User Behavior Prediction of Social Hotspots Based on Multimessage Interaction and Neural Network

Yunpeng Xiao, Jinghua Li, Yangfu Zhu, and Qian Li

***Abstract*—In network public-opinion analysis, the diversity of messages under social hot topics plays an important role in user participation behavior. Considering the interactions among multiple messages and the complex user behaviors, this article proposes a prediction model of user participation behavior during multiple messaging of hot social topics. First, considering the influence of multimessage interaction on user participation behavior, a multimessage interaction influence-driving mechanism was proposed to predict user participation behavior more accurately. Second, in the view of the behavioral complexity of users engaging in multimessage hotspots and the simple structure of backpropagation (BP) neural networks (which can map complex nonlinear relationships), this study proposes a user participant behavior prediction model of social hotspots based on a multimessage interaction-driving mechanism and the BP neural network. Finally, the multimessage interaction has an iterative guiding effect on user behavior, which easily causes overfitting of the BP neural network. To avoid this problem, the traditional BP neural network is optimized by a simulated annealing algorithm to further improve the prediction accuracy. In evaluation experiments, the model not only predicted the user participation behavior in actual situations of multimessage interaction but also further quantified the correlations among multiple messages on hot topics.**

***Index Terms*—Backpropagation (BP) neural network, multimessage interaction, social hotspots, user behavior.**

I. Introduction

WITH the emerging of the Internet era, online social networks such as Twitter and Facebook continue to be popular. People's communication and lifestyle have brought about tremendous changes. The generation and dissemination of hot topics in social media are constantly affecting the daily lives of people. The social hotspots refer to news or topics that are concerned or interested by the public at present. The social network topology and the user's reads and replies to messages in the network promote the dissemination and evolution of information related to the hot topic, that is, the propagation of the network topics [1]. Therefore, mastering user-forwarding participation behavior is important for evaluating the influence of a microblog topic [2], monitoring public opinion through networks [3], [4], and information retrieval [5].

Manuscript received June 10, 2019; revised November 8, 2019 and January 8, 2020; accepted January 20, 2020. Date of publication February 17, 2020; date of current version April 3, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61772098, in part by Chongqing Graduate Education Teaching Reform Project under Grant yjg183081, in part by the Science and Technology Research Program of Chongqing Municipal Education Commission under Grant KJZD-K201900603 and Grant KJQN201900629, and in part by Chongqing Graduate Scientific Research and Innovation Project under Grant cys19263. *(Corresponding author: Yunpeng Xiao.)*

The authors are with the College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China (e-mail: xiaoyp@cqupt.edu.cn; 1581144794@qq.com; 617411376@qq.com; liqian@cqupt.edu.cn).

Digital Object Identifier 10.1109/TCSS.2020.2969484

At present, the prediction of user behavior in social networks mainly includes the following two approaches. The first approach analyzes the structural topology map used for information dissemination in social networks. This approach predicts the path and range of the information propagation [6] and, hence, the user's participation behavior. Which users will participate in the microblog is commonly predicted by dynamic propagation [7] or an infectious disease model [8], [9]. Such predictive models typically classify network nodes as unknowns, communicators, and immunizers [10], [11]. However, this modeling has two main shortcomings. First, it creates a complex topology diagram requiring a large number of calculations. Second, it considers only the relationships of interest among the users, ignoring the differences among users and the frequent changes of topics in social networks. The second modeling approach considers user activity, the number of fans, and the number of messages [3], [12], [13]. Some scholars also make predictions based on the user's microblog interest and microblog information [14]–[16]. The influence of social media platforms (such as Weibo) and the behavior of users are then predicted by machine learning.

The forwarding behavior of online social networks has been extensively studied in recent years. Focusing on the different aspects of the predicted content, prediction models using both approaches have been established. However, despite significant progress in this area of research, there are still some challenges.

1) *The Complexity of the Multimessage Interaction:* Most studies predict either the microparticipation behavior during single messaging or the macropopularity perception during multimessage topics. These studies ignore the complexity of interactions among multiple messages under hot topics that occur in actual situations.
2) *The Ambiguity of Multimessage Mutual Impact Metrics:* The user participation behavior is closely related to the multimessage interaction under a topic. Traditional microparticipation behavior mostly starts from a single message, generally, only analyzes user attributes or network topology, and does not accurately measure the interaction of multiple messages.
3) *The Accuracy of the Predicted Model:* Traditional models cannot correctly capture the nonlinear relationship

between the topic data input and user behavior prediction output. In addition, ordinary neural networks are usually overfitting and prone to local minimums, thus reducing the accuracy of predictions.

When predicting user participation behavior, the model should consider the personal characteristics of users. In addition, the interactions among multiple messages under the same hot topic are vital for improving the prediction results. Multimessage interaction mechanisms and nonlinear relationships can be handled by a backpropagation (BP) neural network model. The BP neural network is a multilayer feedforward network trained by an inverse error propagation algorithm. It can learn and store a large number of input–output mode mapping relationships, without the need to derive mathematical equations for the relationship in advance. However, as mentioned earlier, multiple messages exert an iterative guiding effect on user behavior, which causes overfitting of the neural network. To avoid the overfitting problem, this article applies a simulated annealing algorithm to the BP neural network, which assists the local miniaturization solution of the algorithm and greatly improves the accuracy of the prediction results. The simulated annealing algorithm is derived from the principle of solid annealing and has greatly improved the prediction results in many past instances [17].

The main innovation of this article is that we study the user behavior of social hotspots from the perspective of multimessage interaction at the microlevel. The specific contributions of this article are as follows.

1) A user participation behavior prediction model based on multimessage interaction is constructed. Based on the mapping relationships between the basic user information and participation behavior under the traditional single message, the multimessage interaction-driving mechanism improves the completeness of the prediction results. Meanwhile, it is more realistic to describe the process of message dissemination.
2) A quantization mechanism based on multimessage interaction is proposed. This article can more accurately measure the multimessage selection process within the user community by quantitatively evaluating the mutual influence of messages from the perspective of topics. Meanwhile, the hidden influence under the same topic can be qualitatively measured, which leads to user's participation behavior.
3) The BP neural network was improved by the simulated annealing algorithm. This method fits well with the nonlinear relationship between the topic data input and the user behavior prediction output. Moreover, the neural network overfitting problem is solved by the simulated annealing algorithm, and the prediction accuracy is further improved.

This article is organized as follows. This section introduces the background and status of the research. Section II discusses the work related to our study, and Section III formalizes the research question. Section IV describes the proposed method and its learning algorithms. Section V experimentally evaluates our method on a real-world data set, and Section VI concludes this study.

II. Related Work

Considering the complex user participation behaviors would greatly improve the realism of the predicted behaviors in social network hotspots. Significant progress has been made in this research area. This section analyzes and discusses the prediction of user participation behavior at a single message by traditional models, especially those based on neural networks, in recent years.

In most current models, prediction of user participation behaviors takes into account the user network topology and user basic information while ignoring the impact of messages propagated under hot topics. Sheikhahmadi *et al.* [18] proposed a two-level model that detects and classifies the influence of users by considering the interaction between users. Similarly, Colombo *et al.* [19] established a topological map for studying information dissemination through a social network. Salehi *et al.* [20] extracted the attributes of a multilayer network structure for predicting the probability of microblog forwarding. Other researchers [21], [22] predicted user forwarding behavior through related attributes using a machine-learning method. Grabowicz *et al.* [23] predicted the user forwarding behavior by filtering the factors that are strongly related to user behaviors.

Most of the existing studies predict the nonlinear relationships between the topic data input and the user participation behavior output by traditional machine-learning methods. Lee *et al.* [24] predicted the user forwarding behavior and the time of forwarding by different machine-learning algorithms. Sankaram *et al.* [25] constructed an impact model based on a machine-learning algorithm and predicted the behavior of users and fans. Huang *et al.* [26] measured the user interest in different categories of tweets by a Bayesian model and predicted the forwarding behavior from the interest metrics. Other studies have simulated user participation in messages using infectious disease models, which cannot properly represent the nonlinearity. Huang and Su [27] analyzed the occurrence probability of user behaviors and predicted the user forwarding behavior in a susceptible-infectious-recovered (SIR) model of disease dynamics. Xiong *et al.* [28] proposed a new susceptible, contacted, infected and refractory (SCIR) model (where C denotes "Contacted") that distinguishes and predicts the user browsing behavior and forwarding behavior in detail.

In recent research, user participation behavior has been predicted by neural networks. Although the neural network can fit complex nonlinear relationships between the topic data input and the user participation behavior output, most of the neural network studies are directed at single messages. After considering the driving mechanism of multimessage interactions, the accuracy and complexity of the model would be improved. For example, Yang [29] predicted the online download behavior of users using artificial neural networks. Li *et al.* [30] improved the traditional BP neural network by a genetic algorithm and detected the frequent changes in the social network. Similarly, Liu *et al.* [31] optimized a radial basis function (RBF)-based neural network by cloud theory (a concept in fuzzy mathematics) and predicted the user participation behaviors of single messages. Sharma and Minocha [32] applied neural networks to link prediction in social networks.

The behaviors of user forwarding single messages have also been studied in improved versions of traditional neural networks [33]–[36]. In particular, Zhang *et al.* [35] proposed a convolutional neural network model based on the attention mechanism for retweets prediction of Twitter. This model has good performance.

III. Problem Definition

A. Related Definitions

The present problem considers the basic information $N = \{(n_j, m_i)|m_i \in M\}$ of multiple messages $M = \{m_1, m_2, \ldots, m_n\}$. In the above equation, m_i represents a message under the same hot topic. The basic information $N = \{(n_j, m_i)|m_i \in M\}$ represents the n_j basic features of message m_i, which includes the time of publication, the source and blogger of the message, and the average influence of the message. After extracting the interaction factors of the two messages, the user's personal characteristics are combined to predict that user's participation behavior. The basic definition is given as follows.

Definition 1: The user participation network is defined as $G_U^{m_i} = \{(U, m_i)|U = \{\omega_1, \omega_2, \ldots, \omega_n\}, m_i \in M\}$.

By definition, $G_U^{m_i}$ is the user participation network, which is composed of an user group U that has participated in message m_i related to hot topics. Here, U is a collection of participating ω_i, and m_i is a message under the hot topic.

Definition 2: The user history behavior is defined as $A = \{(a_i, b_i, \omega_i, t)|t \in \psi, \omega_i \in U\}$.

Here, A represents a collection of historical user behaviors, specifically the original and forwarding microblog behavior of user ω_i in the network, who has participated in a hot topic during time period t, and a_i and b_i are the number of original and forwarding microblogs of user ω_i during t, respectively. The time $t \in \psi$ represents the data collection time (one month).

Definition 3: The user tags are denoted as $L = \{(c_i, \omega_i)| \omega_i \in U\}$.

Here, L is the collection of user tags, specifically, the tags of all users who participated in message forwarding in a hot topic, and c_i is the tag of user ω_i participating in the user network.

Definition 4: The influence of multiple messages is expressed as $Influence(m_i, m_j) = isDifT(m_i, m_j) + isSameS(m_i, m_j) + isSameB(m_i, m_j)$

where $Influence(m_i, m_j)$ denotes the influence of message m_i on message m_j, and the $isDifT(m_i, m_j)$ function compares the publication time of messages m_i and m_j. When the value of $isDifT(m_i, m_j)$ is 1, −1, or 0, then message m_j is published after m_i, before m_i, or, simultaneously, with m_i, respectively. The function $isSameS(m_i, m_j)$ indicates whether messages m_i and m_j have the same forwarding source. The function $isSameB(m_i, m_j)$ indicates whether messages m_i and m_j were sent by the same blogger. If m_i and m_j were sent by the same blogger, the function value is 1, otherwise 0. For example, suppose that message A influences message B, message B is published after message A, and both messages were forwarded from the same source by the same blogger. The influence of message A on message B is then quantified as 3.

Definition 5: The multimessage correlation indicator is computed as

$$J(m_i, m_j) = \frac{\sum_{j=1}^{\text{count}(\text{user}(m_i))} \text{output}_{m_j}(\text{user}(m_i))}{\text{count}(\text{user}(m_i))} * \text{Influence}(m_i, m_j), \quad m_i, m_j \in M$$

where $J(m_i, m_j)$ represents the correlation between message m_i and message m_j, count(user(m_i)) represents the number of users participating in message m_i, and $\sum_{j=1}^{\text{count}(\text{user}(m_i))} \text{output}_{m_j}(\text{user}(m_i))$ is the sum of the values after the BP neural network model has participated in message m_i and will participate in message m_j. The function $Influence(m_i, m_j)$ represents the influences of message m_i on message m_j, as defined earlier.

B. Problem Formulation

To formulate the target problem, we first construct the user participation network $G_U^{m_i} = (U, m_i)$ of a message related to a hot topic and the personal attribute (internal driving mechanism) of the user based on the user's historical behaviors $A = \{(a_i, b_i, \omega_i, t)|t \in \psi, \omega_i \in U\}$ and user tags $L = \{(c_i, \omega_i)|\omega_i \in U\}$. Second, the basic information $N = \{(n_j, m_i)|m_i \in M\}$ of the multimessage is statistically processed to obtain the driving mechanism of the multimessage interaction. By comprehensively considering the two driving mechanisms, we can predict whether the user will participate in other messages under the same topic. The mutual influence strength between the multiple messages is then quantified by the multimessage correlation index. More specifically, the problem is defined as follows:

$$\left.\begin{array}{c} M = \{m_1, m_2, \ldots, m_n\} \\ N = \{(n_j, m_i)|m_i \in M\} \\ G_U^{m_i} = (U, m_i) \\ A = \{(a_i, b_i, \omega_i, t)|t \in \psi, \omega_i \in U\} \\ L = \{(c_i, \omega_i)|\omega_i \in U\} \end{array}\right\} \Rightarrow \begin{array}{c} G_U^{m_j}(m_j \in M) \\ J(m_i, m_j)(m_j \neq m_i). \end{array}$$

1) Input: The proposed model requires the following inputs (the definitions are given earlier):

1) multiple messages $M = \{m_1, m_2, \ldots, m_n\}$ under hot topics;
2) basic information $N = \{(n_j, m_i)|m_i \in M\}$ of multiple messages;
3) the user participation network $G_U^{m_i}=(U, m_i)$;
4) the user history behaviors $A = \{(a_i, b_i, \omega_i, t)|t \in \psi, \omega_i \in U\}$;
5) the user tags $L = \{(c_i, \omega_i)|\omega_i \in U\}$.

2) Output: The model extracts the influencing mechanism of a user's multimessage interactions with a hot topic from the multiple messages themselves, $M = \{m_1, m_2, \ldots, m_n\}$, and their basic information, $N = \{(n_j, m_i)|m_i \in M\}$. Meanwhile, the user's personal attribute is extracted from the network $G_U^{m_i}=(U, m_i)$ of the user's participation in the message,

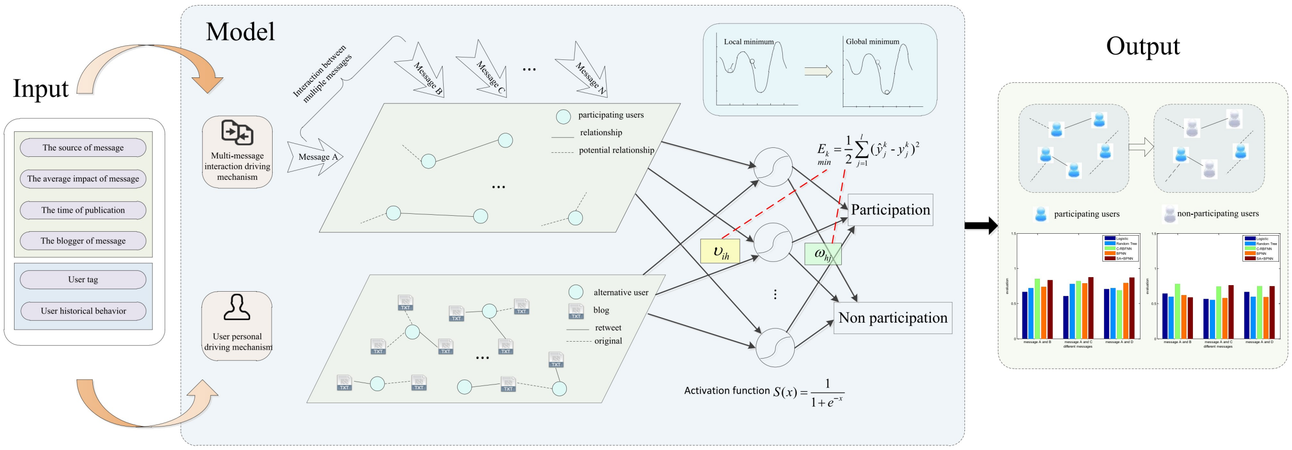


Fig. 1. Framework of user forwarding prediction.

the user's historical behaviors $A = \{(a_i, b_i, \omega_i, t) | t \in \psi, \omega_i \in U\}$, and the user's tag $L = \{(c_i, \omega_i) | \omega_i \in U\}$. Based on the conditions of the drive mechanism, the following problems are solved.

1) How should the multimessage interaction attribute and the user's personal attribute be combined to predict the user's participation (or not) in other messages of the hot topic? In this article, the output of the user behavior prediction matrix is derived from the input to the BP neural network. The behavior prediction matrix is defined as $Y^* = \text{argmax}_i P_i(Y_i | M, N, G, A, L)$. Under the influences of the user's personal attributes and the multimessage interaction, the predicted output is the value assigned a high probability in the behavior matrix.
2) How should the multimessage correlation indicators be calculated from the prediction results and how should the mutual influence strength between multiple messages be determined? This article quantifies the strength of the interaction between messages m_i and m_j by the multiple-message correlation indicator shown in Definition 5 and substitutes the prediction results into the developed formulas.

IV. Proposed Method

To solve the above-mentioned problems, first, the interaction influence-driving mechanism is extracted from the basic information of the multiple messages; next, the user's personal driving mechanism is extracted from the past history and tags of the user. The interactive-influence mechanism of multimessage leads to a more complex nonlinear relationship between the driving mechanisms' input and user participation behavior output, which is resolved by the BP neural network. However, under the interactive influence of multimessages on user behaviors, the BP neural network easily falls into the overfitting. To mitigate this problem, the traditional BP neural network is optimized by a simulated annealing algorithm, and a prediction model of the user participation behavior is established. Finally, the model predicts which user participating in a hot topic and will participate in other messages under that topic, and the multimessage correlation index between the multiple messages is defined. The impact of the topic on user participation behavior is then accurately characterized by quantifying the interaction intensity. The system framework is shown in Fig. 1.

TABLE I
Individual Driving Factors

Symbol	Description
$isRelativeTag(\omega_i)$	Whether user ω_i contain hot topic keywords
$rateOfRetweet(\omega_i)$	The historical forwarding rate of user ω_i
$activity(\omega_i)$	The activity of user ω_i

A. Influence Quantification

Users of social networks participate in the messaging of hot topics for diverse reasons. The purposes are roughly divided into two categories: the user's own interest and multimessage interaction. To quantify the influences of different factors on participation behavior, we need the influence attribute of user's participation behavior, which depends on both the user's personal interest characteristics and the influences of multimessage interactions. Specifically, we define the following driving mechanisms of different categories of influencing factors.

1) Driving Mechanism of User's Personal Characteristics: The behavior of users participating in hot topics is closely related to their individual hobbies, activity levels in the social network platform, and historical forwarding rates. The individual attribute values are shown in Table I.

The function $isRelativeTag(\omega_i)$ indicates whether the tags of user ω_i contain keywords related to a hot topic. The decision is made by a statistical comparison. If relevant keywords are included, $isRelativeTag(\omega_i) = 1$, otherwise $isRelativeTag(\omega_i) = 0$.

$rateOfRetweet(\omega_i)$ is the historical forwarding rate of user ω_i, which is specifically defined as follows: the number of user ω_i forwarding microblogs divided by the total number of microblogs, during the $[t_1, t_2]$ period

$$rateOfRetweet(\omega_i) = \frac{retweetNum(\omega_i)}{wholeNum(\omega_i)}, \quad \text{when } t_1 < t < t_2.$$

TABLE II
MULTIMESSAGE INTERACTION DRIVING FACTORS

Symbol	Description
$isDifT(m_p, m_q)$	Whether message m_p, m_q have same time
$isSameS(m_p, m_q)$	Whether message m_p, m_q have same source
$isSameB(m_p, m_q)$	Whether message m_p, m_q have same blogger
$influence(m_i)$	The average impact of message m_i

The activity of user is $activity(\omega_i) = \alpha * origNum(\omega_i) + \beta * retweetNum(\omega_i)$. The $origNum(\omega_i)$ and the $retweetNum(\omega_i)$ are the numbers of original and forwarding microblogs of user ω_i, respectively, during the $[t_1, t_2]$ period, α and β are the proportionality coefficients, and $\alpha, \beta \in [0, 1]$ are the adjustable parameters.

For descriptive convenience, the above-mentioned user's personal characteristics are unified into a term D_{ik} representing the kth attribute of the ith user ω_i: $D_{i1} = isRelativeTag(\omega_i)$, $D_{i2} = rateOfRetweet(\omega_i)$, and $D_{i3} = activity(\omega_i)$.

2) Driving Mechanism of Multimessage Interaction: Considering the actual situation of multiple messages under hot topics in social networks, the interaction of multiple messages will also affect the user's participation behaviors. Table II gives the attribute values defined by the multimessage interaction.

The functions $isDifT(m_p, m_q)$, $isSameS(m_p, m_q)$, and $isSameB(m_p, m_q)$, respectively, indicate whether the user participated in the message m_p, the forwarding source, and the blogger at the same time as messages m_q.

The average impact $influence(m_i)$ of message m_i in the hot topic is defined as $influence(m_i) = retweetNum(m_i) + comNum(m_i)$. $retweetNum(m_i)$ and $comNum(m_i)$, respectively, represent the average number of forwarding sources and comments in all microblogs by the blogger of message m_i during the $[t_1, t_2]$ period.

For descriptive convenience, the above-mentioned attributes of the multimessage interaction influence are unified into a term N_k representing the kth attribute. N_k has four elements: $N_1 = isDifT(m_p, m_q)$, $N_2 = isSameS(m_p, m_q)$, $N_3 = isSameB(m_p, m_q)$, and $N_4 = influence(m_i)$.

B. Details of the Method

The developed prediction model optimizes the weights and thresholds of a traditional BP neural network by a simulated annealing algorithm. Besides completing the nonlinear mapping of the multimessage interaction and user attributes to the user participation behavior, the annealing algorithm avoids the overfitting problem in the BP neural network. The optimization and prediction processes are detailed in the following.

1) Parameter Optimization of the Prediction Model: Although the traditional BP neural network can effectively fit the nonlinear relationships between the driving mechanisms' input and the user participation behavior output, it is easily overfitting because the multimessage interaction influences the iterative guidance of the user behaviors. To avoid this problem, the weights and threshold adjusted by the BP neural network are optimized by the simulated annealing algorithm. In this optimization, a local minimum is abandoned with a certain probability, and the solution approaches the global minimum, which improves the accuracy of the neural network prediction. The core part of the parameter optimization algorithm is based on Metropolis acceptance criteria that are formulated as follows:

$$p = \begin{cases} C(x_{\text{new}}) < C(x_{\text{old}}) \\ \exp\left(-\dfrac{C(x_{\text{new}}) - C(x_{\text{old}})}{T}\right) & C(x_{\text{new}}) \geq C(x_{\text{old}}). \end{cases}$$

Here, x_{new} and x_{old} represent the preoptimized and postoptimized states, respectively, and $C(x)$ is the evaluation function. The parameter optimization proceeds by the following steps.

1) Initialize the number of samples num, the temperature T, the temperature drop ratio α, the maximum annealing time L, and the termination temperature T_{min}. Assign the optimized weights and thresholds of the BP neural network to the initial value x_{old} of the simulated annealing algorithm.
2) Generate the new value x_{new} by inserting the current value x_{old} into a certain formula (which should be as simple as possible).
3) Calculate the increment $\Delta T = C(x_{\text{new}}) - C(x_{\text{old}})$ between the current value and the new value. The evaluation function $C(x)$ is generally determined by a plurality of parameters, such as the input and output of the neural network.
4) Determine whether to accept or reject the new value. The acceptance criterion is the Metropolis criterion: if $\Delta T < 0$, accept the new state x_{new} as the new current value; otherwise, accept x_{new} as the new current value with probability $\exp(-\Delta T/T)$.
5) If the temperature reaches the termination temperature, or if $k = L$, the current weight and threshold are output as the optimal values, and the program is terminated. The algorithm is shown in Algorithm 1.

2) Forwarding Prediction: First, the captured data are processed and quantified according to the driving mechanisms of the multimessage interaction and the user. This processing gives the value of each attribute of the input layer and the result set of whether the user will participate in other messages related to the topic. Some of the results recorded in the result set ($y_i = 1$ or $y_i = 0$) are selected as the training data. The remainders are made unknown ($y_k =?$). Based on expert experience, the learning rate η was set to 0.1, and all the connection weights and thresholds in the network were randomly initialized within the range $(0, 1)$. The BP neural network model is then trained on the training data set. As a training example in this network, suppose that the neural network output for (x_k, y_k) is $\hat{y}_k = (\hat{y}_1^k, \hat{y}_2^k, \ldots, \hat{y}_l^k)$:

$$\hat{y}_j^k = f(\beta_j - \theta_j) \tag{1}$$

where $\hat{y}_j^k$ represents the jth dimension of the actual output of the kth training case, and β_j and θ_j denote the input and threshold of the jth neuron of the output layer. The activation function f uses a sigmoidal function sigmoid $(x) = 1/(1 + e^{-x})$.

Algorithm 1 Parameter Optimization Algorithm

Input: Weights and thresholds of the BP neural network
Output: Current value $x_{current}$
// initialization and construction
Initialize num, T, α, L, T_{min}, x_{old};
for $k = 1 : L$ **do**
  $x_{new} = x_{old} + rand * (sum(abs(BPoutput - output\_exp)))/num$;
  // $(sum(abs(BPoutput - output\_exp)))/num$ is used to calculate the average error
  The evaluation function $C(x)$ is generally determined by a plurality of parameters, such as the input and output of the neural network;
  then calculate $\Delta T = C(x_{new}) - C(x_{old})$;
  if $\Delta T < 0$
    $x_{current} = x_{new}$;
  else if $rand < exp(-\Delta T/T)$
    $x_{current} = x_{new}$;
  else
    $x_{current} = x_{old}$;
  end if
  $T = \alpha * T$;
  // whether stop iterating according to threshold T_{min}
  if $T \leq T_{min}$
    break;
  end if
  $k = k + 1$;
end for

The gradient term g_j of the output layer neurons is calculated from the actual and expected outputs of the training examples. The calculation formula is as follows:

$$g_j = y_j^k\left(1 - \hat{y}_j^k\right)\left(y_j^k - \hat{y}_j^k\right) \tag{2}$$

where $\hat{y}_j^k$ and y_j^k represent the actual and expected outputs of the training case, respectively.

Similarly, the gradient term e_h of the hidden-layer neurons is calculated from the hidden and output layers by the following formula:

$$e_h = b_h(1 - b_h)\sum_{j=1}^{l}\lambda_{hj}g_j \tag{3}$$

where e_h represents the output of the hth neuron of the hidden layer, and λ_{hj} is the connection weight between the hth neuron of the hidden layer and the jth neuron of the output layer. The connection weights λ_{hj} and υ_{ih} and the thresholds θ_j and γ_h are updated based on the obtained gradient term and the input of the training set. The updated formulas are as follows:

$$\Delta\lambda_{hj} = \eta g_j b_h, \quad \Delta\upsilon_{ih} = \eta e_h x_i, \quad \Delta\theta_j = -\eta g_j, \quad \Delta\gamma_h = -\eta e_h. \tag{4}$$

The update process loops until the training error reaches a small preset value. After optimizing the weights and thresholds by the simulated annealing algorithm, the multilayer feed-forward neural network is constructed. The test samples are then input to the model, and the user's behavior-discrimination matrix is obtained. The entire prediction algorithm is shown in Algorithm 2.

Algorithm 2 Forwarding Prediction Algorithm

Input: Multiple messages $M = \{m_1, m_2, ..., m_n\}$ under hot topics
  Basic information $N = \{(n_j, m_i)|m_i \in M\}$ of multiple messages
  The user participation network $G_U^{m_i} = (U, m_i)$
  The user history behaviors $A = \{(a_i, b_i, \omega_i, t)|t \in \psi, \omega_i \in U\}$
  The user tags $L = \{(c_i, \omega_i)|\omega_i \in U\}$
Output: The behavior prediction matrix $Y^* = argmax_i P_i(Y_i|M, N, G, A, L)$
// initialization and construction
Initialize λ_n, θ, η, n;
repeat
  for $all(x_k, y_k) \in trainingdata$ **do**
    The output $\hat{y}_j^k$ of the current sample is calculated based on the initialization parameters and (1);
    The gradient term g_j of the neurons in the output layer is calculated based on (2);
    The gradient term e_h of the neurons in the hidden layer is calculated based on (3);
    The weights λ_n and υ_{ih} and the thresholds θ_j and γ_h are updated based on (4);
  end for
until converge
// The updated weights and thresholds are assigned to the initial solution of the simulated annealing algorithm, and simulated annealing is used to optimize the weights and thresholds
  for $all x_k \in testdata$
    predict $Y^* = argmax_i P_i(Y_i|M, N, G, A, L)$;
  end for

C. Algorithm Complexity

Considering the basic information $N = \{(n_j, m_i)|m_i \in M\}$ of the multiple messages in the popular discussion, the multi-message interaction-driving mechanism is extracted from the multiple messages $M = \{m_1, m_2, \ldots, m_n\}$ and the user participation network $G_U^{m_i} = (U, m_i)$ under the hot topic. Meanwhile, the user's personal driving mechanism is extracted from the historical behaviors data $A = \{(a_i, b_i, \omega_i, t)|t \in \psi, \omega_i \in U\}$ and tags $L = \{(c_i, \omega_i)|\omega_i \in U\}$ of the participants. The computational complexity is $O(N)$. In turn, when predicting the user participation behavior, the complexity of the improved BP neural network is $O(N^2)$. Therefore, the complexity of the whole algorithm is $O(N) + O(N^2) \sim O(N^2)$.

V. Experiments

A. Experimental Settings

This section introduces the experimental data and compares the performances of the improved BP neural network and

TABLE III
STATISTICS OF FOUR MESSAGES

Dataset	Message A	Message B	Message C	Message D	Topic
Blogger	*Xiu Xianlu*	*Lai Quzhijian*	*Sina entertainment*	*Xiu Xianlu*	*
Source	*Li Yutong*	*Li Yutong*	*Xue Zhiqian*	*Xue Zhiqian*	*
Time	9.13 22:39	9.13 00:41	9.15 21:36	9.15 21:46	*
Forward userNum	1,878	1,000	1,186	1,447	5511
Histrionic action	223	860	900	223	2206
ForwardNum of action	484,516	176,304	1,899,276	484,516	3044612
CommentNum of action	172,950	83,692	1,618,300	172,950	2047892

several baseline models. The proposed model is then tested by several evaluation methods.

1) Experimental Data: The experimental data set was collected from the Sina Weibo platform. Weibo is currently the largest social platform in China, attracting the largest number of people and discussion topics. For establishing a prediction model of user behaviors when multiple messages related to a common topic interact, this article extracted four messages related to the hot topic Xue Zhiqian and Li Yutong on the Sina Weibo platform. The selected messages were Li Yutong forwarded by Xiu Xianlu (Message A), Li Yutong forwarded by Lai Quzhijian (Message B), Xue Zhiqian forwarded by Sina Entertainment (Message C), and Xue Zhiqian forwarded by Xiu Xianlu (Message D). The entire topic propagation process includes 5511 message propagation nodes, and the user attributes and historical behaviors data are accumulated in 5 093 504. The prediction model of message behavior was established and evaluated by combining Message A, Message B, Message C, and Message D into three sets of messages. Table III shows the statistics of each message.

Message A: The Weibo user Xiu Xianlu forwarded the message of Weibo user Li Yutong at 22:39 on September 13, 2017. The number of participating users was 1878. From August 16 to September 17 of 2017, Xiu Xianlu published 223 microblogs, totaling 484 516 reposts and 172 950 comments.

Message B: The Weibo user Lai Quzhijian forwarded the message of Weibo user Li Yutong at 00:41 on September 13, 2017. There were 1000 participating users. Between August 17 and September 18 of 2017, Lai Quzhijian published 860 microblogs, totaling 176 304 reposts and 83 692 comments.

Message C: The Weibo user Sina Entertainment forwarded the message of Weibo user Xue Zhiqian at 21:36 on September 15, 2017. There were 1186 participating users. Sina Entertainment released a total of 900 microblogs from August 17 to September 18 of 2017, totaling 1 899 276 reposts and 1 618 300 comments.

Message D: The Weibo user Xiu Xianlu forwarded the message of Weibo user Xue Zhiqian at 21:46 on September 15, 2017. There were 1447 participating users. From August 16 to September 17 of 2017, Xiu Xianlu published 223 microblogs, totaling 484 516 reposts and 172 950 comments.

In this article, Messages A–D were constructed into different multimessage combinations. We extracted the influencing factors of the multimessage interactions and the personal attributes and historical behaviors of the participating users and predicted whether the users participating in (for example) Message A would also participate in Messages B–D under the influence of multiple messages.

2) Baseline Methods: To evaluate the BP neural network model optimized by the simulated annealing algorithm, its performance was compared with those of the following baseline methods.

C-RBF Neural Network: This model predicts the user forwarding behavior by an RBF neural network. To accommodate the uncertainty of the mapping relationships between the user attributes and forwarding behavior, the activation function of the hidden layer in the RBF neural network is optimized by cloud theory in fuzzy mathematics, creating the C-RBF neural network [31].

BP Neural Network: The BP neural network transforms complex problems into conventional classification problems in machine learning. The model accounts for the interaction between multiple messages and the user's basic information and obtains a predictive value of the user's behavior by an activation function.

Convolutional Neural Networks: CNN [35] is a feedforward neural network that contains convolution calculations and has a deep structure. We input topic features for convolution operations, extract more abstract features at the top, and then sort it out.

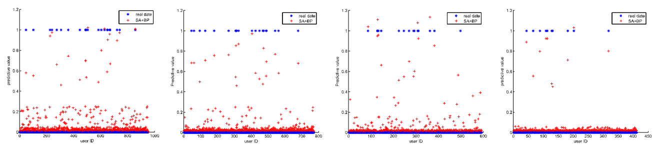
Long Short-Term Memory Networks: User behavior prediction is an experiment of time-series analysis. Due to time dependence, we naturally think about whether we can use the LSTM network [36] to classify time-series data.

Other Classifiers: Forward prediction can be regarded as a classification problem that marks each message as a positive instance (to be forwarded) or a negative instance (not to be forwarded). As baseline methods for the comparisons, we adopted two classic classifiers (logistic regression and random tree).

3) Evaluation Metrics: The BP neural network prediction model improved by the simulated annealing algorithm was competed against the baseline method. When the probability of participation in the output matrix is greater than the probability of not participating, the instance was recorded as a positive example "1." The vice-versa situation was reported as a negative example "0." The prediction results were expressed as a confusion matrix (see Table IV). The prediction results

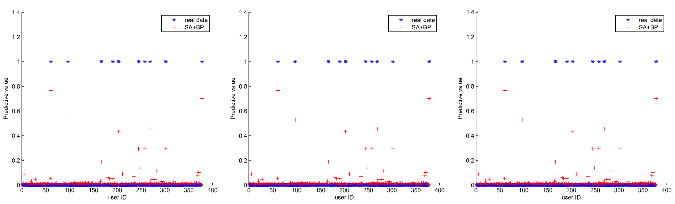
TABLE IV
CONFUSION MATRIX OF PREDICTION

actual class	predicted class	
	involved ("1")	not involved ("0")
involved ("1")	TP (True "1")	FN (False "0")
not involved ("0")	FP (False "1")	TN (True "0")



(a) (b) (c) (d)

Fig. 2. Prediction results under different proportions in messages A and C. (a) 50% and 50%. (b) 60% and 40%. (c) 70% and 30%. (d) 80% and 20%.



(a) (b) (c)

Fig. 3. Prediction results under different multimessage interactions. (a) Messages A and B. (b) Messages A and C. (c) Messages A and D.

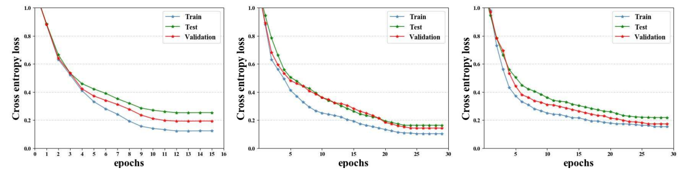
were evaluated by the precision, recall, and F1-measure indices that represent the accuracy, comprehensiveness, and comprehensive evaluation indicators of the user behavior prediction model.

B. Performance Analysis

1) Prediction Results: We first divided the data into a training set for training the BP neural network and a test set for prediction. The training and test data were divided as 50% and 50%, 60% and 40%, 70% and 30%, and 80% and 20%. As an example, Fig. 2 shows the prediction results of the multimessage combination of messages A and C. The blue "*" and red "+" are the real data and the predictions of the BP neural network model with the simulated annealing algorithm, respectively. The real data are the actual participation behavior of the user. If the user participates in the message set to 1, the user does not participate in the message set to 0.

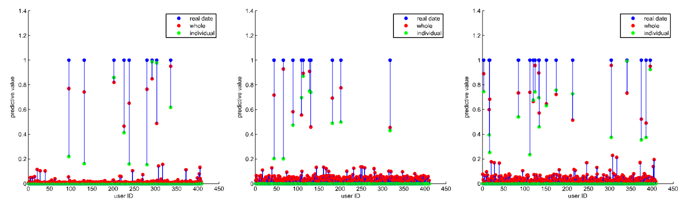
The experimental results of dividing the data set into 80% training data and 20% test data were relatively stable, so the subsequent experiments were performed at this ratio. The prediction result is shown in Fig. 3.

2) Ideal Target: Fig. 4 shows the variation of the cross entropy loss with a number of iterations during the training process. The blue, red, and green lines represent the loss performances of the BP training process, the BP cross-validation process, and the BP test process, respectively, in each generation. The most ideal training results were obtained after a certain number of iterations.



(a) (b) (c)

Fig. 4. Iterative process of the cross entropy loss indicator. (a) Messages A and B. (b) Messages A and C. (c) Messages A and D.



(a) (b) (c)

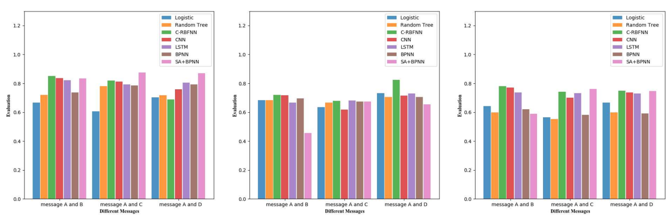
Fig. 5. Comparison of adding and not adding the multimessage interaction. (a) Messages A and B. (b) Messages A and C. (c) Messages A and D.

3) Multiple-Message Interaction Analysis: In Section 5.1, we analyzed the relationship between multimessage interaction and the number of participating users. The multimessage interaction-driving mechanism exerted certain impact on the prediction results. When the prediction results were analyzed without adding the multiple-message interaction effects, the predicted user participation behaviors were less accurate than when adding the multiple-message interaction effects (see Fig. 5).

4) Prediction Performance Analysis: The performance of the BP neural network prediction model optimized by the simulated annealing algorithm was compared with various baseline methods, namely the traditional neural network model and other classical classification algorithms. The performance was evaluated in more detail using the three evaluation metrics (precision, recall, and F1-measure introduced in Section V-A.3). The results are shown in Fig. 6. The horizontal and vertical coordinates in Fig. 6 represent the values of the multimessage combinations and evaluation indicators, respectively. Because the C-RBF model predicts the user behaviors during single messaging, it failed to resolve the overfitting problem during multiple-message interactions, so the prediction accuracy during multimessaging was low. The convolution kernel of CNN emphasizes the window in space. Due to the applicability of data density, the accuracy of using CNN to predict message propagation is low. Due to the advantages of LSTM for time-series data processing, it can learn the user's historical behavior well. However, the multimessage attribute is not fully captured. In contrast, the SA + BP model well predicted the user participation behavior during multimessaging. Fig. 6 can only measure the performance indicators of the model unilaterally. To comprehensively evaluate the predictive performance, this article adds a comparative experiment of the receiver operating characteristic (ROC) curves. In Fig. 7, the abscissa and ordinate represent the false positive rate (FPR) and true positive rate (TPR), respectively. For different topic combinations, the ROC curves of the SA + BP model

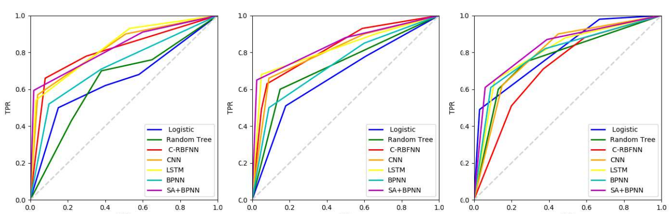
TABLE V
COMPARISON OF EXPERIMENTAL RESULTS OF DIFFERENT METHODS

Method	Message A and B			Message A and C			Message A and D		
	Prec.	Rec.	F1	Prec.	Rec.	F1	Prec.	Rec.	F1
Logistic	0.667	0.683	0.643	0.606	0.636	0.564	0.703	0.733	0.668
RandomTree	0.719	0.683	0.600	0.780	0.667	0.554	0.718	0.705	0.599
C-RBFNN	0.852	0.720	0.780	0.819	0.680	0.743	0.689	0.824	0.750
CNN	0.837	0.717	0.772	0.812	0.619	0.702	0.759	0.716	0.736
LSTM	0.821	0.668	0.736	0.793	0.681	0.732	0.805	0.671	0.731
BPNN	0.737	0.695	0.622	0.786	0.674	0.581	0.793	0.705	0.591
SA+BPNN	0.833	0.455	0.589	0.875	0.674	0.761	0.870	0.654	0.747



(a) (b) (c)

Fig. 6. Comparison of improved BPNN with baseline methods. (a) Precision. (b) Recall. (c) F1-measure.



(a) (b) (c)

Fig. 7. Comparison of different methods in ROC. (a) Messages A and B. (b) Messages A and C. (c) Messages A and D.

TABLE VI
MULTIMESSAGE RELEVANCE

Messages	Influence	Relativity
A and B	1.00	0.0226
A and C	2.00	0.086
A and D	2.73	0.148

remained very close to the upper left corner. Based on the good characteristics of the ROC curve, we conclude that the SA + BP model yielded the best predictive effect. The specific data are shown in Table V.

5) Correlation Analysis: We additionally predicted whether a user participating in one message (Message A) will also participate in other messages on the same topic (Messages B–D). To this end, we calculated the mutual influence strength between the quantized multimessages based on the predicted result and the correlation index, calculated as $J(m_i, m_j) = (\sum_{j=1}^{\text{count(user}(m_i))} t_{m_j}(\text{user}(\mathrm{m_i})))/(\text{count}(\text{user}(m_i)) * Influence(m_i, m_j), m_i, m_j) \in M$. The calculation results are shown in Table VI.

From the results, it is speculated that the correlation between multiple messages is proportional to the influence between them. In other words, the greater the influence, the stronger the correlation.

VI. CONCLUSION

From the user behavior data and the basic information data of multiple messages under a hot topic being discussed on a social network, this article extracted the driving mechanisms of both the user and the multimessage interaction and proposed a prediction model of the user's participation behavior in the discussed topic. First, the user's participation behavior was predicted by a BP neural network model, which copes with the complex nonlinear relationships between the input of the driving mechanisms of the user and the multimessage interaction and user behaviors' prediction output. Meanwhile, due to the iterative guidance of multiinformation interaction on user behavior, the BP neural network was degraded by the overfitting problem. After correcting the overfitting by a simulated annealing algorithm, the accuracy of the prediction was improved. Finally, we defined the multiple-message correlation metrics, statistically analyzed the model outputs, and estimated the proportion of users participating in one message, who also participated in other messages. The calculation results quantified the mutual influence strength between the multiple messages and accurately represented the influence of the hot topic on user participation behaviors.

The proposed method was experimentally evaluated on multimessage data under a hot topic discussed on the online social network, Sina Weibo. The model not only accurately predicted the user's participation behaviors but also quantified the intensity of the mutual influence between the multiple messages. Moreover, it dynamically perceived the situational changes in the hot topic, providing strong support for public opinion control.

REFERENCES

[1] M. Takayasu, K. Sato, Y. Sano, K. Yamada, W. Miura, and H. Takayasu, "Rumor diffusion and convergence during the 3.11 earthquake: A Twitter case study," *PLoS ONE*, vol. 10, no. 4, Apr. 2015, Art. no. e0121443.

[2] Y. Du, Y. He, Y. Tian, Q. Chen, and L. Lin, "Microblog bursty topic detection based on user relationship," in *Proc. 6th IEEE Joint Int. Inf. Technol. Artif. Intell. Conf.*, vol. 1, Aug. 2011, pp. 260–263.

[3] S. Gaglio, G. Lo Re, and M. Morana, "A framework for real-time Twitter data analysis," *Comput. Commun.*, vol. 73, pp. 236–242, Jan. 2016.

[4] G. Verma, A. Swami, and K. Chan, "The impact of competing zealots on opinion dynamics," *Phys. A, Stat. Mech. Appl.*, vol. 395, pp. 310–331, Feb. 2014.

[5] L.-L. Ma, C. Ma, H.-F. Zhang, and B.-H. Wang, "Identifying influential spreaders in complex networks based on gravity formula," *Phys. A, Stat. Mech. Appl.*, vol. 451, pp. 205–212, Jun. 2016.

[6] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and analysis of online social networks," in *Proc. 7th ACM SIGCOMM Conf. Internet Meas.*, 2007, pp. 29–42.

[7] X. Yun-Peng, L. Song-Yang, and L. Yan-Bing, "An information diffusion dynamic model based on social influence and mean-field theory," *Acta Phys. Sinica*, vol. 66, no. 3, 2017.

[8] Q. Su, J. Huang, and X. Zhao, "An information propagation model considering incomplete reading behavior in microblog," *Phys. A, Stat. Mech. Appl.*, vol. 419, pp. 55–63, Feb. 2015.

[9] L.-L. Xia, G.-P. Jiang, B. Song, and Y.-R. Song, "Rumor spreading model considering hesitating mechanism in complex social networks," *Phys. A, Stat. Mech. Appl.*, vol. 437, pp. 295–303, Nov. 2015.

[10] R. Jie, J. Qiao, G. Xu, and Y. Meng, "A study on the interaction between two rumors in homogeneous complex networks under symmetric conditions," *Phys. A, Stat. Mech. Appl.*, vol. 454, pp. 129–142, Jul. 2016.

[11] R.-Y. Tian, X.-F. Zhang, and Y.-J. Liu, "SSIC model: A multi-layer model for intervention of online rumors spreading," *Phys. A, Stat. Mech. Appl.*, vol. 427, pp. 181–191, Jun. 2015.

[12] L. Zhang, Z. Qi, L. Guo, and L. Xu, "Research on online social network information diffusion detection node selection algorithm based on the random walk model," *J. Comput. Theor. Nanosci.*, vol. 13, no. 1, pp. 971–981, Jan. 2016.

[13] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from twitter," in *Proc. IEEE 3rd Int. Conf. Privacy, Secur., Risk Trust, IEEE 3rd Int. Conf. Social Comput.*, Oct. 2011, pp. 149–156.

[14] G. Shen, W. Yang, W. Wang, and M. Yu, "Burst topic detection oriented large-scale microblogs streams," *J. Comput. Res. Develop.*, vol. 52, no. 2, pp. 512–521, 2015.

[15] C. Wang, Z. Zhang, J. Zhou, Y. He, J. Cui, and C. Jiang, "Modeling interest-driven data dissemination in online social networks," in *Proc. 12th Int. Conf. Mobile Ad-Hoc Sensor Netw. (MSN)*, Dec. 2016, pp. 290–295.

[16] M. Wang, W. Zuo, and Y. Wang, "A multidimensional nonnegative matrix factorization model for retweeting behavior prediction," *Math. Problems Eng.*, vol. 2015, pp. 1–10, Mar. 2015.

[17] L. Wei, Z. Zhang, D. Zhang, and S. C. Leung, "A simulated annealing algorithm for the capacitated vehicle routing problem with two-dimensional loading constraints," *Eur. J. Oper. Res.*, vol. 265, no. 3, pp. 843–859, Mar. 2018.

[18] A. Sheikhahmadi, M. A. Nematbakhsh, and A. Zareie, "Identification of influential users by neighbors in online social networks," *Phys. A, Stat. Mech. Appl.*, vol. 486, pp. 517–534, Nov. 2017.

[19] G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield, "Analysing the connectivity and communication of suicidal users on twitter," *Comput. Commun.*, vol. 73, pp. 291–300, Jan. 2016.

[20] M. Salehi, R. Sharma, M. Marzolla, M. Magnani, P. Siyari, and D. Montesi, "Spreading processes in multilayer networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 2, no. 2, pp. 65–83, Apr. 2015.

[21] Z. Luo, M. Osborne, J. Tang, and T. Wang, "Who will retweet me?: Finding retweeters in twitter," in *Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2013, pp. 869–872.

[22] Y. Artzi, P. Pantel, and M. Gamon, "Predicting responses to microblog posts," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2012, pp. 602–606.

[23] P. A. Grabowicz, J. J. Ramasco, B. Gonçalves, and V. M. Eguíluz, "Entangling mobility and interactions in social media," *PLoS ONE*, vol. 9, no. 3, p. e92196, 2014.

[24] K. Lee, J. Mahmud, J. Chen, M. Zhou, and J. Nichols, "Who will retweet this?: Automatically identifying and engaging strangers on twitter to spread information," in *Proc. 19th Int. Conf. Intell. User Interfaces*, 2014, pp. 247–256.

[25] K. Sankaram and M. F. Schober, "Reading a Blog when empowered to comment: Posting, lurking, and non-interactive reading," *Discourse Process.*, vol. 52, nos. 5–6, pp. 406–433, Jul. 2015.

[26] D. Huang, J. Zhou, D. Mu, and F. Yang, "Retweet behavior prediction in twitter," in *Proc. 7th Int. Symp. Comput. Intell. Design*, vol. 2, Dec. 2014, pp. 30–33.

[27] J. Huang and Q. Su, "A rumor spreading model based on user browsing behavior analysis in microblog," in *Proc. 10th Int. Conf. Service Syst. Service Manage.*, Jul. 2013, pp. 170–173.

[28] F. Xiong, Y. Liu, Z.-J. Zhang, J. Zhu, and Y. Zhang, "An information diffusion model based on retweeting mechanism for online social media," *Phys. Lett. A*, vol. 376, nos. 30–31, pp. 2103–2108, Jun. 2012.

[29] Z. Yang, "User-online load movement forecasting for social network site based on bp artificial neural network," *J. Comput.*, vol. 8, no. 12, pp. 3176–3183, 2013.

[30] Z. Li, D.-Y. Sun, J. Li, and Z.-F. Li, "Social network change detection using a genetic algorithm based back propagation neural network model," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2016, pp. 1386–1387.

[31] Y. Liu, J. Zhao, and Y. Xiao, "C-RBFNN: A user retweet behavior prediction method for hotspot topics based on improved RBF neural network," *Neurocomputing*, vol. 275, pp. 733–746, Jan. 2018.

[32] U. Sharma and B. Minocha, "Link prediction in social networks: A similarity score based neural network approach," in *Proc. 2nd Int. Conf. Inf. Commun. Technol. Competitive Strategies*, 2016, p. 90.

[33] D. Qing, M. Yefeng, L. Yi, and Z. Hui, "Prediction of retweet counts by a back propagation neural network," *J. Tsinghua Univ. (Sci. Technol.)*, vol. 55, no. 12, pp. 1342–1347, 2016.

[34] X. Huang, C. Shen, X. Boix, and Q. Zhao, "Salicon: Reducing the semantic gap in saliency prediction by adapting deep neural networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 262–270.

[35] Q. Zhang, Y. Gong, J. Wu, H. Huang, and X. Huang, "Retweet prediction with attention-based deep neural network," in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, 2016, pp. 75–84.

[36] M. Sundermeyer, R. Schlüter, and H. Ney, "LSTM neural networks for language modeling," in *Proc. 3th Annu. Conf. Int. Speech Commun. Assoc.*, 2012.

Yunpeng Xiao received the Ph.D. degree in computer science from the Beijing University of Posts and Telecommunications, Beijing, China, in 2013.

He is currently a Professor with the Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include social networks and machine learning.

Jinghua Li received the B.S. degree in software engineering from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2015, and the M.S. degree in 2019.

Her research interests include social networks and machine learning.

Yangfu Zhu received the B.S. degree in communications engineering from Xizang Minzu University, Xianyang, China, in 2014. He is currently pursuing the M.S. degree in computer science and technology with the Chongqing University of Posts and Telecommunications, Chongqing, China.

His research interests include social networks, deep learning, and its application.

Qian Li received the M.S. degree in computer science from the Chongqing University of Posts and Telecommunications, Chongqing, China, in 2008. She is currently pursuing the Ph.D. degree with the School of Computer, Beijing University of Posts and Telecommunications, Beijing, China.

Her research interests include social network, machine learning, and information dissemination dynamics.