Received May 29, 2019, accepted June 30, 2019, date of publication July 2, 2019, date of current version September 4, 2019.
Digital Object Identifier 10.1109/ACCESS.2019.2926439

Modeling and Analyzing the Influence of Multi-Information Coexistence on Attention

NAN ZHAO, SIWEI FANG, NAN CHEN, AND CHANGXING PEI
State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China
Corresponding author: Nan Zhao (zhaonan@xidian.edu.cn)

This work was supported in part by the Natural Science Foundation of Shanxi Province under Grant 2018JM6053, in part by the National Natural Science Foundation of China under Grant 61372076 and Grant 61301171, and in part by the 111 Project under Grant B08038.

ABSTRACT The research of information dissemination in online social networks has always been a hot topic in the field of network public opinion research. The traditional information dissemination model has a strong dependence on the complete network topology structure and the information of participants' neighbors. In real-network hot events, it is difficult to obtain the real-network topology structure, and most participants who participate in network events do not have social relationships. We propose a multiple information coexistence attention model (MICAM), which does not depend on the network topology and the information of participants' neighbors. According to the real news data, the corresponding mapping table of model parameters is established. The model is more in line with the actual situation of social networks due to parameter mapping tables. The MICAM model is adopted to study the changes of various types of information attention in the case of multi-information coexistence. The accuracy of the model is proved by comparing the simulation results with the actual data. In addition, the factors that affect the attention of participants are also studied.

INDEX TERMS Attention of information, multi-information dissemination, parameter mapping table, social network.

I. INTRODUCTION

In the modern society, online social networks (such as Facebook, Sina Weibo, etc.), news portals (such as Netease, Sina, etc.), mobile news client APP (such as Headline today APP, Netease news APP, etc.) provide great convenience for the hot issues dissemination. The news and information can be obtained instantly, and the amount of information accessed can increases dramatically for a period of time. People can show their views and opinions, and influence the spread of hot news in the network. Therefore, the behavior that people pay attention to different information has aroused the interest of many researchers. Research multi-information spreading and multi-information attention are conducive to further reveal the law of information dissemination among participants, but also help to formulate business promotion plans and communication strategies of media information.

The associate editor coordinating the review of this manuscript and approving it for publication was Hua-Wei Shen.

Information dissemination has been extensively researched. A large number of information dissemination models are applied to study the information dissemination process on the internet. In 1978, Mark Granovetter proposed a threshold model [8] by studying the phenomenon that individuals are influenced by other participants' behavior. In this model, the central node is continuously affected by those active from the neighbor nodes. Once the sum of the effects exceeds the activation threshold, the central node will be activated. Rahimkhani *et al.* [16] have deeply studied the threshold model. They combined the threshold model with the theory of maximizing influence to propose an improved algorithm that the most influential individuals can be located quickly. Then, independent cascades model [6] was originally introduced by Jacob Goldenberg in 2001. It is a typical probability model, where the nodes in the active state may affect their neighbors independently, and the probability that nodes interact with each other is also irrelevant. The researchers including C. Wang focus on the independent cascades model and apply it to the study of scalable influence maximization in large-scale social networks [19]. C. Castellano combines infectious disease model, which is proposed at the beginning of the 20th century, with information dissemination to analyze the process of information dissemination, the scope of influence and the law of action, etc. [2], [9], [11], [13]. They transplant the law of infectious diseases spreading among individuals into information transmission. The dynamics of information transmission is analyzed by simulating the process of infectious disease spreading, which is more intuitive and convenient to be understood. Some extended models [5], [7], [10], [17], [20] extensively arise in recent years. These related models have brought brand-new ideas to the research on the field of information dissemination. However, most of the models are based on static networks structure and features information from participants and their neighbors, the performance of the models is most probable to be limited by the actual network environment. In fact, it is difficult to gather timely accurate information about the network structures and participants in practice. In addition, the exchange of views and the information dissemination are not necessary to be based on the premise that the participants know or follow each other in reality.

Some of classical information dissemination models such as SIR model, independent cascades(IC) model, and linear threshold (LT) models concentrate on the dissemination of single information. Nevertheless, in most situation, a number of hot issues in online social networks emerge and disseminate at the same time. Therefore, the single information dissemination model cannot satisfy the actual demands frequently. In recent years, the influence of the propagation of multiple information and the coexistence of multi-information on individuals attention has gradually gained attention and been studied. P. Xu et al. study the competition of topics in social networks. The factors of information competition in social networks are considered and analyzed [21]. Zhang *et al.* and Myers and Leskovec [15], [23] do in-depth research on information cooperation and competition in social networks. In addition, cooperation in information dissemination is investigated. Zhao and Cui [24] combine personal interests with information dissemination to explore in more detail the characteristics in the social network. Sun and Yao [18] applies evolutionary game theory with some certain innovative significance to the process of social network information dissemination. Bhargava and Ghosh [1] predicts the popularity of information dissemination by analyzing the process of information dissemination in social networks.

The models above are constructed primarily according to the network topology and information attributes. To a certain extent, the overall trend of information dissemination in the whole network can be depicted, and the regularity or characteristics of information dissemination can be presented. However, the performance of the models and the change of some parameters rely on the network topology. It limits practicability of the model, because it is very difficult to obtain the complete topology of information dissemination in social network, and the timeliness is not easy to be guaranteed. It is obvious that, in the previous studies the contribution from the combination of attributes of participants and news is not considered particularly, while human as a significant component of information dissemination can directly impact on the results.

Our study is based on the attributes of attributes of human and news isolated from network topology. We are devoted to improving the practicability and accuracy of the model and propose a novel information dissemination model named as MICAM (Multiple Information Coexistence Attention Model). The particular attention is paid to the following issues.

(1) The spread situation of multiple information in the absence of specific network topology.

(2) The factors that affect the attention of participants to information and the degree of impact.

Specifically, we analyze the subjective decision-making process of human comprehensively in dominant social network. Meanwhile, we establish parameter mapping tables with real data to determine the value of subjective parameters in the model, contributing to enhance the applicability of the model.

The rest of the paper is presented as follows. In Section 2, we collected four real news s data to verify the accuracy of our model. In Section 3, we propose Multiple Information Coexistence Attention Model (MICAM) to analyze the change of participants' attention when there were many kinds of information in the network. In Section 4, we simulate the MICAM model, and compare the simulation results with real data to verify the accuracy of our model. In addition, we further simulate and analyze the factors affecting the attention of participants. Finally, we present our conclusions in Section 5.

II. DATA COLLECTION

We collected data about four different hot news, including "D&G humiliating China" incident, "Jiang Jinfu domestic violence" incident, "Golden Horse Award" incident, "National flag handing in Marathon" incident, through the "Micro Hot" big data application service platform. These data are from various platforms including online forums, Sina Weibo, Wechat, news website, etc. We measure the attention of news by detecting the amount of original information that consists of related keywords over a period of time.

(1) News 1: "D&G humiliating China" incident

On November 17, 2018, Dolce & Gabbana, an Italian luxury brand, released a short video promotional film suspected of racial discrimination against Chinese people, which aroused heated discussion among Chinese netizens. On November 21, one of the founders of Dolce & Gabbana publicly published extreme comments insulting Chinese people on social networks, which aroused strong dissatisfaction among Chinese netizens. On November 23, two founders of Dolce & Gabbana released apology videos.

TABLE 1. "D&G humiliating china" incident.

date	11.18	11.19	11.20	11.21	11.22	11.23	11.24	11.25	11.26	11.27	11.28	11.29	11.30	12.01	12.02
amount	6492	2259	1567	7713	24778	19761	14026	3190	2247	1547	1119	2247	24778	19761	14026

TABLE 2. "Jiang Jinfu domestic violence" incident.

date	11.20	11.21	11.22	11.23	11.24	11.25	11.26	11.27	11.28	11.29	11.30	12.01	12.02
amount	113434	31536	61396	49002	33705	30735	1632	65582	22354	45931	36379	24001	30683

On November 29, Dolce & Gabbana withdrew apology videos from its social accounts and appealed to Italians to boycott Chinese products.

(2) News 2: "Jiang Jinfu domestic violence" incident

On November 20, 2018, Jiang Jinfu, a well-known Chinese artist, admitted to having domestic violence against his girlfriend in his social account, he apologized and confessed to the public; On the morning of November 22, Jiang Jinfu's friend disclosed in his social account that Jiang Jinfu's girlfriend had chaotic private life, false pregnancy, and that her family was related to Japanese gangsters. On the morning of November 27, Japanese media reported that Jiang Jinfu was suspected of domestic violence and said that "blood was everywhere in his family". On November 28, Hu Ge, a well-known Chinese artist and friend of Jiang Jinfu, made comments on his social account, encouraging Jiang to wait for the truth.

(3) News 3: "Golden Horse Award" incident

On the evening of November 17, 2018, "Golden Horse Award" Film Award Ceremony was held in Taiwan Province of China. A Taiwanese female director made a speech of "Taiwan independence" when she came on stage to receive the award, Chinese actors and directors boycotted it collectively; On the evening of November 19, the media reported that Gong Li, a famous Chinese actress, refused to come on stage to award in "Golden Horse Award" ceremony; On November 26, participants disclosed that Chinese actress Xu Qing left in anger during the award presentation.

(4) News 4: "National flag handing in Marathon" incident

In the Suzhou International Marathon on November 17, 2018, Chinese women athlete were handed the national flag by volunteers near the end of the race, which affected their performance and eventually led to the loss of the champion. On November 25, the situation that athlete were handed the national flag by volunteers appeared again in the Shaoxing International Marathon.

The statistical charts of the four types of news s are as follows,

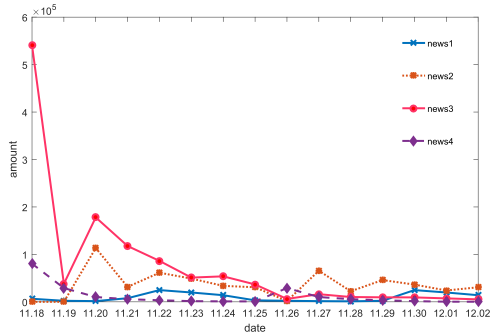


FIGURE 1. Daily amount of four types of news messages with keywords.

III. MULTI-INFORMATION COEXISTENCE ATTENTION MODEL

A. RELEVANT ASSUMPTIONS OF MODEL

(1)When the news are attractive to participant i, the participant may participate in hot events by forwarding information or comment information. We use $1 \times m$ matrix $F_i(t)$ to denote the attractiveness of m type information to participants i in the current network environment, that is, the participant's attention to m type information. $F_i(t)$ is defined as follows,

$$F_i(t) = \{f_1 \quad f_2 \ldots f_m\}, \tag{1}$$

where f_k represents the attractiveness of information k $(k = 1, 2, \ldots M)$ to participant i. After forwarding information k, participant i can forward the information k again if it satisfies the attention value greater than the attention threshold θ, as follows,

$$f_k(t) > \theta. \tag{2}$$

If the subsequent update information of news k appears, the premise for participant i to forward the update information is that the attraction of updated information to participant i must be greater than that of the last forwarding information, as follows,

$$f_k(t) < f_k(t + \Delta t). \tag{3}$$

The formula (3) means the attraction of the updateds to participant i must be greater than that of information recently released.

The state of attention of participant i, we adopt state matrix $Si(t)$ to represent.

$$S_i(t) = \{s_1 s_2 \ldots s_m\}, \tag{4}$$

$$s_k = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} \quad (k = 1, 2, \ldots, m). \tag{5}$$

In formula (5)(6), 0 indicates that participant i is not concerned about information k, 1 indicates that participant i is

TABLE 3. "Golden horse award" incident.

date	11.18	11.19	11.20	11.21	11.22	11.23	11.24	11.25	11.26	11.27	11.28	11.29	11.30	12.01	12.02
amount	540835	38038	178008	117657	86312	51119	53774	36709	6171	16165	10216	9569	9394	7302	5259

TABLE 4. "National flag handing in marathon" incident.

date	11.18	11.19	11.20	11.21	11.22	11.23	11.24	11.25	11.26	11.27	11.28	11.29	11.30	12.01	12.02
amount	80457	29063	10176	5602	3392	1535	1043	950	28221	9883	4672	3059	1366	319	212

concerned about information k. For the participants, they can pay attention to multiple information in unit time.

$$\begin{cases} f_k(t) \leq \theta \quad or\, f_k(t) \geq f_k(t+\Delta t) \Rightarrow s_k = 0 \\ f_k(t) \geq \theta \quad or\, f_k(t) \leq f_k(t+\Delta t) \Rightarrow s_k = 1. \end{cases} \tag{6}$$

B. PARAMETER OF MODEL

In the attention function $F(t)$, we focus on several factors for detailed analysis, and build a multi-information coexistence attention model (MICAM) according to the results of analysis.

1) INITIAL OCCURRENCE TIME OF INFORMATION

The initial occurrence time of information refers to the initial time when the hot news was first disclosed or reported by network media, which is expressed by the parameter $\tau = \{\tau_0\ \tau_1\ \ldots\ \tau_m\ \}$. According to the article [12], the relationship between participant's attention and initial occurrence time of information τ can be modeled by an exponential function $T(\tau)$, which is expressed as follows,

$$F_i(t) \propto T_{1\times m}(\tau), \tag{7}$$

where

$$T_k(\tau_k) = e^{\tau_k - t} \quad (\tau_k \leq t_{\leq} T_{end}) \tag{8}$$

In formula (7), m denotes the number of different types of information in the current network. In formula (8), τ_k denotes the initial occurrence time of news k, and T_{end} denotes the time when the heat of information dissemination basically disappears. In this paper, For initial news information k, the subsequent updates news of k is represented by parameter $\tau^k_{(r)}$. $\tau^k_{(r)}$ represents the r-th update news of k.

2) CONTENT MATCHING DEGREE

In social networks, participants usually pay attention to the hot events that they are interested in. Moreover, participants with different attributes may focus different types of news that participants are interested in. We adopt $C_i(\psi, \zeta)$ to express the matching degree between information content and participant's interest, as follows,

$$F_i(t) \propto C_i(\psi, \zeta), \tag{9}$$

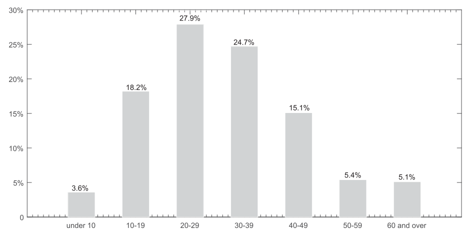


FIGURE 2. The age Y structure of Chinese participants in 2018 from CNNIC statistical report on internet development in china.

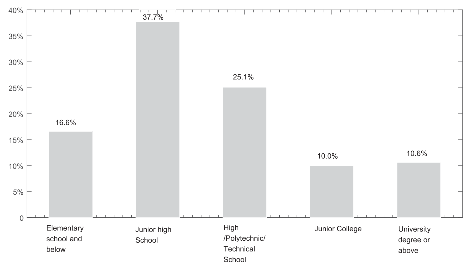


FIGURE 3. The education background E structure of Chinese participants in 2018 from CNNIC statistical report on internet development in china.

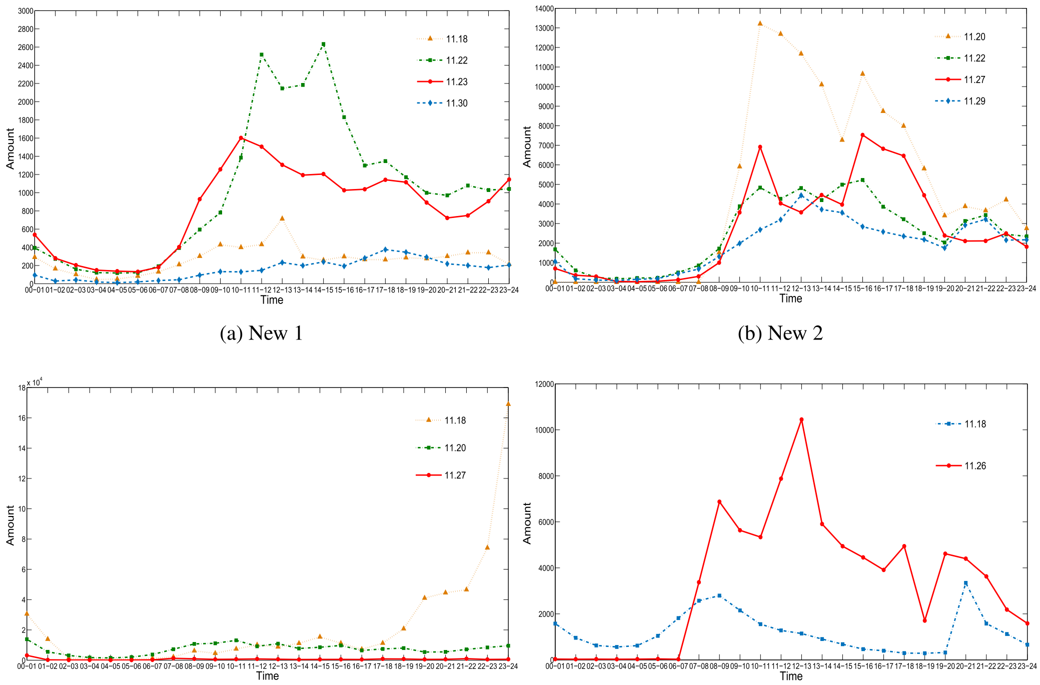
where ψ represents the participant's personal attribute matrix in the current network, as shown in formula(10).

$$\psi = \begin{pmatrix} \varphi_{1,1}(E,Y) & \varphi_{1,2}(E,Y) & \ldots & \varphi_{1,N}(E,Y) \\ \varphi_{2,1}(E,Y) & \varphi_{2,2}(E,Y) & \ldots & \varphi_{2,N}(E,Y) \\ \ldots & \ldots & \ldots & \ldots \\ \varphi_{N,1}(E,Y) & \varphi_{N,2}(E,Y) & \ldots & \varphi_{N,N}(E,Y) \end{pmatrix}. \tag{10}$$

In formula (10), E denotes the educational level of participants, Y denotes the age of participants, and N denotes the number and scale of participants in the current network. According to [4], in the model, E and Y follow the distribution trend of the figure2 and figure3 respectively. ζ represents the information attribute matrix in the network, as follows,

$$\zeta = \{\xi_0(tp,l)\xi_1(tp,l)\ldots\xi_m(tp,l)\}. \tag{11}$$

In formula (11), ξ denotes the content function of information, and tp denotes the content type of information, including political type, social and livelihood type, international diplomacy type, military type, etc. l denotes the content sensitivity



(a) New 1

(b) New 2

(c) New 3

(d) New 4

FIGURE 4. The statistical figures of 24-hour information release.

of corresponding information. m denotes the number of different types of information in the current network. For the content sensitivity of information, we define as follows,

$$\frac{l_\upsilon}{l_\kappa} = \frac{\sum\limits_{\Delta t = p_\upsilon - 1}^{P_\upsilon + 1} I(\Delta t)}{\sum\limits_{\Delta t = p_\kappa - 1}^{P_\kappa + 1} I(\Delta t)}, \tag{12}$$

$$P \in \Delta_{day} = \arg\max\{I(\Delta_{day})\}, \tag{13}$$

where $I(\Delta t)$ represents the number of information release within $t - th$ hours. $I(\Delta_{day})$ indicates the number of information release within the $day - th$ day, and P represents the time period (1h) when the number of information release reaches peak. When news and subsequent events occur, we use content sensitivity of information to measure the attraction level of news to participants. For the above four news and subsequent news, we give tables to record the number of information release in the day or next day when the news happened. We define $l_{(r)}^k$ is the content sensitivity of the $r - th$ update information of news k. $l_{(0)}^k$ represents the content sensitivity of the original news. We expect to find out the relationship between content sensitivity of information and corresponding news through historical real data, so that we can establish relational table of news-sensitivity. Then we can find the corresponding values of content sensitivity for other news, which will facilitate the quantification of content sensitivity. In this paper, the content sensitivity of all news s is determined according to the real data statistics figure 4(a)(b)(c)(d). In the figure 4, abscissa represents a period of time, and ordinate represents the amount of news released during that period.

TABLE 5. The age structure of news participants.

Age	Under 25	25-34	35-44	Over 44
Percentage	37.9%	26.0%	22.8%	13.3%

(1) news 1

According to the statistics of figure 4(a), we get

$$l_{(0)}^1 : l_{(1)}^1 : l_{(2)}^1 : l_{(3)}^1 \approx 1.44 : 6.63 : 4.36 : 1. \tag{14}$$

(2) news 2

According to the statistics of figure 4(b), we get

$$l_{(0)}^2 : l_{(1)}^2 : l_{(2)}^2 : l_{(3)}^2 \approx 2.80 : 1.24 : 1.61 : 1. \tag{15}$$

(3) news 3

According to the statistics of figure 4(c), we get

$$l_{(0)}^3 : l_{(1)}^3 : l_{(2)}^3 \approx 71.33 : 8.24 : 1. \tag{16}$$

(4) news 4

According to the statistics of figure 4(d), we get

$$l_{(0)}^4 : l_{(1)}^4 \approx 3.45 : 1. \tag{17}$$

According to figure 4(a)(b)(c)(d), we get the formula (18).

$$\begin{cases} l_{(0)}^1 : l_{(0)}^2 : l_{(0)}^3 : l_{(0)}^4 \approx 1 : 22.08 : 201.16 : 16.83 \\ l_{(1)}^1 : l_{(1)}^2 : l_{(1)}^3 : l_{(1)}^4 \approx 1 : 2.12 : 5.03 : 1.13 \\ l_{(3)}^1 : l_{(3)}^2 : l_{(3)}^3 \approx 1.07 : 4.51 : 1 \\ l_{(4)}^1 : l_{(4)}^2 \approx 1 : 11.38. \end{cases} \tag{18}$$

TABLE 6. The education background structure of news participants.

Education background	Elementary School and below	Junior High School	HighPolytechnicTechnical School	Junior College	University degree or above
Percentage	3.7%	13.2%	23.7%	23.5%	35.9%

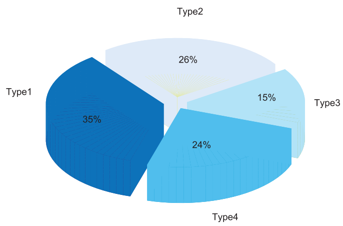


FIGURE 5. Distribution of network news participants .

TABLE 7. The age structure of type 1 participants.

Age	Under 25	25-34
Percentage	84.9%	15.1%

For news k, the $C_i(\psi, \zeta)$ of participant i is defined as follows,

$$C_i(\psi_i, \zeta_k) = \mu_k \cdot Map(Y_i, E_i \rightarrow tp_k) \cdot e^{\lg l_{(0)}^k} \cdot \varepsilon(t - \tau_k), \tag{19}$$

where μ_k indicates the adjustment coefficient. Taking into account the difference between the type of standard news and the type of actual news k, we use the adjustment coefficient to balance the difference between them. $\varepsilon(t - \tau_k)$ is the step function of time t.

According to [3], we give statistics on the age structure and educational background of internet news participants in China, and divide the internet news participants into four types. The statistics of these four kinds of internet participants' attributes and their concerns are respectively given by *table*5 - *table*18.

Similarly, we expect to find out the relationship between content matching degree and corresponding news through historical real data, so that we can establish mapping tables of news content matching degree. That will facilitate the quantification of content matching degree. In this paper, the content matching degree of all news s is determined according to *Table* 19.

One point needs special explanation. Figures 2 and 3 are statistical charts of the age and educational level of all Chinese netizens. The basis of parameters value of Y and E in the MICAM are statistical distributions in figures 2 and 3. Tables 5 and 6 are statistical tables of age and education level of netizens who are concerned about news. Tables 5 and 6 are proposed just for the purpose of establishing Table 19 and are not related to Y and E in MICAM.

In order to express convenience, we set

A1 = Under 25; A2 = 25-34; A3 = 35-44; A4 = Over 44. B1 = Elementary School and below; B2 = Junior; B3 = High/Polytechnic/Technical School; B4 = Junior College; B5 = University degree or above.

According to $table5 - table18$, we give the mapping table, we get the final mapping table $Map(Y_i, E_i \rightarrow tp)$ as shown $table19$.

3) SCALE OF PARTICIPANTS

We define the number of Internet participants who focus on hot news in the network as $N_m(t)$. Obviously, the degree of attention of participants for hot news is related to the number of participants that concern about the news in the current network.

$$N_{1\times m}(t) = \{n_1(t)n_2(t)..n_m(t) \quad (n_k(t) < N^2, k = 1, 2 \ldots m) \tag{20}$$

4) UPDATE INFORMATION

Relevant subsequent information of network hot news is often disclosed after a certain period of time. It may be supplement to the initial information, or may be very different from the initial information. We call the subsequent information as "updated information" $\gamma_{1\times m} = \{\gamma_1(t)\ \gamma_2(t) \ldots \gamma_m(t)\}$, there is the following formula,

$$F_{\rm i} \propto \sum_{t=t_{oc}+1} \gamma_{1\times m}(t), \tag{21}$$

where t_{oc} indicates the time of updating information occurred. Considering the lag time of information dissemination [14], [22], we set the time of updating information occurred is t_{oc+1}.

In addition,

$$\begin{aligned}\gamma_k(t) &= C_i(\psi, |\eta - \eta_\theta|\,\zeta) \\ &= |\eta - \eta_\theta|\, e^{\lg l_{(r)}^k} \cdot \mu_k \cdot Map(Y_i, E_i \rightarrow tp_k),\end{aligned} \tag{22}$$

for η, $\eta \in (-1, 0) \cup (0, 1)$, η indicates the positive/negative correlation between the update information and the initial information. Taking into account the difference between the updated information content and the source information content, the greater the difference is, it often attracts more people's attention. We set the content difference threshold value η_θ be used to adjusted.

5) PROBABILITY OF INFORMATION ACQUISITION

There is a relative lag in the acquisition of information for network participants. Normally, hot news need some time to be known to the public. One of important function of social media platform is the latest news push (e.g. the hot search list of Sina Weibo, etc.), and the main basis is the scale of participants. The probability that a participant gets news k at

TABLE 8. The education background structure of type 1 participants.

Education background	Elementary School and below	Junior High School	High/Polytechnic/Technical School	Junior College	University degree or above
Percentage	4.4%	13.7%	26.4%	27.5%	28.0%

TABLE 9. The news types concerned structure of type 1 participants.

news types	Social news	Entertainment News	Life News	current political news	Technology News	Financial news	Sports news
Percentage	74.4%	62.9%	56.2%	42.5%	39.8%	30.5%	30.1%

t time is $P_k(t)$, it is defined as follows,

$$P_k(t) = \frac{\sum_{t=\tau_k}^{t} n_k(t)}{\sum_{j=1}^{m}\sum_{t=\tau_k}^{t} n_j(t)}. \tag{23}$$

In particular, if participant i knows news k at t_0 time, when $t > t_0$, for participant i, $P_k(t) = 1$.

C. DECISION-STATE MATRIX

On the basis of 3.1 and 3.2, we give the participant decision function and the decision matrix of the model based on the decision function. Firstly, we give the participant decision vector $(F_i(t))_{1\times m}$ as follows,

$$\begin{aligned}(F_i(t))_{1\times m} = ((P_k(t))_{1\times m}\cdot u_i)\cdot[(\mu_k\cdot C_i(\psi,\zeta)_{1\times m}\\ +\sum_{\tau_{k_t}}\varepsilon(t-\tau_{k_t})\cdot\gamma_{1\times m}(t))\cdot T_{1\times m}(t)].\end{aligned} \tag{24}$$

where u_i denotes participant i,

$$P_k(t)\cdot u_i = \begin{cases}0\\1\end{cases}$$
$$if\ P_k(t_0)\cdot u_i = 1 \rightarrow P_k(t)\cdot u_i = 1(t \geq t_0), \tag{25}$$

where $P_k(t)\cdot u_i = 1$ means that participant i knows news k at time t, otherwise he does not know.

The formula (24) is composed of two parts, attention of initial information and attention of updated information. For the part of attention of initial information, we focus on netizens' attention when the news happens, and this is related to the attributes of the initial news and netizens. For the part of attention of updated information, we study how news updates affect the attention of netizens, and this is related to the attributes of the news updates and netizens.The formula reflects the changing process of netizens' attention to multi-information under the influence of various factors.

Next, we give the decision matrix of MICAM.

$$F(t) = \begin{pmatrix} F_{1,1}(t) & F_{1,2}(t) & \dots & F_{1,N}(t)\\ F_{2,1}(t) & F_{2,2}(t) & \dots & F_{2,N}(t)\\ \dots & \dots & \dots & \dots\\ F_{N,1}(t) & F_{N,2}(t) & \dots & F_{N,N}(t)\end{pmatrix}. \tag{26}$$

According to the decision matrix and $F_i(t)- > S_i(t)$ mapping relation in 3.1, we can get the state matrix of the MICAM finally, as follows,

TABLE 10. The age structure of type 2 participants.

Age	Under 25	25-34	35-44	Over 44
Percentage	84.9%	60.3%	8.0%	0.9%

$$S(t) = \begin{pmatrix} S_{1,1}(t) & S_{1,2}(t) & \dots & S_{1,N}(t)\\ S_{2,1}(t) & S_{2,2}(t) & \dots & S_{2,N}(t)\\ \dots & \dots & \dots & \dots\\ S_{N,1}(t) & S_{N,2}(t) & \dots & S_{N,N}(t)\end{pmatrix}. \tag{27}$$

So far, we have completed the construction of MICAM. We hope to apply MICAM to specific network cases. Therefore, we briefly describe the application process of MICAM.

In the practical application of the model, we only need to collect and analyze the data of information existing in the current network. Meanwhile, we can sample some of the netizens in the current network, and take the attribute distribution of these netizens as the attribute distribution parameters of all netizens. Finally, the information and the attributes of netizens are brought into the model to predict and analyze the change of multi-information attention in the current network environment.

IV. MODEL GENERATION ALGORITHM

Based on the content of the MICAM presented above, we present the generation algorithm of the model to prepare for subsequent simulation experiments and results analysis. The generation algorithm as follows Algorithm 1.

The core of the algorithm is to show the construction process of the model in the form of concise pseudo-code. However, it is difficult to describe the whole process of model building in detail here and some parameters in the MICAM are not included in the algorithm, because there are many parameters involved in the model and the defining expressions of some parameters are complex. What we show in the algorithm is the important steps and parameters of the construction of the MICAM, and that enable readers to have an overall understanding of the formation process of the model. At the same time, for the next research contents, it provides a basis for simulation experiments and result analysis of the model.

TABLE 11. The education background structure of type 2 participants.

Education background	Elementary School and below	Junior High School	High/Polytechnic/Technical School	Junior College	University degree or above
Percentage	1.2%	6.6%	14.1%	21.5%	56.6%

TABLE 12. The news types concerned structure of type 2 participants.

news types	Social news	Entertainment News	Life News	current political news	Technology News	Financial news	Sports news
Percentage	76.5%	62.2%	56.2%	59.0%	45.5%	42.6%	39.5%

Algorithm 1 MICAM Generation Algorithm

Input: $T_k(\tau_k)$, $C_i(\psi, \zeta)$, N, $\sum_{t=t_{oc}+1} \gamma_{1\times m}(t)$, μ_k, θ

Output: $\{N_{1\times m}\}_{t\times 1}$

initialize attribute parameters of participant i;
while $Stop == 0$ **do**
  for $t = 1; ;t++$ **do**
    for each $i \in N^2$ **do**
      Get $F_i(1)_{1\times m}$;
      if $f_k(1) > \theta$ **then**
        $s_k(1) = 1$;
      else
        $s_k(1) = 0$;
      end if
    end for
    Get $S_i(1) = \{s_k(1)\}_{1\times m}$;
    for $t = 1$ to T_{end} **do**
      for each $i \in N$ **do**
        Get $(F_i(t))_{1\times m}$;
        if Update information of news k appears **then**
          if $f_k(t) > f_k(t - \Delta t)$ **then**
            $s_i(t) = 1$;
          else
            $s_i(t) = 0$
          end if
        end if
        if Update information of news k did not appears **then**
          if $f_k(t) > \theta$ **then**
            $s_i(t) = 1$;
          else
            $s_i(t) = 0$;
          end if
        end if
        $S_i(t) = \{s_k(t)\}_{1\times m}$;
      end for
    end for
  end for
end while
return $\{N_{1\times m}\}_{t\times 1}$;

A. ALGORITHMIC DESCRIPTION

(1) Information and data collection. Relevant attribute data of target news information are collected, including the initial

TABLE 13. The age structure of type 3 participants.

Age	25-34	35-44	Over 44
Percentage	19.0%	76.8%	4.2%

TABLE 16. The age structure of type 4 participants.

Age	25-34	35-44	Over 44
Percentage	9.4%	38.1%	52.5%

occurrence time of information, the occurrence time of subsequent update information, news type and information content. Relevant attribute data of participants were collected, including the size, age distribution and education level distribution of participants.

(2) Model parameter establishment and assignment. Six types of model parameters (the occurrence time of information, the content matching degree, the update information, the scale of participants and the probability of information acquisition) are established, and the model parameters are assigned directly or indirectly through the collected data and the model parameter mapping table.

(3) Comprehensive decision-making mechanism. The decision function is adopted to construct the individual's comprehensive decision mechanism, and whether the individual pays attention to the information depends on the judgment result of the comprehensive decision mechanism.

(4) Individual − multi-information attention mapping. The one-to-many mapping relationship of attention between individuals and news is obtained by iterating and looping through the individual attention-state matrix.

(5) Group − multi-information attention mapping. The many-to-many relationship mapping of attention between groups and news information is obtained by iterating and looping through netizens in the current network, until the external termination condition is satisfied to end the cycle.

V. EMPIRICAL ANALYSIS

In order to verify the accuracy of the model, we compare the simulation results of the model with the actual statistical figure of attention of news.

A. MODEL SIMULATION

We categorize news 1 as social news, news 2 as entertainment news, news 3 and 4 as current political news. For the

TABLE 14. The education background structure of type 3 participants.

Education background	Elementary School and below	Junior High School	High/Polytechnic/Technical School	Junior College	University degree or above
Percentage	1.0%	4.6%	9.5%	30.8%	54.1%

TABLE 15. The news types concerned structure of type 3 participants.

news types	Social news	Entertainment News	Life News	current political news	Technology News	Financial news	Sports news
Percentage	76.4%	55.0%	55.0%	60.6%	39.2%	43.8%	31.7%

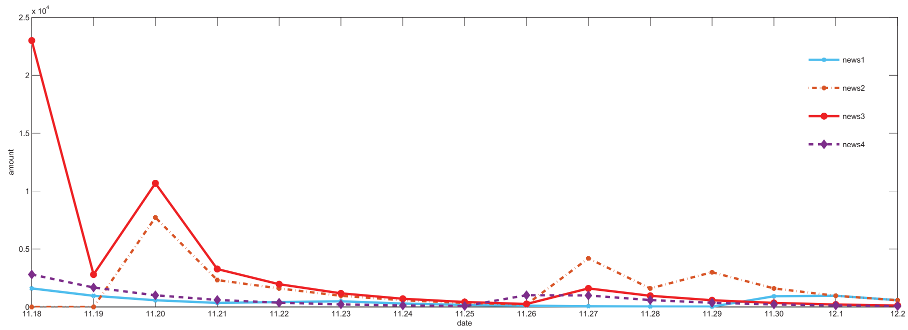


FIGURE 6. Simulation of daily amount of four types of news messages with keywords.

positive/negative correlation of content η, we set the threshold $\eta_\theta = 0.2$. For the positive/negative correlation of content η of updating news, we adopt the matrix as follows,

$$(\eta^k_{(r)})_{4\times 3} = \begin{bmatrix} \eta^1_{(1)} & \eta^1_{(2)} & \eta^1_{(3)} \\ \eta^2_{(1)} & \eta^2_{(2)} & \eta^2_{(3)} \\ \eta^3_{(1)} & \eta^3_{(2)} & \\ \eta^4_{(1)} & & \end{bmatrix}. \quad (28)$$

Here, $\eta^k_{(r)}$ denotes the positive/negative correlation of the content of the $r-th$ update information of news k. There are some differences between the four types of news mentioned above and the corresponding types of standard news. We adopt the adjustment coefficient μ_k to adjust. For the occurrence time of initial information $\tau^k_{(0)}$ of news k, We set the matrix to express, as follows,

$$(\tau^k_{(0)})_{1\times 4} = \begin{bmatrix} \tau^1_{(0)} & \tau^2_{(0)} & \tau^3_{(0)} & \tau^4_{(0)} \end{bmatrix}. \quad (29)$$

The occurrence time of updating news $\tau^k_{(r)}$, represented by a matrix as follows,

$$(\tau^k_{(r)})_{4\times 3} = \begin{bmatrix} \tau^1_{(1)} & \tau^1_{(2)} & \tau^1_{(3)} \\ \tau^2_{(1)} & \tau^2_{(2)} & \tau^2_{(3)} \\ \tau^3_{(1)} & \tau^3_{(2)} & \\ \tau^4_{(1)} & & \end{bmatrix}. \quad (30)$$

Through the description above, we construct the MICAM model. Then, we will simulate the changing process of attention of four types of news (news 1,2,3,4) in dissemination process. The simulation result is shown in figure 6.

B. MODEL COMPARISON AND VALIDATION

After getting the simulation results, we compare data of real news events with the simulation results and analyze the error rate in order to verify the relative accuracy of the model. Firstly, for four types of news, the proportion of the number of information release each day to the total number of information release on that day is calculated (retaining the last three decimal points) as following figure 7 (a)(b)(c)(d).

According to figure 7(a)(b)(c)(d), the error of proportion of the real data and the simulation data corresponding to the news $k(k \in \{1, 2, 3, 4\})$ on day $t(t \in [1, 15])$are calculated (the results are kept in three decimal places), as shown in *table*20.

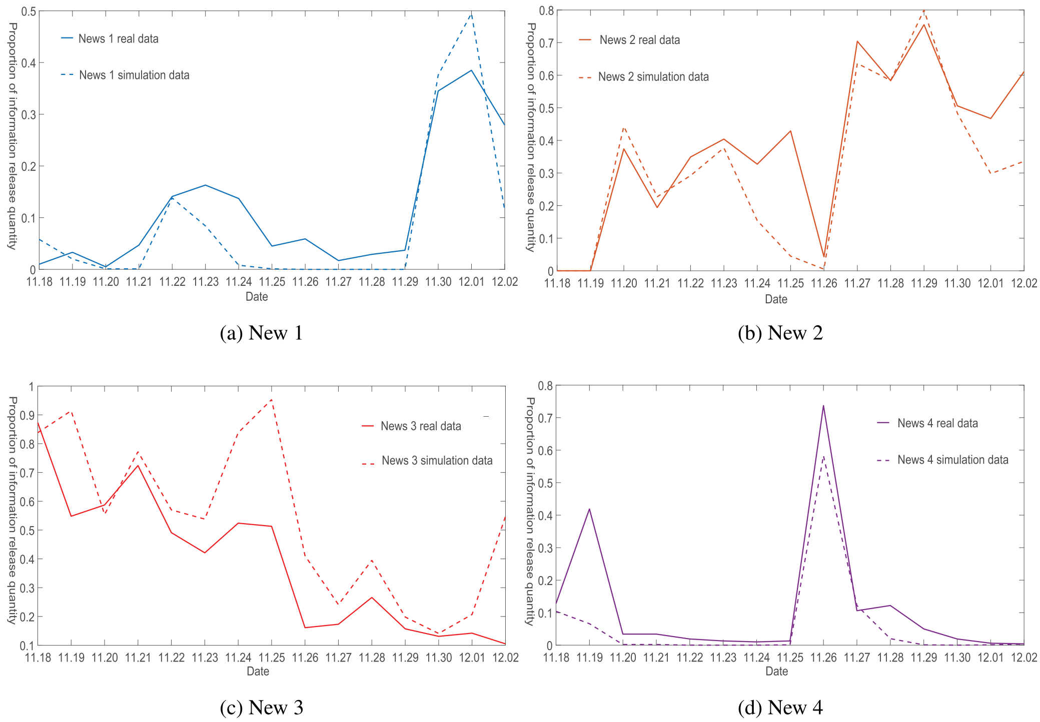
According to the statistical results of error statistics table, the average errors of all news s on t day $\Delta_{mean\_er}(t)$ are calculated. The expressions are as follows,

$$\Delta_{mean\_er}(t) = \frac{1}{4}\sum_{k=1}^{4}\Delta_{\mathrm{er\_k}}(t) \quad (31)$$

Finally, the comprehensive average error can be obtained by taking average $\Delta_{mean\_er}(t)$ $(t \in [1, 15])$

$$\begin{aligned} \Delta_{mean\_er} &= \frac{1}{15}\sum_{t=1}^{15}\Delta_{mean\_er}(t) \\ &= \frac{1}{15}\times\frac{1}{4}\sum_{t=1}^{15}\sum_{k=1}^{4}\Delta_{er\_k} \\ &\approx 0.093 = 9.3\%. \end{aligned} \quad (32)$$

By comparing the results of model simulation and actual data, the comprehensive average error is 9.3%. It shows that the



(a) New 1 (b) New 2 (c) New 3 (d) New 4

FIGURE 7. The proportion of the number of information on the day to the total number of information on the day.

TABLE 17. The education background structure of type 4 participants.

Education background	Elementary School and below	Junior High School	High/Polytechnic/Technical School	Junior College	University degree or above
Percentage	6.9%	25.3%	39.8%	15.2%	12.8%

TABLE 18. The news types concerned structure of type 4 participants.

news types	Social news	Entertainment News	Life News	current political news	Technology News	Financial news	Sports news
Percentage	71.3%	36.2%	50.3%	64.2%	38.9%	40.1%	30.0%

general trend of model simulation is similar to the actual trend, and the error rate is still acceptable. However, there are some small errors in local details. When there are multiple types of news information in social networks, the model can effectively analyze the changing trend of network participants' attention to news events to a certain extent.

For the occurrence of local slight deviation, we believe that there are the following reasons.

a: It is unable to judge the number of participants in the actual network for real news. Therefore, there are some errors in scale of participants that we adopt in simulation, resulting in some errors between the results of the simulation and the actual change trend. This belongs to the systematic error.

b: There are a lot of information in the follow up updates of news. We can only find the important and influential updates as model parameters to simulate. There may be other undiscovered follow-up updates which lead to deviation compared with the actual situation.

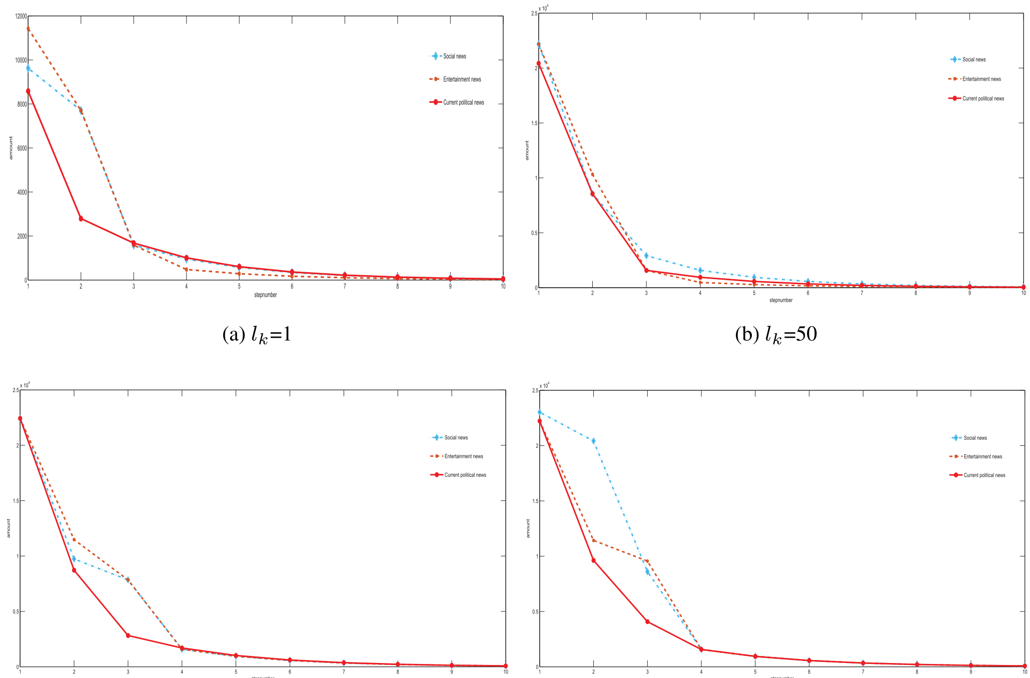
c: The network environment is complex and changeable, and there are many unpredictable environmental variables (including the factors of social network platform itself, the quantitative deviation of content sensitivity, the adjustment deviation of news content type, etc.) which will affect the attention of news, and belong to systematic error.

C. ANALYZING THE INFLUENCING FACTORS OF MICAM

After studying the different attributes of news, we analyze the effects of various factors on the attention of news. In this part, we adopt standard news types to carry out simulation experiments.

1) THE INFLUENCES OF TYPE OF NEWS ON PARTICIPANTS' ATTENTION

We analyze the influence of social news, entertainment news and current political news on participants' attention. In order to control the variables, we set the content sensitivity l_k of three types of news to be the same in each simulation,



(a) l_k=1 (b) l_k=50

(c) l_k=100 (d) l_k=150

FIGURE 8. The influence of types of news on participants' attention with different content sensitivity (l_k).

TABLE 19. $Map(Y_i, E_i \rightarrow tp)$ mapping table of four types participants.

Mapping	Social news	Entertainment News	current political news
(A1,B1)	0.17	0.14	0.10
(A1,B2)	2.07	1.73	1.29
(A1,B3)	7.73	6.41	4.86
(A1,B4)	9.06	7.55	5.85
(A1,B5)	22.44	18.52	15.53
(A2,B1)	0.04	0.03	0.03
(A2,B2)	0.63	0.47	0.47
(A2,B3)	2.20	1.63	1.58
(A2,B4)	2.80	2.21	2.05
(A2,B5)	7.44	6.91	6.52
(A3,B1)	0.05	0.02	0.04
(A3,B2)	0.65	0.60	0.57
(A3,B3)	1.97	1.14	1.71
(A3,B4)	2.17	1.46	1.77
(A3,B5)	5.32	3.75	4.27
(A4,B1)	0.03	0.02	0.03
(A4,B2)	0.41	0.21	0.36
(A4,B3)	1.15	5.74	1.03
(A4,B4)	0.49	0.26	0.43
(A4,B5)	0.72	0.41	0.63

the initial occurrence time $\tau_{(0)}^k$ of information is the same as 1, three types of news are standard types, and the adjustment coefficient μ_k is set to 1, there is no follow-up news update. The simulation results are as follows,

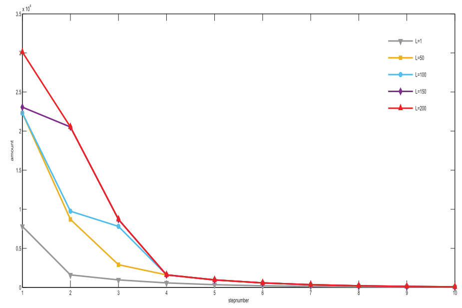
Analysis of simulation results;

From the results of Figure 8 (a) (b) (c) (d), we can draw conclusions,

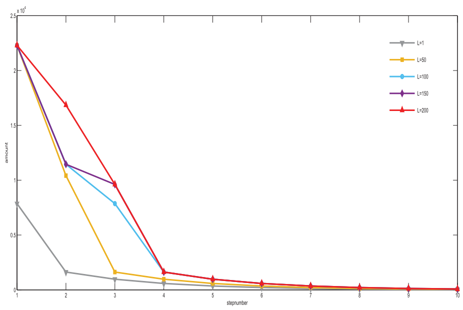
TABLE 20. The error of the proportion of the real news data and the simulation news data.

date	11.18	11.19	11.20	11.21	11.22
error of news 1	0.048	0.013	0.004	0.046	0.208
error of news 2	0.000	0.000	0.069	0.033	0.057
error of news 3	0.035	0.366	0.033	0.048	0.079
error of news 4	0.024	0.353	0.032	0.032	0.019
date	11.23	11.24	11.25	11.26	11.27
error of news 1	0.079	0.129	0.044	0.059	0.017
error of news 2	0.027	0.173	0.384	0.038	0.068
error of news 3	0.117	0.314	0.440	0.250	0.068
error of news 4	0.013	0.010	0.012	0.154	0.016
date	11.28	11.29	11.30	12.01	12.02
error of news 1	0.037	0.031	0.110	0.165	0.029
error of news 2	0.045	0.023	0.169	0.275	0.001
error of news 3	0.041	0.010	0.065	0.442	0.078
error of news 4	0.049	0.019	0.005	0.001	0.102

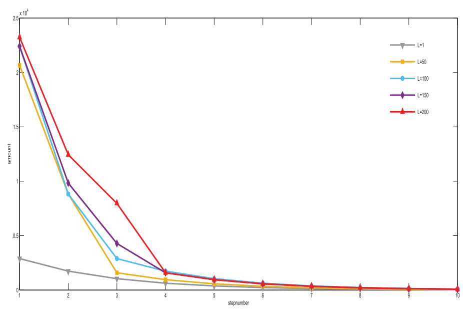
(1) Each simulation has the same sensitivity to the content of three types of news, when the content sensitivity of the three types of news is high, there is no obvious difference in their initial attention. The result indicates that the participants have strong interest in highly sensitive news, and the type of news has little influence on the attention of high sensitivity news. But for those news events with low sensitivity, the participants have much higher interest in entertainment news than current political news and social news.



(a) Social news



(b) Entertainment news



(c) Current political news

FIGURE 9. The influence of content sensitivity on the attention of news.

(2) With the increase of content sensitivity, participants pay more attention to social news and current political news, especially for social news.

2) THE INFLUENCE OF CONTENT SENSITIVITY ON PARTICIPANTS' ATTENTION

We set different content sensitivities of the same kind of news events, and study the influence of content sensitivities on attention of news. The simulations are as follows,

Analysis of simulations results;

Analyzing Figure 9 (a) (b) (c), we can draw the following conclusions,

(1) The impact of content sensitivity on participants' attention is briefly summed up as follows. The higher the content sensitivity is, the higher the participants' attention is. The higher the content sensitivity is, the slower the decline of participants' attention is.

(2) Participants may be very interested in entertainment news and current political news with relatively high content sensitivity. When content sensitivity is higher than 50, the entertainment news and the current political news receive similar attention from participants. For social news, from the simulations results, the significant variation trend of participants' attention is presented with the different sensitivities of social news content.

(3) When the content sensitivity is relatively low, the attentions of the three types of news events decline rapidly, indicating that the duration of non-sensitive events is very short. When the content sensitivity is relatively high, the attentions of the three types of news events decline much slower than those with low content sensitivity. The attentions of social news and current political news decline faster than entertainment news, and the pace of decline is the fastest for current political news, indicating that highly sensitive entertainment news is the most concerned by participants, followed by social news, and finally current political news.

VI. CONCLUSION

We propose an information dissemination model named as MICAM to analyze the change of information attention in the case of multi-information coexistence. The model is more practical because it is not dependent on the network topology and the information from neighbor nodes. Compared with the traditional single information dissemination model, the MICAM model has more generality and extensiveness in the study of multi-information dissemination. According to the real issues data, we formulate a corresponding model parameter mapping table, which makes the model more suitable to the actual situation of information dissemination in social networks and improves the accuracy of the model. The comparison between the simulation results and the actual data proves that the MICAM model can accurately describe the changes in participants' attention to various types of information in the case of multi-information coexistence. Further, we analyze and explain the reasons for the existing error in the application of the model in detail. In addition, we make an analysis of two critical factors affecting participants' attention, and draws corresponding conclusions.

REFERENCES

[1] S. Bhargava and S. Ghosh, "A study on online news popularity prediction techniques," *Int. J. Online Sci.*, vol. 4, no. 9, p. 11, Sep. 2018.

[2] C. Castellano and R. Pastor-Satorras, "Competing activation mechanisms in epidemics on networks," *Sci. Rep.*, vol. 2, Apr. 2012, Art. no. 371.

[3] *China Internet News Market Research Report 2016*, CNNIC, Beijing, China, 2017.

[4] *The 42th China Statistical Report on Internet Development*, CNNIC, Beijing, China, 2018.

[5] J. P. Gleeson, K. P. O'Sullivan, R. A. Baños, and Y. Moreno, "Effects of network structure, competition and memory time on social spreading phenomena," *Phys. Rev. X*, vol. 6, no. 2, Apr./Jun. 2016, Art. no. 021019.

[6] J. Goldenberg, B. Libai, and E. Müller, "Talk of the network: A complex systems look at the underlying process of word-of-mouth," *Marketing Lett.*, vol. 12, no. 3, pp. 211–223, 2001.

[7] C. Granell and S. Gómez, and A. Arenas, "Competing spreading processes on multiplex networks: Awareness and epidemics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 1, 2014, Art. no. 012808.

[8] M. Granovetter, "Threshold models of collective behavior," *Amer. J. Sociol.*, vol. 83, no. 6, pp. 1420–1443, May 1978.

[9] H. W. Hethcote, "The mathematics of infectious diseases," *SIAM Rev.*, vol. 42, no. 4, pp. 599–653, 2000.

[10] S. Hosseini and M. A. Azgomi, "A model for malware propagation in scale-free networks based on rumor spreading process," *Comput. Netw.*, vol. 108, pp. 97–107, Oct. 2016.

[11] J.-Q. Kan and H.-F. Zhang, "Effects of awareness diffusion and self-initiated awareness behavior on epidemic spreading—An approach based on multiplex networks," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 44, pp. 193–203, Mar. 2017.

[12] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jun. 2009, pp. 497–506.

[13] J. Leskovec, M. McGlohon, C. Faloutsos, N. Glance, and M. Hurst, "Patterns of cascading behavior in large blog graphs," in *Proc. SIAM Int. Conf. Data Mining*, 2007, pp. 551–556.

[14] E. Mussumeci and F. C. Coelho, "Reconstructing news spread networks and studying its dynamics," *Social Netw. Anal. Mining*, vol. 8, no. 1, p. 6, Dec. 2018.

[15] S. A. Myers and J. Leskovec, "Clash of the contagions: Cooperation and competition in information diffusion," in *Proc. IEEE 12th Int. Conf. Data Mining*, Dec. 2012, pp. 539–548.

[16] K. Rahimkhani, A. Aleahmad, M. Rahgozar, and A. Moeini, "A fast algorithm for finding most influential people based on the linear threshold model," *Expert Syst. Appl.*, vol. 42, no. 3, pp. 1353–1361, Feb. 2015.

[17] F. Siwei, Z. Nan, C. Nan, X. Fei, and Y. Yunhui, "Analyzing and predicting network public opinion evolution based on group persuasion force of populism," *Phys. A, Stat. Mech. Appl.*, vol. 525C, pp. 809–824, Jul. 2019.

[18] Q. Sun and Z. Yao, "Evolutionary game analysis of competitive information dissemination on social networks: An agent-based computational approach," *Math. Problems Eng.*, vol. 2015, Dec. 2015, Art. no. 679726.

[19] C. Wang, W. Chen, and Y. Wang, "Scalable influence maximization for independent cascade model in large-scale social networks," *Data Mining Knowl. Discovery*, vol. 25, no. 3, pp. 545–576, Nov. 2012.

[20] X. Wei, S. Chen, X. Wu, J. Feng, and J.-A. Lu, "A unified framework of interplay between two spreading processes in multiplex networks," *EPL*, vol. 114, no. 2, 2016, Art. no. 26006.

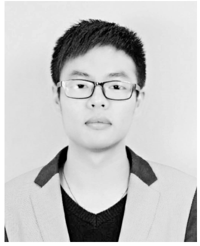
[21] P. Xu, Y. Wu, E. Wei, T.-Q. Peng, S. Liu, J. J. Zhu, and H. Qu, "Visual analysis of topic competition on social media," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 12, pp. 2012–2021, Dec. 2013.

[22] L. Zhang, S. Fu, and B. Li, "Research on stock price forecast based on news sentiment analysis—A case study of alibaba," in *Proc. Int. Conf. Comput. Sci.* Cham, Switzerland: Springer, 2018, pp. 429–442.

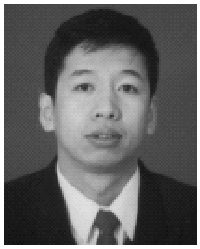
[23] Y. Zhang, C. Tang, and L. Weigang, "Cooperative and competitive dynamics model for information propagation in online social networks," *J. Appl. Math.*, vol. 2014, no. 2, Jul. 2014, Art. no. 610382.

[24] N. Zhao and X. Cui, "Impact of individual interest shift on information dissemination in modular networks," *Phys. A, Stat. Mech. Appl.*, vol. 466, pp. 232–242, Jan. 2017.

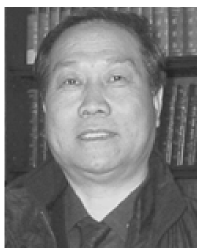
NAN ZHAO was born in Xi'an, Shaanxi, China, in 1980. He received the Ph.D. degree in telecommunication and information system from Xidian University, Xi'an, in 2014, where he is currently an Associate Professor and a Master Student Supervisor with the College of Telecommunication Engineering. His research interests include social media data mining and information processing.



SIWEI FANG was born in Huangshan, Anhui, China, in 1993. He is currently pursuing the master's degree with the College of Telecommunication Engineering, Xidian University. His research interests include network public opinion and information dissemination.



NAN CHEN was born in Xi'an, Shaanxi, China, in 1965. He received the M.S. degree in telecommunication and electronic systems from Xidian University, Xi'an, in 1993, where he is currently a Professor, the Director of the Department of Electronic Technology, and a Master Student Supervisor with the College of Telecommunication Engineering. He is also the Deputy Director of the National Basic Course Electrotechnics and Electronics Teaching Base. His research interests include public opinion analysis in social networks and information dissemination.



CHANGXING PEI is currently a Professor and an Advisor for doctoral students. He holds two patents and applied for two. He has authored 11 technical books and has published over 100 papers in key journals and academic conferences. He is mainly involved in the fields of radio communication, anti-jamming communication, network measurement, television systems, high-frequency electronic circuits, and modern filter networks. He is a Senior Member of the China Institute of Communication, the China Institute of Electronics, the Vice-Director of the Electronics Institute of Shaanxi Province, and the Director of the Working Committee for Academic Exchange. He received the Prize of Scientific and Technical Accomplishment three times.