，在开始演化之后元胞状态保持 1 不变。因此，该用户在演化次数为 t 次时所对应的元胞状态见公式（4.1）。

对于媒体认证用户的值域为[-1，1]。其中，当 0.09＜𝑌𝑖𝑗(𝑡)≤1 时，该用户的元胞状态为 1， 呈现积极态度；当-1≤𝑌𝑖𝑗(𝑡)＜-0.09 时，该用户的元胞状态为-1，呈现消极态度；当 -0.09≤𝑌𝑖𝑗(𝑡)≤0.09 时，该用户的元胞状态为 0，呈现中性态度。因此，该用户在演 化次数为 t 次时所对应的元胞状态见公式（4.2）。

最后，我们经过多次迭代得到各类情感倾向的人数随迭代次数的变化，以及预测高影响力人物的最终情感。.

**3.2 用户影响力评估模型s**

由于微博上存在着“互相关注”的风气,一些实际上并不很重要的用户由于大量关注他人而获得了较多的“回礼式”关注,提升了粉丝数量。因此真正有较高影响力的用户应该是关注用户数量相对较小而粉丝数量较多的用户。据此,关注用户数量应视为负向属性,用以修正模型。0经过上面的假设和分析,对于微博客用户U来说,重要性评分Socre(U)大体应与粉丝数量(FansU)、关注用户数量(FocusU)、发布微博总数(TU)和是否被验证(VU)这几个指标相关。本文的微博客用户重要性评分模型如公式(1)所示

Score(U)=A(FoucsU)×B(TU)×FansU×(1+αVU)   (1)

其中A函数为针对关注用户数量FocusU的一个修正函数。B为针对发布微博数量TU的一个修正函数。0对于A函数,目的是惩罚“互相关注”这种行为,所以设定一个门限值,若关注用户数目和粉丝数目之比高于门限值,则给予惩罚。如公式(2):

A(FocusU)={1+β⋅FocusUFocusUFansU>0.21 else   (2)

对于B函数,需要惩罚“话痨”用户,所以设定置信区间,进而“话痨”用户给予惩罚。如公式(3):

B(tU)=⎧⎩⎨⎪⎪⎪⎪⎪⎪tU tU<c1c1+χ(tU−c1)c1c2 c1<tU<c2c2+δ(tU−c2)c2max c2<tU   (2)

其中,根据本文数据集合的统计信息,c1和c2的取值分别为226和996,α和β取经验值0.2。

3.2. 1 Prior Neighboring Emotion Influence.

3.2.2 Prior Personal Emotion Influence.

3.2.3 Current Neighboring Emotion Influence.

MoMM ’20, November 30-December 2, 2020, Chiang Mai, Thailand

3.2.4 Historical Emotion Influence.

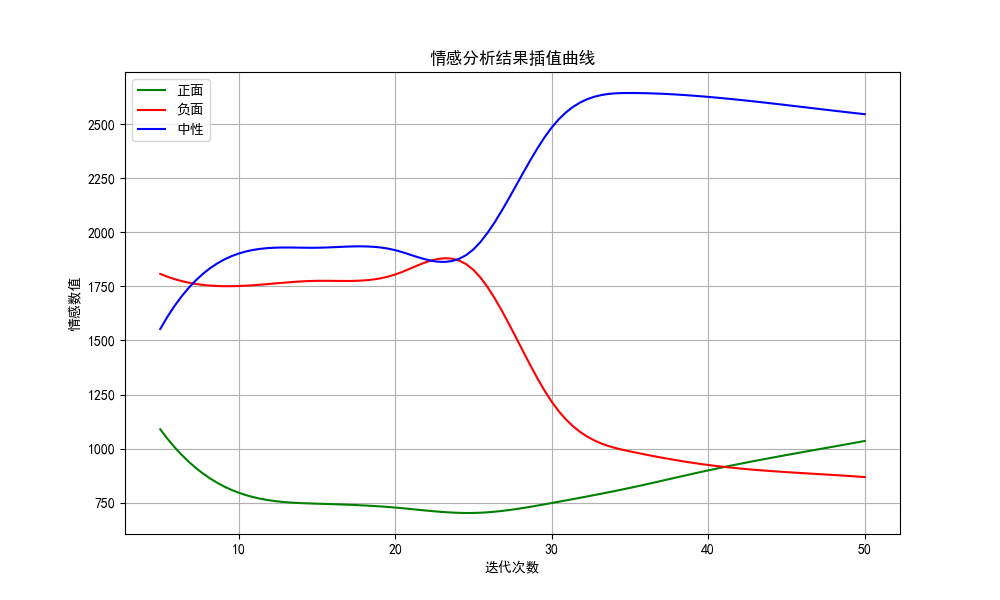
3.2.5 Recurrent Neural Network.

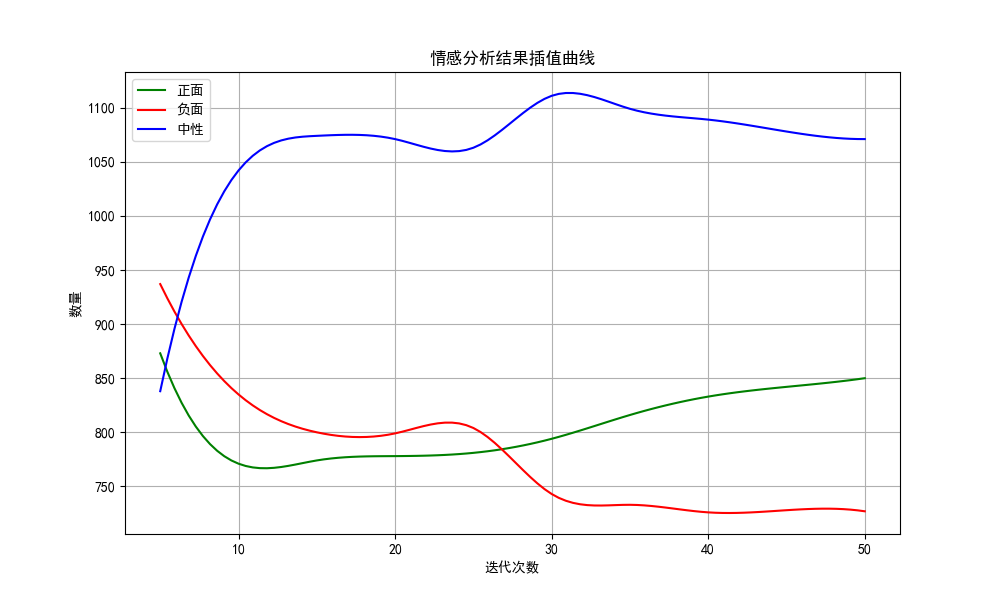
3.2.6 Gated Recurrent Unit.

Naskar, et al.

**3.3 Emotion-based Prediction**

为了预测涉华事件未来的舆情状态，我们通过多次迭代，首先针对新冠疫情事件，得到的迭代结果如图所示：





**3.4 Parameter Estimation**

To train our model, we utilise the gradients by using the Back Propagation Through Time (BPTT) approach [[14] and maximize](#bookmark18) the log-likelihood of the sequences of emotions. To implement this algorithm, we use Tensor-[Flow1](#bookmark17)open-source software library with a few lines of code.

L =  Σi log P(Su(i)| ↓ NRn,i(t), ↑ PRt (t)) (12)

**4 EXPERIMENTAL EVALUATION**

**4.1 Dataset Baselines**

For the purpose of our study, we collected our datasets with the Search-API related to four different events through the hashtags #BlackMoneyDebate, #Brexit, #GrenfellTower and #SyriaGasAttack respectively. It consists of 6.80K users over a total of 14.50M tweets. From the raw datasets, we filtered out users who sent 0 or only one message and tweets that did not show any emotion (null values of valence and arousal). Table [1](#bookmark10) shown the final figures after filtering. For validation of our methods, we compare it with the three baseline methods, i.e. DeGroot model [[6], Flocking model](#bookmark19) [[9] and Voter](#bookmark20) model [[10]](#bookmark21).

**4.2 Evaluation Metrics on Emotions Prediction**

Given the sequence of u’s emotion, we split the data into training dataset and test dataset according to the posting time. The training dataset is made by emotional sequence from the previous steps Cu,i(t − 1) for each user ui ∈ U. Since the length of emotional sequences are not the equal, therefore we predict the last opinion for each user u after learning from the training set.

The emotion prediction performance for all methods are eval- uated after optimizing from P(Su(i)|Cu,i(t)) to the sequence of emotion state Su(i) of the users. To evaluate the performance of our proposed method on a different topic, we have used two differ- ent measures of error. One is MSE (Mean Square Error):

t1 (Ru(i) − Su(i))2

(13)

MsE =

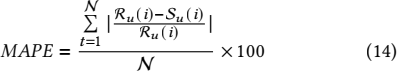
N

Another one is MAPE (Mean Absolute Percentage Error):

1<https://www.tensorflow.org/>

**Table 2: Performance metrics of emotions prediction for each dataset. The first each column of the each dataset is forecasting** **error in terms of MSE and second one is MAPE with percentage.**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Topic** | **#BlackMoneyDebate** | | **#Brexit** | | **#GrenfellTower** | | **#SyriaGasAttack** | |
| **MSE** | **MAPE(%)** | **MSE** | **MAPE(%)** | **MSE** | **MAPE(%)** | **MSE** | **MAPE(%)** |
| DeGroot | 1.3412 | 3.61% | 2.5112 | 2.80% | 1.1160 | 2.20% | 4.0420 | 7.34% |
| Flocking | 1.8551 | 4.58% | 3.3331 | 3.02% | 1.2751 | 2.84% | 4.0782 | 7.91% |
| Voter | 1.9321 | 5.37% | 4.3421 | 3.21% | 1.6342 | 3.92% | 6.0554 | 8.83% |
| **E-USIM** | **0.0521** | **1.32%** | **0.0415** | **0.94%** | **0.0513** | **1.44%** | **0.0811** | **2.42%** |



where for both measures, Ru(i) is the actual value, Su(i) is the predicted value and N is the number of users on each emotional sequence.

Table [2](#bookmark1)represents a comparative analysis of the prediction error (MSE and MAPE) of three state-of-the-art algorithms along with our proposed method.

Compare with other baseline methods, the prediction perfor- mance of DeGroot is consistently good for each dataset. But during the training phase, this model was iterated multiple times to con- verge and predict the emotion. Flocking is comparatively better than the voter model and this model updates the emotion of a user by calculating the average value of his/her neighbours. If any user does not have any neighbours, then this model makes difficult to predict the right emotion of that user. The performance of voter model is not much impressive among all methods. Since this model update user’s emotion randomly, thus it cannot judge the actual emotional sequence for prediction. Thus, the performance of this model is unsatisfactory. E − USIM performs consistently much better than all baseline methods over each dataset. If we look in an individual topic, E − USIM method for #Brexit perform well with minimum error (i.e. 0.0415 MSE and 0.94% of MAPE). Most importantly, E − USIM properly captures all the past history of emotion influence for each user. As a result, it provides better result in terms of forecasting emotion even at a distant future. Due insuf- ficient historical information over the network, it’s difficult to learn the actual influence for emotion prediction. Therefore, prediction performance may even harm the results for the topic with small dataset.

**5 CONCLUSION**

本文聚焦于高影响力人物，运用 Python 爬虫技术、SnowNLP情感分析算法和元胞自动机仿真建模等多种方法，充分获取twitter平台中高影响力人物发表事件相关的各类数据信息，并对网络舆情的几个基本问题进行了深入分析与探讨。探究高影响力人物的行为特征和情感倾向。此外，提出基于高影响力人物的舆情引导策略，以达到及时控制舆情，防止负面情绪传染，降低舆情危机可能性的目的。

根据仿真模型影响因素分析结果显示，高影响力对公众人物的情感有一定的影响，通过与其他方法对比，发现基于元胞自动机的仿真模型精确度更高，能够更好的模拟用户的情感变化。

需要认识到的是，网络舆情治理离不开多方努力。政府机构一方面需要通过提高舆情响应速度来对突发公共事件进行积极引导，保证在舆情事件传播全周期内信息的公开透明。另一方面，社交媒体应当担负舆情谣言治理的责任，处罚违法账号，塑造积极健康的网络舆情环境。广大人民群众作为参与者则更应该强化自我信息安全意识，提供信息素养能力，学会分辨不实信息，不能盲目听从不良信息。在面对浩如烟海的网络世界，人们需要能“以不变应万变”，坚定不移跟着党 的领导，共同营造和谐友爱、积极向上的互联网环境。

.

**6 ACKNOWLEDGEMENTS**

This work is partially supported by the Spanish MINECO project TIN2017-88476-C2-1-R.

**REFERENCES**

[1]黄炜, 余辉, 李岳峰, 陆薇.网络舆情事件演化的仿真实践研究.现代情报. 2017, 37(8): 65-73.

[2]强韶华, 吴鹏. 突发事件网络舆情演变过程中网民群体行为仿真研究 . 现代图书情报技术, 2014, 30(6): 7l-78.

[3] White S H, del Rey A M, Sanchez G R. Modeling epidemics using cellular automata [J].Applied Mathematics and Computation, 2007, 186(1):193-202.

[4]张明新. 国内网络舆情建模与仿真研究综述[J]. 系统仿真学报, 2019, 31(10): 1983-1994.

[5] 周思锦, 陈棣成, 涂耿, 姜大志. 基于个性化和记忆机制的多模态情感计算模型[J]. 系统仿真学报, 2022, 34(4): 745-758.

[6] Eksin C. Control of stochastic disease network games via influential individuals [C]. In: 2019IEEE 58th Conference on Decision and Control (CDC). 2019:6893-6898.

[7]余雷,薛惠锋,高晓燕,等.基于元胞自动机的传染病传播模型研究[J].计算机工程与应用,2007,(02):196-198+237.

[8]HU Changjun, XU Wenwen, HU Ying, FANG Mingzhe, LIU Feng. Review of Information Diffusion in Online Social Networks[J]. Journal of Electronics & Information Technology, 2017, 39(4): 794-804.

[9]Hui Li, Jiang-Tao Cui, Jian-Feng Ma. Social Influence Study in Online Networks: A Three-Level Review[J].Journal of Computer Science and Technology, 2015, 30(1): 184-199.

[10] Gergely Palla，Imre Derényi，Illés Farkas，and TamásVicsek，“Uncovering the overlapping community structure of complex networks in nature and society，”[J]. Nature，435 : 814-818，Jun. 2005.

[11]张侃.自媒体时代网络舆情演化与治理策略分析——基于框架理论的视角[J].传媒论坛,2024,7(19):32-35.

[12]王茜,何轲,贾伟.基于主题—情感融合分析的高校突发网络舆情管理研究[J].网络安全技术与应用,2024,(10):89-94.

[13] [1]高晓宁,张孟伟,杨蕴琦,等.基于多案例数据的突发事件网络舆情传播组态路径及其影响研究[J/OL].现代情报,1-21[2024-10-26].

[14]徐会杰,靳华.面向舆论场的网络舆情传播模型建模与仿真[J].计算机仿真,2024,41(07):399-405.

[15]Macdonald D .White Racial Identity and Preferences for (Non) White Immigrants in the United States[J].American Politics Research,2024,52(6):677-688.

.