\



Social bots and mass media manipulated public opinion through dual opinion climate

Chun Cheng(程纯), Yun Luo(罗云), Chang-bin Yu(于长斌), and Wei-ping Ding(丁卫平)
Citation: Chin. Phys. B, 2022, 31 (**1**): 018701. DOI: 10.1088/1674-1056/ac0baa
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Fitness of others' evaluation effect promotes cooperation in spatial public goods game

Jian-Wei Wang(王建伟), Rong Wang(王蓉), and Feng-Yuan Yu(于逢源)
Chin. Phys. B, 2021, 30 (**12**): 128701. DOI: 10.1088/1674-1056/ac05ae

Benefit community promotes evolution of cooperation in prisoners' dilemma game

Jianwei Wang(王建伟), Jialu He(何佳陆), Fengyuan Yu(于逢源), Wei Chen(陈伟), Rong Wang(王蓉), Ke Yu(于可)
Chin. Phys. B, 2019, 28 (**10**): 108703. DOI: 10.1088/1674-1056/ab3f1e

Modeling for heterogeneous multi-stage information propagation networks and maximizing information

Ren-Jie Mei(梅人杰), Li Ding(丁李), Xu-Ming An(安栩明), Ping Hu(胡萍)
Chin. Phys. B, 2019, 28 (**2**): 028701. DOI: 10.1088/1674-1056/28/2/028701

The most common friend first immunization

Fu-Zhong Nian(年福忠), Cha-Sheng Hu(胡茶升)
Chin. Phys. B, 2016, 25 (**12**): 128702. DOI: 10.1088/1674-1056/25/12/128702

Ising model on evolution networks and its application on opinion formation

Zhu Xiao-Long, Zhang Hai-Tian, Sang Jian-Ping, Huang Sheng-You, Zou Xian-Wu
Chin. Phys. B, 2014, 23 (**6**): 068701. DOI: 10.1088/1674-1056/23/6/068701

Social bots and mass media manipulated public opinion through dual opinion climate

Chun Cheng(程纯)[1,3], Yun Luo(罗云)[2,3], Chang-bin Yu(于长斌)[3], and Wei-ping Ding(丁卫平)[1,†]

[1] *School of Information Science and Technology, Nantong University, Nantong* 226019, *China*
[2] *School of Computer Science and Technology, Zhejiang University, Hangzhou* 310024, *China*
[3] *School of Engineering, Westlake University, Hangzhou* 310024, *China*

(Received 19 April 2021; revised manuscript received 6 June 2021; accepted manuscript online 16 June 2021)

Opinion dynamics models based on the multi-agent method commonly assume that interactions between individuals in a social network result in changes in their opinions. However, formation of public opinion in a social network is a macroscopic statistical result of opinions of all expressive individuals (corresponding to silent individuals). Therefore, public opinion can be manipulated not only by changing individuals' opinions, but also by changing their states of expression (or silence) which can be interpreted as the phenomenon "spiral of silence" in social psychology. Based on this theory, we establish a "dual opinion climate" model, involving social bots and mass media through a multi-agent method, to describe mechanism for manipulation of public opinion in social networks. We find that both social bots (as local variables) and mass media (as a global variable) can interfere with the formation of public opinion, cause a significant superposition effect when they act in the same direction, and inhibit each other when they act in opposite directions.

Keywords: public opinion, social bots, mass media, spiral of silence

PACS: 87.23.Ge, 02.50.Le **DOI:** 10.1088/1674-1056/ac0baa

1. Introduction

Participants in online social network discussions tend to divide into groups with opposite opinions,[1–3] which represent their believes or behaviors or attitudes towards a certain issue or phenomena. From a sociological perspective, the polarization of opinions is caused by a phenomenon called individuals' "selective exposure",[4–6] which widely exists in the social environment. Research shows that convenient and open access to information and expression often leads to the opposite effect: individuals tend to create connections with those who share similar opinions, which will also lead to the creation of echo-chambers.[6,7] How public opinion is formed by individuals with binary opinions in social networks and how it will be affected by intervene of social bots and medias cause our research interests.

Today's social networks evolve into a hybrid space in which human users, social bots, and mass media coexist.[8,9] On the one hand, the automated programs driven by algorithms-social bots play a role in social networks, in details, research shows that 19% of all interactions in networks are directed from social bots to humans.[10] Evidence from multiple investigations shows that social bots (as a type of zealots[11]) are used to manipulate public opinion.[12] Therefore, influence of social bots cannot be ignored when studying opinion formation and expression in social networks. On the other hand, mass media is the primary means to communicate with the majority of individuals in the general public,[13] and also play an important role in guiding public opinion.

For a long time, classical opinion dynamics models based on multi-agent methods can effectively describe the process of opinion interaction, negotiation, and decision making procedure among group members,[14–22] which means that they are very suitable for small group networks. However, classical models focus on the changes in individual opinions caused by interaction behaviors, while ignoring other important factors that lead to the formation of public opinion at the macro level. Therefore, a new exploration of public opinion formation in large social networks is necessary, for example, Noelle-Neumann's spiral of silence theory[23] provides a new perspective for studying the formation of public opinion. The changes between the expression and silence states of all individuals in the network determine the final formation of public opinion, as the macroscopic statistical result of the opinions of all expressive individuals. Noelle-Neumann summarized two influential factors for individual perceptions of the majority opinion:[23] "first hand observation of reality and observation of reality through the eyes of the media", which has been generalized to the idea of dual opinion climate by Ref. [24]. This theory is widely used in research of political elections and polarization of public opinion, and the dual opinion climate has inspired us to focus on the differences in public opinion manipulation mechanisms between social bots and mass media.

Different from the classical opinion dynamics models, which only focus on the factors of the change of individuals' opinion values in the group. In this paper, inspired by the spiral of silence theory, we pay attention to the change of indi-

†Corresponding author. E-mail: dwp9988@163.com

© 2022 Chinese Physical Society and IOP Publishing Ltd

vidual states (expression or silence) and attempt to re-examine the process of opinion formation. A key notion is that not everyone expresses their opinions in social networks, and the formation of public opinion only depends on those expressed opinions. Furthermore, people's willingness to express themselves is influenced by expressed opinions rather than those unexpressed, even though the latter exist and remain constant. By introducing heterogeneous individuals such as mass media and social bots, we attempt to use the multi-agent method to establish an opinion dynamics model, to capture the interaction behavior of different individuals, study the role of the dual opinion climate formed by social bots (as local variables) and mass media (as a global variable), and reveal the mechanism of public opinion formation, expression and manipulation in social network.

The rest of this paper is organized as follows: First, we introduce the sociological theory of spiral of silence, and modeling methods in Section 2. In Section 3, we investigate the influence of social bots and mass media on public opinion in networks through numerical simulation. Finally, discussions and the conclusions of this study are presented in Section 4.

2. Model description

2.1. The spiral of silence

The spiral of silence is a sociological theory proposed by Noelle-Neumann, which is based on crowd psychology and convergence behavior. In details, it states that individuals in a group who oppose the dominant opinions remain silent for fear of being isolated. As a result, the inferior opinions are silenced and the superior opinions are expressed and accessible to others; the dominant opinions expand and strengthen in a spiral manner, and eventually manipulate public opinion. As social medias become increasingly popular as online spaces for social interaction and political participation, the spiral of silence theory on social media has emerged as a popular subject of research, and many researchers[25–28] have attempted to establish mathematical models to explain the spiral of silence in networks.

2.2. Opinion expression under dual opinion climate

The formation of public opinion in social networks is not only determined by the interaction between individuals, but also influenced by mass media. In this paper, we attempt to establish a dual opinion climate model based on human users, social bots, and mass media, to capture the interaction between individuals (human or bots), and analyze the factors for formation of public opinion. Consider a society of N individuals, each of which has an opinion (or attitude) on a debatable topic, such as same-sex marriage. We assume that there are only two possible opinions on each topic: positive and negative. Therefore, as shown in Fig. 1, according to the spiral of silence theory, the opinion evolution at each time can be described as a three-step process:

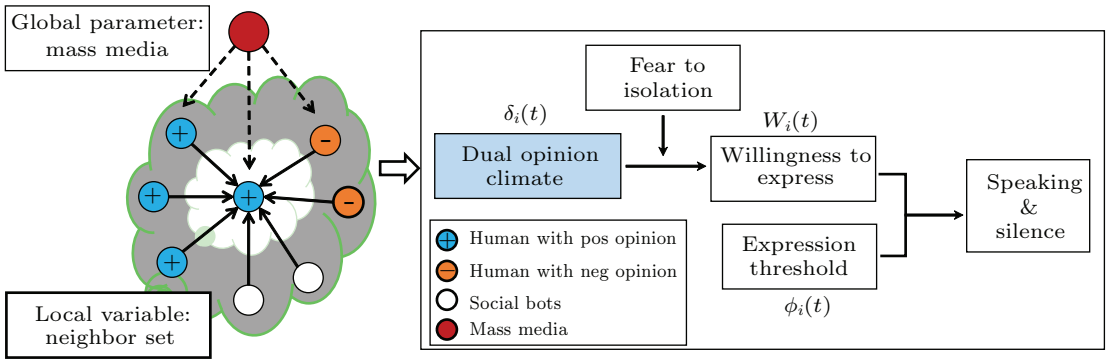


Fig. 1. Block diagram of opinion evolution under the dual opinion climate.

Firstly, the dual opinion climate of individual i comes from two sources: the local opinion climate, formed by sensory statistics of the opinions of neighbors (human or bots), and the global opinion climate, perceived by the (global variable) mass media as a reference.

Secondly, because of fear of being isolated, individual i's willingness to express his/her opinion is driven by the current climate.

Finally, the willingness of individual i to express an opinion, and his/her own expression threshold, jointly determine his/her choice of "speaking" or maintaining a "silence" state.

The dual opinion climate of individual i at time t in terms of the local opinion climate and global opinion climate is expressed by

$$\delta_i(t) = \begin{cases} \frac{n_s(t)\omega - n_o(t)}{n_s(t)+n_o(t)}, & \text{if} \quad n_s(t)+n_o(t) > 0, \\ 0, & \text{if} \quad n_s(t)+n_o(t) = 0, \end{cases} \tag{1}$$

where $n_s(t)$ and $n_o(t)$ represent the numbers of neighbors with supportive and opposing opinions, respectively, observed by individual i in a local social network at time t; ω denotes the weight reflecting the media influence perceived by individual i, according to the definition of Sohn[13] each individual is as-

sumed to have a susceptibility to mass media m, in the range [0,1]. The media influence weight ω can then be described by m as follows:

$$\omega = \begin{cases} 1+m_i, & \text{congruent media exposure,} \\ 1-m_i, & \text{incongruent media exposure.} \end{cases} \tag{2}$$

Under the congruent media exposure condition, in which the individual's opinion is the same as that of the mass media, the proportion of congruent neighbors increases according to their degree of media dependency. In contrast, under the incongruent media exposure condition, the individual's dependence on the congruent neighbors is weakened.

Further, to limit the value of $\delta_i(t)$ to the range $[-1, 1]$, we rewrite the impact of dual opinion climate $I_i(t)$ as an S-curve equation,

$$I_i(t) = 2(1+\mathrm{e}^{-\delta_i(t)})^{-1} - 1. \tag{3}$$

The willingness of individual i to express an opinion $W_i(t)$ can be defined as a linear combination of prior willingness $W_i(t-1)$ and the impact of the dual opinion climate $I_i(t)$ at the current time, as shown in Eq. (4). Accordingly, if $I_i(t) > 0$, indicating that individual i perceives a favorable double opinion climate, then his/her willingness to express $W_i(t)$ increases; on the contrary, $W_i(t)$ reduces if $I_i(t) < 0$; if $I_i(t) = 0$, $W_i(t)$ remains unchanged.

$$W_i(t) = \begin{cases} W_i(t-1) + (1 - W_i(t-1))I_i(t), & \text{if } I_i(t) \geq 0, \\ W_i(t-1) + W_i(t-1)I_i(t), & \text{if } I_i(t) < 0. \end{cases} \tag{4}$$

The intention to delimit $W_i(t)$ in the range [0,1] is mainly to match the individual's implicit expression threshold ϕ_i, which is constant and randomly distributed in [0,1]. Because the susceptibility to public opinion may vary between human users, we assume that each person has a hidden expression threshold ϕ_i. The expression threshold refers to the intensity of discourse in society: for example, the expression threshold of open (talkative) individuals may be low, whereas, for conservative (reticent) individuals, the parameter will tend to be high. Therefore, compared with a conservative person, an open person is more inclined to express an opinion when they have the same degree of willingness. Finally, individuals choose to speak out or remain silent in each step of the process, by comparing their willingness to express $W_i(t)$ with their threshold of expression ϕ_i.

If social bots are added to the network composed of human users, we expect to observe the bots' manipulation and interference with public opinion from the micro level. Unlike human users, goal-oriented social bots are not affected by the spiral of silence: they ignore external influences and continue to express their own opinions. At the same time, mass media, as a global variable, can also influence human users' willingness to express. In the following section, we use simulation to verify the phenomenon of public opinion manipulation by the joint action of social bots and mass media in small-world networks.

3. Simulation analysis and results

For simulation, we create a small-world network with $N = 1000$ agents and a random reconnection probability of $P_r = 0.5$ using the Netlogo programmable environment. In the initial state, the numbers of human users holding positive and negative opinions are the same, and social bots are randomly distributed in the network in the proportion λ (the bots ratio), which varies from 1% to 10%. Without loss of generality, we assume that all human users share the same media impact factor $m_i \equiv m$, which is varied from 0 to 1. For each set of parameters, the statistical average value of 1000 independent simulations is calculated and taken as the result.

3.1. Bots and media acting separately

Experiments are begun with the situation in which bots and media act separately. Because of the difference in connectivity density of different small-world networks, we simulate the distribution of public opinion over time in networks with different values of mean degree $\langle k \rangle$. As shown in Fig 2(a), when only bots are involved in the network ($\lambda \neq 0$, $m = 0$), as the number of bots increases, the proportion of humans holding the same opinion as the bots gradually increases,

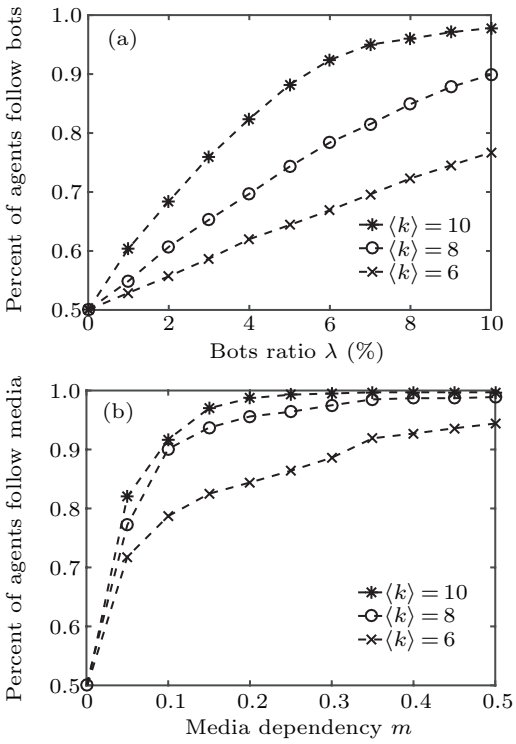


Fig. 2. Influence of different factors in small-world networks: (a) bots, (b) media.

and public opinion is gradually manipulated by the bots. As a result, the concerted action of even a small percentage of bots in the network is sufficient to disrupt the opinion climate among humans and tip public opinion in favor of the bots. It can be observed from Fig 2(b) that, when only media existed in the network ($\lambda = 0$, $m \neq 0$), as the media influence, the proportion of humans holding the same opinions as the media increased, and public opinion is guided by the media.

We find that both bots and media in the network can independently influence the opinion climate perceived by human users. The difference is that, as agents in the network, bots interfere with the opinion climate of their human neighbors through local interaction behaviors, and gradually spread their influence in the network through an information cascade effect. As a global variable, the media interfere with the opinion climate of all human users in the network directly and synchronously, to guide the public opinion of the network.

3.2. Bots and media with same opinion

We now investigate public opinion evolution in networks where bots and media act in concert, for example, when both hold positive opinions. Figure 3 shows comparative results of the effects of manipulating public opinion in three situations: (a) only bots exist, (b) only media exist, and (c) bots and media coexist. The parameters are set as follows: $N = 1000$, $\langle k \rangle = 6$, $\lambda = 5\%$, and $m = 0.05$. The simulation results show that the media and bots have a superimposed effect on opinion evolution and increase the probability of public opinion being manipulated. In different network structures, as P_r varied from 0 to 1, the combined effect of bots and media is significantly greater than the effect of bots or media alone.

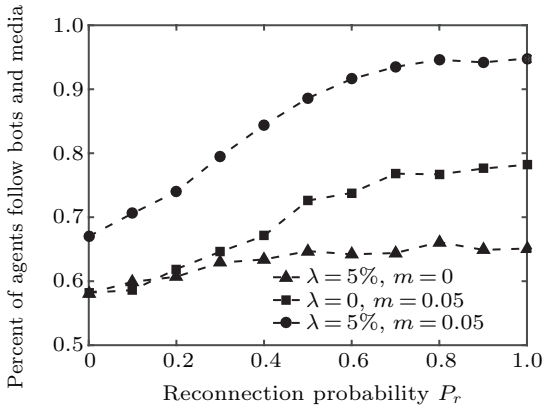


Fig. 3. Combined effects of bots and media acting in concert.

3.3. Bots and media with opposite opinions

Because bots can increase the risk of exposure to, and dissemination of, information with low credibility,[9] mass media take the role of clarifying rumors, to some extent, through their discourse power. Therefore, we now focus on the phenomenon of bots and media manipulating public opinion in networks together, when their opinions are polarized. The network size was $N = 1000$, of which the bots ratio is $\lambda = 5\%$, the opinions of human users are randomly selected from the binary values of positive and negative, bots are assigned negative opinions throughout, and the media holds positive opinions. Figure 4 shows how the proportion of human users with negative opinions changes with the media influence parameter m for networks with different average densities ($\langle k \rangle$).

We find that, when $m = 0$ (i.e., only bots exist), a 5% bots ratio is sufficient to manipulate public opinion. As m increases, the proportion of humans holding negative opinions decreases. This shows that, as the human user's dependence on media increases, the manipulation effect of bots on public opinion is weakened. Furthermore, the influence of media constantly increases, and finally the trend of public opinion is reversed, making the previous minority gradually become the majority of society.

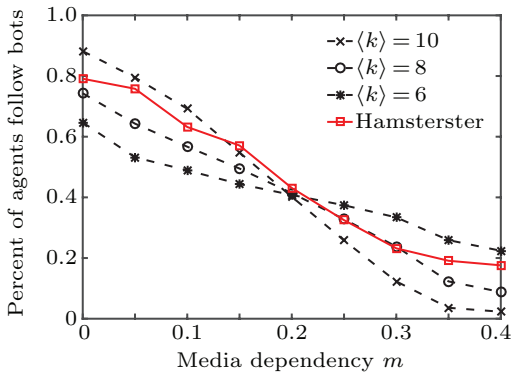


Fig. 4. Inhibiting effect of media on bots.

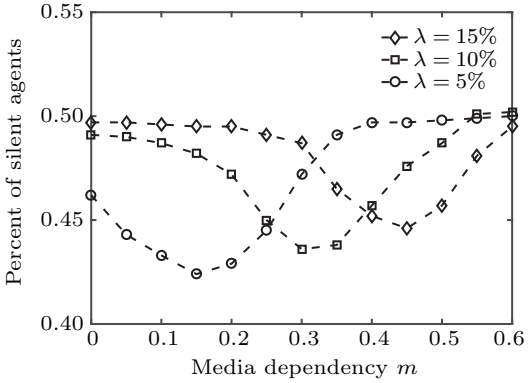


Fig.5. Proportion of silent humans with the influence of media.

In addition, we attempt to verify the evolution of public opinion in accordance with the spiral of silence in real small-world social networks, using the "Hamsterster friendships" dataset (a friendship network for 1858 users of hamsterster.com). As shown by the red line in Fig. 4, by comparison with artificial small-world networks, we find that the dual opinion climate model constructed by bots and media shows the same evolution trend in the real network.

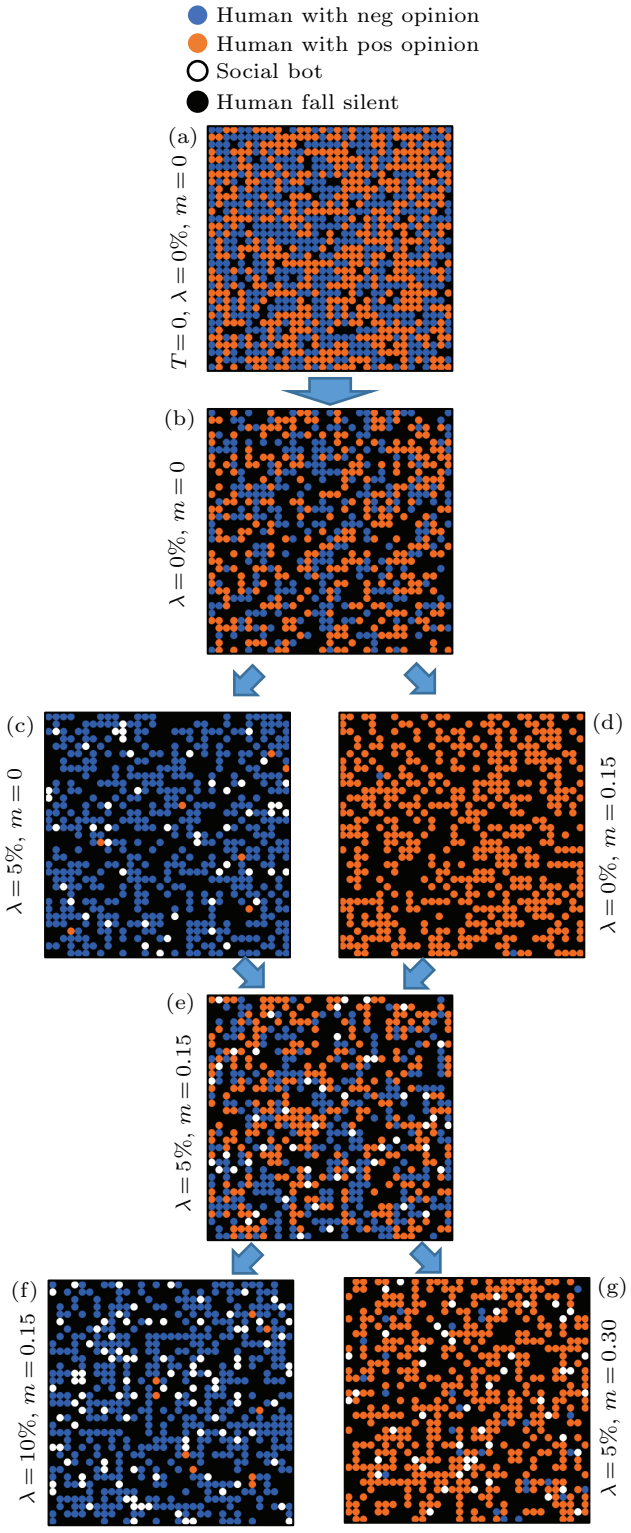


Fig. 6. Snapshot illustrating the evolution of the opinions of $N = 1000$ agents in a small-world network, when the media opinion is positive. Human users have positive opinion (orange) or negative opinion (blue), bots have negative opinion (white), and some human users fall silent (black).

We now investigate the change in the number of silent humans in the population in accordance with the spiral of silence theory, to explain the inhibitory effects of bots and media on the expression willingness of human users with opposite opinions. Figure 5 shows the percentage of silent humans, with respect to media influence m, with different bots ratios. The results show that, on the one hand, the participation of media can effectively weaken the bots' inhibition of humans with positive opinions, to effectively reduce the number of silent humans; on the other hand, as the influence of media increases, the media exerts a reverse inhibition on humans with negative opinions, so that the number of silent humans also increases in reverse. Therefore, the percentage of silent humans in the network is jointly affected by the bots ratio λ and the media influence m. As the effect of manipulating public opinion becomes stronger, the spiral of silence phenomenon becomes more obvious, and the number of silent human users increases.

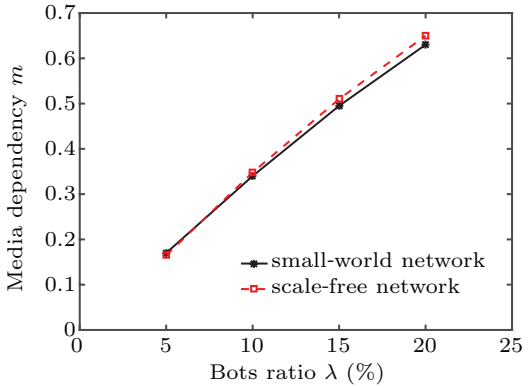


Fig. 7. Media dependency (m) as a function of bots ratio (λ) to ensure that public opinion remains stable.

The snapshots in Fig. 6 illustrate the opinion evolution in the dual opinion climate model, where the white nodes represent bots, and the orange and blue nodes represent humans with positive and negative opinions, respectively. When a human user becomes silent (because $W_i(t) < \phi_i$), the node turns to black. As shown in Fig. 6(a), in the initial state, 1000 agents are arranged in a 33×33 grid, with the bots and humans randomly distributed. In Fig. 6(b), $\lambda = 0$, $m = 0$ (with no bots or media), the numbers of humans with positive and negative opinion are approximately the same, the spiral of silence does not appear, and public opinion is not manipulated. In Fig. 6(c), $\lambda = 5\%$, $m = 0$ (only bots exist), the number of humans holding negative opinions increases significantly, the spiral of silence appears, and public opinion is manipulated by the bots. In Fig. 6(d), $\lambda = 0$, $m = 0.15$ (only media exists), the number of humans holding positive opinions increases significantly, the spiral of silence appears, and public opinion is manipulated by the media. In Fig. 6(e), $\lambda = 5\%$, $m = 0.15$ (both bots and media exist), the balance of the numbers of humans with positive and negative opinions is restored, and the effects of bots and media cancel each other out. Further, as shown in Figs. 6(f)–6(g), a continuous unilateral increase in either bots or media can upset the balance again and guide the evolution of public opinion in a certain direction.

Finally, we investigate the correlation between the media influence m and the bots ratio λ that is required to maintain a balance of public opinion (i.e., approximately the same number of humans with positive and negative opinions), in the case

that bots and media have opposite opinions. Figure 7 shows how m changes with λ in an artificial small-world network and the Hamsterster friendships network. Comparison of the two networks' simulation results shows that there is an approximately linear relationship between the media influence m and bots ratio λ.

4. Conclusions

In summary, we have introduced the spiral of silence theory of social psychology to attempt to explain the reasons for the formation of public opinion. In contrast to the classical opinion dynamics model, to explore the factors that change individuals' opinion values, we start with the change in individuals' states (speaking or silence) to analyze the distribution of opinions at the macro level of the network, to provide a new idea for research on public opinion. Further, we use a multi-agent modeling method to construct a dual opinion climate model with the joint participation of bots and media, to reveal the mechanism for manipulating public opinion.

Through experiments, we find that both bots and media can manipulate public opinion by promoting the expression willingness of humans who hold the same opinions and inhibiting those who hold opposite opinions, thus expanding the number of silent humans in the network, leading to the spiral of silence phenomenon, and ultimately manipulating public opinion. The difference between media and bots is that the latter exacerbates the spiral of silence from the perspective of local interaction in the network, by promoting or inhibiting a neighbor's expression willingness to change his/her state, which gradually spreads to other humans in the network. As a global variable, the media can exert influence on all humans in the network synchronously to achieve an effect of public opinion. When bots and media hold the same opinion, they produce a superimposed effect, further intensifying the manipulation of public opinion. However, when bots and media become polarized, the reverse action of the media can effectively offset the effect of bots on public opinion, and finally reverse the direction of the change in public opinion. This conclusion can provide a reference for simulating rumor propagation in online social networks.

However, there are several limitations on parameter selection to this work. For example, considering the feasibility of the comparative experiment of media and bots simulation, we adopt a simplified treatment, assuming that the media sensitivity of individuals is the same and varies between the interval [0,1]. Since media is a global variable and exerts influence on all individuals at the same time, we believe that this assumption can be considered from the macro level of the group, but the local influence on the network from the micro level is ignored. Therefore, we only discuss the distribution of public opinion. In future work, it can be interesting to further distinguish individuals with different expression threshold attributes. Similar studies include classifying individuals into open-minded, moderate-minded, and close-minded subgroups, and we will further investigate the impact of different attribute groups.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Nos. 61976120 and 62006128) and the Humanities and Social Science Fund of Ministry of Education of China (Grant No. 21YJCZH013).

References

[1] Chen X and De Bie T 2018 *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery and data mining*, August 19–23, 2018, London, UK, p. 1197
[2] Musco C and Tsourakakis C E 2018 *Proceedings of the 2018 World Wide Web conference*, April 23–27, 2018, Lyon, France, p. 369
[3] Das A, Gollapudi S and Munagala K 2014 *Proceedings of the 7th ACM international conference on web search and data mining*, February 24–28, 2014, New York, USA, p. 403
[4] Bakshy E, Messing S and Adamic L A 2015 *Science* **348** 1130
[5] Garimella K, De Francisci Morales G, Gionis A and Mathioudakis M 2018 *Proceedings of the 2018 World Wide Web conference*, April 23-27, 2018, Lyon, France, p. 913
[6] Matakos A, Terzi E and Tsaparas P 2017 *Data. Min. Knowl. Disc.* **31** 1480
[7] Chitra U and Musco C 2020 *Proceedings of the 13th International conference on web search and data mining*, February 3–7, 2020, Houston, USA, p.115
[8] Ferrara E, Varol O, Davis C, Menczer F and Flammini A 2016 *Commun. Acm.* **59** 96
[9] Shao C, Ciampaglia G L, Varol O, Yang K, Flammini A and Menczer F 2018 *Nat. Commun.* **9** 1
[10] Stella M, Ferrara E and De Domenico M. 2018 *Proc. Natl. Acad. Sci. USA* **115** 12435
[11] Luo Y, Cheng C, Li Y and Yu C. 2021 *Commun. Nonlinear Sci.* **98** 105772
[12] Ross B, Pilz L, Cabrera B, Brachten F, Neubaum G and Stieglitz S 2019 *Eur. J. Inform. Syst.* **28** 394
[13] Sohn D and Geidner N 2016 *Int. J. Public. Opin. Res.* **28** 1
[14] Anderson B D O and Ye M 2019 *Int. J. Autom. Comput.* **16** 129
[15] Wang L, Tian Y and Du J 2018 *Sci. China. Inform. Sci.* **48** 3 (in Chinese)
[16] Wang Y, Guo T, Li W and Chen B 2020 *Chin. Phys. B* **29** 100204
[17] Xia W, Ye M, Liu J, Cao M and Sun X 2020 *Automatica* **120** 109113
[18] Ye M, Zino L, Mlakar Z, Bolderdijk J W, Risselada H, Fennis B and Cao M 2021 *Nat. Commun.* **12** 1
[19] Hou J, Li W and Jiang M 2021 *Physica A* **574** 125968
[20] Li H, Bu Z, Wang Z and Cao J 2019 *IEEE Trans. Industrial Informatics* **16** 5327
[21] Zhang W, Zhang Z, Zeadally S and Leung V 2019 *IEEE Trans. Industrial Informatics* **15** 4216
[22] Li H, Wang Q, Liu S and Hu J 2020 *Physica A* **542** 123514
[23] Noelle-Neumann E 1993 *The spiral of silence: Public opinion, our social skin*, 1st edn (University of Chicago Press)
[24] Chengjun W 2012 *Proceedings of the 62nd annual conference of International Communication Association (ICA)*, May 24–28, 2012, Arizona, USA
[25] Chen H T 2018 *New. Media Soc.* **20** 3917
[26] Kennamer J D 1990 *Commun. Res.* **17** 393
[27] Sohn D and Geidner N 2016 *Int. J. Public Opin. Res.* **28** 25
[28] Cheng C, Luo Y and Yu C 2020 *Physica A* **551** 124163