Physica A 490 (2018) 185–202

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

An emotional contagion model for heterogeneous social media with multiple behaviors

Xi Xiong [a,1], Yuanyuan Li [b,1], Shaojie Qiao [a,*], Nan Han [c], Yue Wu [d], Jing Peng [e], Binyong Li [a]

[a] School of CyberSecurity, Chengdu University of Information Technology, Chengdu, 610225, China
[b] Mental Health Center, West China Hospital of Sichuan University, Chengdu, 610041, China
[c] School of Management, Chengdu University of Information Technology, Chengdu, 610103, China
[d] School of Computer and Software Engineering, Xihua University, Chengdu, 610039, China
[e] Department of Science and Technology Information, Sichuan Provincial Public Security Bureau, Chengdu, 610014, China

HIGHLIGHTS

- The transmissibility is introduced to measure the capability of spreading emotion.
- A model is proposed to describe the spatio-temporal features of emotion contagion.
- The transmissibility changes with interactions types and community structures.
- The proposed model shows better performance than other models of emotion contagion.
- The simulation results of our model reveal some interesting characteristics of social media.

ARTICLE INFO

Article history:
Received 23 April 2017
Received in revised form 13 August 2017

Keywords:
Emotional contagion
Multilayer network
Independent cascade model
User behavior
Social media

ABSTRACT

The emotion varies and propagates with the spatial and temporal information of individuals through social media, which uncovers several interaction mechanisms and features the community structure in order to facilitate individuals' communication and emotional contagion in social networks. Aiming to show the detailed process and characteristics of emotional contagion within social media, we propose an emotional independent cascade model in which individual emotion can affect the subsequent emotion of his/her friends. The transmissibility is introduced to measure the capability of propagating emotion with respect to an individual in social networks. By analyzing the patterns of emotional contagion on Twitter data, we find that the value of transmissibility differs on different layers and on different community structures. Extensive experiments were conducted and the results reveal that, the polar emotion of hub users can lead to the disappearance of opposite emotion, and the transmissibility makes no sense. The final emotional distribution depends on the initial emotional distribution and the transmissibilities. Individuals from a small community are more likely to change their mood by the influence of community leaders. In addition, we compared the proposed model with two other models, the emotion-based spreader–ignorant–stifler model and the standard independent cascade model. The results demonstrate that the proposed model can reflect the real-world situation of emotional

* Corresponding author.
E-mail address: sjqiao@cuit.edu.cn (S. Qiao).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.physa.2017.08.025

contagion for heterogeneous social media while the computational complexities of all these three models are similar.

1. Introduction

As a complicated psychological experience, the emotion has been proved to be transferred directly from one individual to another one via mimicry and copying of emotion-relevant bodily actions like facial expressions [1]. The individual emotion is affected by various non-verbal clues. Currently, the study on emotion has aroused the interest in a wide range of disciplines, including economics, neuroscience and psychology. Empirical studies have demonstrated that people can "catch" emotional states they observe from others over time ranging from seconds to months [2], and the probability of emotional contagion between strangers, even those in ephemeral contact, has been documented by the effects of "service with a smile" on customer satisfaction and tipping [3]. With the shift in the usage of the Web from information consumption to information production and sharing, numerous social media services have emerged. Internet has accumulated a large volume of user generated content. The online social media have been the main platform for communication and emotional expression. One of the examples is shown as follows:

Example 1. Jeff Bezos, the founder of Amazon, announced on Twitter that they had implemented the soft landing of rockets in 2015. The message was reposted and spread in a quick speed and his followers were immersed in an excited mood. Simultaneously, unhappiness and depression emerged and spread in the followers of SpaceX and its CEO Elon Musk, who expressed his unhappiness and responded later in his tweets that his rocket had completed sub-orbital flight for six times three years ago. His followers were inspired by these words and became exited in a short period of time.

We can see from the above example that emotion propagation in online social networks have the characteristics such as diversity, broad societal impact and unpredictability.

The motivations inspiring this work are given as follows:

(1) The individual emotion of social media varies with the time and the distance between individuals. To retrieve spatio-temporal features from large-scale Web data is an essential work, which can improve the accuracy of forecasting the spreading trend of emotion via social media.
(2) Social media feature several interaction mechanisms to facilitate the communication and the emotional contagion among users. The effect of different behaviors on emotional contagion is a very important issue to be considered.
(3) Users from the same online community often have specific relations (e.g. classmates or workmates) or share common interests. High frequent contacts in the same community make the emotion spread more quickly than across the communities. It plays an essential role of analyzing the effect of community structures on emotional contagion.

Based on the heterogeneity of social media, we propose an emotional independent cascade model to show the detailed process and characteristics of emotional contagion on social media. In this study, the original contributions are given as follows:

(1) The transmissibility, a network parameter, is introduced to measure the capability of spreading emotion between two users. The statistical results on a Twitter dataset show that the transmissibility value does change on different interaction layers as well as different community structures.
(2) In our proposed model, individual emotion can affect his/her friends' subsequent emotion. The proposed model demonstrates better performance than other typical models for emotional contagion, such as the emotion-based spreader–ignorant–stifler model [4] and the standard independent cascade model [5].
(3) The experiments reveal that the polar emotion of hub users leads to the disappearance of the opposite emotion, irrelevant to the transmissibility value. The emotional distribution depends on the initial emotional distribution and the transmissibility values. Users in a small community are more likely to change their moods by the influence of community leaders.

The remainder of this paper is organized as follows: Section 2 surveys related work in information diffusion, emotional contagion, heterogeneous networks and sentimental analysis. Section 3 provides the problem statement and the preliminary results and definitions. We introduce the emotional independent cascade model in Section 4. The experimental results are presented in Section 5 and we compare the different models in Section 6. Lastly, we conclude this study in Section 7.

2. Related works

In this section, we will introduce the relevant works from four research areas as given below.

2.1. Information diffusion

Information diffusion offers a basic condition for emotional contagion. The models on information diffusion have been intensively studied. Existing models are categorized into two classes: graph and non-graph [5]. Graph models include the independent cascade (IC) model [6,7] and the linear threshold (LT) model [8]. In the IC model, the probability on each edge of activating neighbors to spread information was proposed. Non-graph approaches use an epidemiological process to model information spreading and partition users into several categories that contain the probability for a user of one category changing into another one. Typical examples of non-graph models include the susceptible–infective–removed (SIR) [9] and spreader–ignorant–stifler (SIS) [10] models.

Recently, new models were proposed to adapt various situations [11,12]. To characterize the information transmitting in online social media, the authors [13] proposed a diffusion model (SCIR) which contains four possible states: susceptible, contacted, infected and refractory. Agents that read the information but have not decided to forward it will stay in the contacted state. They may become infected or refractory, and both the infected and refractory states are stable. An emotion-based spreader–ignorant–stifler (ESIS) model [4] categorizes information cascades into fine-grained classes, and the proportion of retweets among users for each kind of emotion is viewed as the weight of a edge. The probability of information adoption is based on the spreading probability as well as retweeting strength among users. Though these models can summarize some characteristics of information diffusion and emotional contagion on social media, they might overlook some significant information such as spatio-temporal information.

2.2. Emotional contagion

The spread of emotion such as happiness [14], depression [15] and loneliness [16] on social media was extensively studied based on the data of Framingham Heart Study (FHS) [17], including the 20-year longitudinal quantized data of these kinds of emotion and the interpersonal relations. Moreover, they analyzed the correlation of emotion between friends by generalizing estimation formulation, and finally concluded that emotion can be transmitted via social media with long-term effects. Various contributions have been made based on the observation that emotion can be propagated via online interactions [18]. A method for measuring the contagion of emotional expression shows that rainfall can directly influences the emotional content of human's status messages, and it also affects friends' status messages in other cities who are not experiencing rainfall [1]. A recent study from Facebook suggests that emotional contagion occurs online even in absence of non-verbal cues typical of interpersonal interactions [19,20]. The study performed a controlled experiment by choosing a sample of users and manipulating the content conveyed from their contacts, and finally found that these users' emotion were affected [21].

According to the aforementioned literatures, the authors often assume that emotional contagion phenomenon is often caused by the intrinsic characteristics of messages and the application scenarios of social media. However, the influence of structures and user behaviors has not been taken into careful consideration.

2.3. Heterogeneous networks

The heterogeneity of social media involves two aspects: the multiple interaction mechanisms and the community structure. These mechanisms reflect in different layers through which users can communicate and propagate emotion. Multilayer networks [22] can help us understand a myriad of complex systems, from social networks to biological systems, and most of them do not operate in an isolation layer but through multiple interconnected layers. Thus, the main advantage of this new methodology is that it incorporates multiple channels of connectivity, which makes it suitable to describe such systems in which the properties and neighbors of each vertex vary across layers. Borondo et al. [23] highlighted how the elite exert influence across layers, and it is considered to be a factor when studying the Venezuelan online protest of 2010 and the 2011 Spanish general elections. Individuals receive and diffuse information through three layers of different interaction mechanisms, including following, retweeting and mention behaviors. Hence, information diffusion of social media does not take place through a single channel. However, another important behavior of users is replying, which has been overlooked. Yagan et al. [24] studied the information diffusion in overlaying social–physical networks and indicated that information tends to spread among friends within a social community. However, it has been proved to be a very difficult work to collect data from physical networks.

2.4. Sentimental analysis

The emotional information from short texts in a tweet can be leveraged under various circumstances [25,26]. Several sentimental analysis algorithms can be used to capture positive and negative sentiment in short informal texts [27,28]. A tool called SentiStrength [29,30] is utilized in our work to annotate tweets with positive and negative sentimental scores. By Comparing with other tools, SentiStrength provides several advantages: it is designed for short informal texts with abbreviations and slang (features commonly observed in Twitter), and it employs linguistic rules for negations, amplifications, booster words, emotions, spelling corrections, which are particularly well-suited to process social media data. SentiStrength was proven able to capture positive emotion with 60.6% accuracy and negative emotion with 72.8% accuracy.

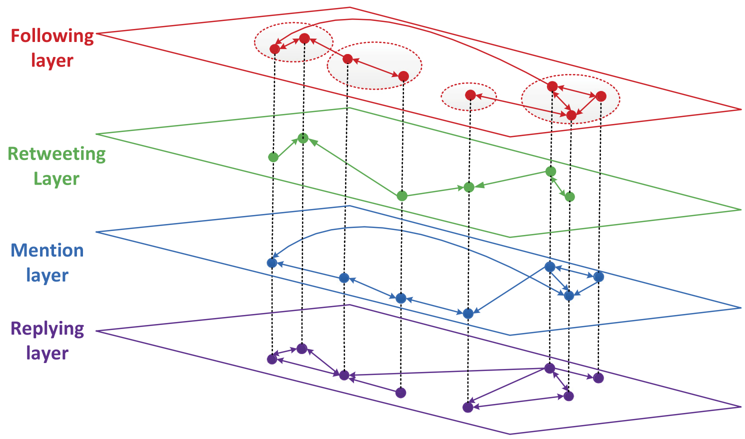


Fig. 1. The multilayer structure of Twitter. The network contains four layers: following, retweeting, mention and replying. Not all users have to be present on all layers, and edges are not necessarily repeated on more than one layer.

The aforementioned models revealed the features of the process of emotional contagion on social media. However, to our best knowledge, the influence of different behaviors and a heterogeneous structure on emotional contagion has not been studied properly. Our proposed model takes a user's capability of spreading emotion into consideration, and involves the spatio-temporal relations among users.

3. Problem statement and preliminaries

The structural and dynamic characteristics of social media are more complex than those of the simple networks, e.g., the random networks, the small-world networks and the scale-free networks. Multiple factors including the different behaviors and the heterogeneous structures of social media play an essential role in studying emotional contagion. The objectives of this paper can be addressed as follows:

(1) The primary objective of this paper is to establish a new model to describe emotional contagion by taking into account the important effects hidden in social media.
(2) The secondary objective is to discover the rules of emotional contagion on social media and forecast its trend via the proposed model.

We propose an emotional independent cascade model (eIC for short) to predict the process of emotional contagion. Based on the standard independent cascade model (IC for short), a sliding time-window is introduced into the model. It is worthwhile to notice that the edge weight represents the interaction strength between two adjacent vertices. Neighbors with high edge weights might be viewed as close friends or relatives, whose mutual influences are stronger than those of neighbors with low weights, representing less significant relations in individuals' social networks. Edge weight is a suitable parameter that can extend the basic property of human social relationships to the emotion domain.

Social media feature several interaction mechanisms to facilitate the communication and emotional contagion. The first interaction mechanism is introduced in [31], a passive mechanism that allows users to receive the messages written by people they have followed at any time. Meanwhile, the users automatically deliver their posted messages to their followers. Twitter also allows users to retransmit or retweet messages posted by others [32]. The retweet mechanism allows individual messages to distribute via social media. By mentioning someone's username in the message text, people are able to send directed messages to other users. Whenever a user is mentioned on a tweet, he/she gets notified about it, as it appears in his/her private inbox, significantly increasing the chance of reading it. A user allows his/her followers to reply or make comments on his/her tweets. It is the direct mechanism to discuss a topic. When a tweet is replied by someone, the author can be notified automatically.

By distinguishing the different interaction mechanisms available among users, a social network like Twitter is modeled as a multilayer directed graph, $G = \cup_{\alpha=1}^{4} G_\alpha$, where G_α is a topology graph on layer α. As show in Fig. 1, the four layers are represented as follow (α=1), retweet (α=2), mention (α=3) and reply (α=4). In a specific layer, $V_\alpha = \{v_1, \dots, v_{m_\alpha}\}$ denotes a set of users, called vertices, where m_α is the number of users. $E_\alpha = \{e_1, \dots, e_{n_\alpha}\}$ represents a set of directed edges. The emotion diffuses along the edges.

Here, we present some important concepts used in the proposed eIC model.

Definition 1 (*Vertices*)**.** The users on different layers are represented as a set of vertices. Among the four levels, the following layer provides the environment for users to communicate, and all users are included. The vertices on the other three layers are a subset of the vertices on the following layer.

Table 1
The meanings of $A \rightarrow B$ on different layers.

Action	The meaning of $A \rightarrow B$
Follow	*A* follows *B*
Retweet	*A* retweets a tweet to *B*
Mention	*B* is mentioned in *A*′s tweet
Reply	*A* replies to *B*′s tweet

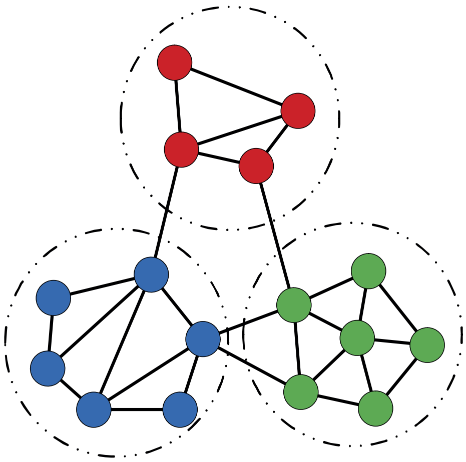


Fig. 2. The community structure of Twitter. The classical *GN* algorithm [33] is used to divide users into different communities.

Definition 2 (*Following Layer*)**.** As shown in Fig. 1, the top layer represents the following layer, where a directed edge represents who followed whom. The edges on this layer can establish a substratum by which emotion is delivered.

Definition 3 (*Retweeting Layer*)**.** The second top layer represents the retweeting layer. A directed edge on this layer indicates who retweeted a tweet to whom. The topology of the following layer has a big impact on retweeting, as it raises a high level of disparity in the message reception and the information spreading process. This layer is considerably sparser than the following layer. This fact shows that users are much more selective when actively spreading information, than when just receiving or reading it [31].

Definition 4 (*Mention Layer*)**.** The third layer represents the mention layer. On this layer, a directed edge indicates who mentioned whom in his/her tweets text. This mechanism is often used for establishing conversations between users, or just referring someone [20]. Hence, a new edge appeared when a new message posted by user A made a mention of user B (denoted by A→B). B was notified about the new tweet, which significantly increased the chance of reading it.

Definition 5 (*Replying Layer*)**.** The bottom layer (bottom) represents the replying layer. On this layer, a directed edge indicates who replied to whose tweets.

As illustrated in Fig. 1, not all users have to appear on all layers. Similarly, edges are not necessarily repeated on more than one layer. Table 1 shows the different meanings of the expression, user *A* points to user *B*, on different layers.

We can conclude that users with different positions show various probabilities to infect or be infected by friends. One of the most significant structural characteristics of the networks is communities.

Definition 6 (*Community*)**.** Communities are groups of connected vertices with a high density, while the connections between groups are sparse. The follow relations between users are static and relatively stable. The following layer contains all the users, and shows better community structures than the other layers. As is shown in figure Fig. 2, we use the classical *GN* algorithm [33] to divide users into different communities.

Users in the same community usually have the specific relation such as classmates or workmates, or share the common hobbies. High frequent contacts in the same community make the emotion spread more quickly than across the communities.

Definition 7 (*Transmissibility*)**.** Transmissibility is introduced to measure the ability of emotion propagation among users, who may be located in the same community, or belong to different communities. The transmissibility is also be affected by users' behaviors.

Table 2
The statistical summary of the linear regression.

Parameter	Value
Significant level	$8 * 10^{-6}$
Standard deviation	828
Correlation of x and y	0.946
A	(0.35,0.13;1.17,0.41;0.92,0.33)
b	46.91

Table 3
The transmissibilities of different behaviors and positions.

	Same community	Different communities
Retweet	0.35	0.13
Mention	1.17	0.41
Reply	0.92	0.33

If we want to classify and measure the features of emotional contagion quantitatively, the tools for sentimental analysis like SentiStrength are necessary. Each tweet is classified as positive or negative emotion and assigned a positive $S^+(t)$ and negative $S^-(t)$ sentimental score. Both scores are on a scale ranging between 1 (neutral) and 5 (strongly positive or negative).

Definition 8 (*Polarity Score*)**.** To capture the sentiment expressed by each tweet in one single measure, we define the polarity score $S(t)$ as the sum of positive and negative sentimental scores assigned to a tweet in the following equation:

$$S(t) = S^+(t) - S(t)^- \tag{1}$$

The polarity score $S(t)$ ranges between -4 (extremely negative: $S^+(t) = 1$ and $S^-(t) = 5$) to +4 (extremely positive: $S^+(t) = 5$ and $S^-(t) = 1$). When the positive and negative sentimental scores for a tweet are the same ($S^+(t) = S^-(t)$), we say that the polarity of the tweet is neutral ($S(t) = 0$).

When emotion is weak, i.e., the polarity score is close to 0, it is considered to be neutral. The emotional tendency is defined based on the polarity score.

Definition 9 (*Discrete Emotional Tendency*)**.** Discrete emotional tendencies are the types of emotion based on the emotional polarity score. When $S(t)$ ranges from -4 to -2, the emotional tendency is negative and the emotion is viewed as negative; when $S(t)$ ranges from -1 to 1, the emotional tendency is neutral and the emotion is viewed as neutral; when $S(t)$ ranges from 2 to 4, the emotional tendency is positive and the emotion is called positive.

Definition 10 (*Continuous Emotional Tendency*)**.** A continuous value of emotion is a necessary in the time-varying model. We use θ_1 as the boundary between the positive emotion and the neutral emotion, and θ_2 as the boundary between the negative emotion and the neutral emotion. A uniform distribution of continuous emotional polarity score between -4 and 4 in the proposed model requires θ_1 =1.33 and θ_2 = -1.33.

Definition 11 (*Time Step*)**.** In order to show the evolution process of the model clearly, the continuous time is converted into a set of short periods. One of the periods is called a time step.

Each user influences the emotion of his/her neighbors by the three types of behaviors. The details of the dataset are described in Section 5.1. Fig. 3 shows the time distribution of the three behaviors and the change of emotion in a whole day. It is obvious the peaks do exist in the morning, in the afternoon and in the evening, respectively. The users' emotion changes with their activities and the network structure. All the data is the average value of inconsecutive three days. In every two hours (a time step), the emotion of each user may change from -4 to 4. We adopt the absolute value to measure the influence of different behaviors and different communities. The change of absolute emotion value of users can be indicated as:

$$y = \sum_{k \in V} \sum_{t} \mid \Delta E \mid_{kt} = sum(AX) + b \tag{2}$$

where ΔE represents the change of a user's emotion in one time step and the two summation operations are calculated for all the vertices and all the time steps. The symbol sum($\bullet$) denotes the summation of all the elements in the matrix. Each element of the matrix $X = (x_{ij})_{2\times 3}$ counts the occurrence of each behavior in (i=0) or out (i=1) of the community. The constant matrix A has a 3×2 dimension and each element a_{ji} represents the effect of different behaviors and positions on users' emotion.

We adopted the *linear regression* to get the values of matrix A. Table 2 is the statistical summary while the confidence coefficient is 95%. Table 3 shows that the emotional transmissibilities of the three behaviors (retweet, mention and reply) in the same community are 0.35, 1.17 and 0.92, and the values between different communities are 0.13, 0.41 and 0.33.

4. The emotional independent cascade model

The independent cascade model (IC model), models the propagation process in discrete time steps by using a weighted graph. At a given time step, an active vertex v attempts to influence his/her inactive neighbor w. w may become active with a probability equaling to the weight of the link from v to w (without regard to the state of any other neighbors). The edge weight models the influence a vertex has on its neighbors. If w is not active, v can make no more attempts to influence w. One might consider neighbors with high edge weights as close friends or relatives and have stronger influence. Otherwise, neighbors with low weights means less significant. The Edge weight is such an important parameter that it models a basic property of human social networks and is easy to be utilized in the emotion domain.

According to Twitter data, we found that users on social media were able to affect their neighbors' emotion. In addition, only friends within the distance of three and the time interval of three days are relevant. Different interaction mechanisms and the community structure have various effects on emotional contagion.

By taking these characteristics into consideration, we present the following emotional independent cascade model (eIC model). It is assumed that the emotional states of all the vertices are updated successively, not simultaneously. At a step, each type of interactions (retweet, mention and reply) happens only once respectively between two vertices according to the topology of each layer. In order to describe the model clearly, some notations in the proposed model is shown in Table 4.

On one of the layers such as the retweeting layer, the interaction weight from user i to user j is computed by:

$$w_{ij}^{\alpha}(t) = \frac{n_{ij}^{\alpha}(t, \Delta t)}{D_j^{\alpha}(t, \Delta t)} \tag{3}$$

where $n_{ij}^{\alpha}(t, \Delta t)$ and D_j^{α} represents the occurrence number of interaction from i to j and the total occurrence number of interaction from j's neighbors to j within $[t - \Delta t, t]$ on layer α, respectively. If user i takes a behavior such as retweeting a message towards j, the emotion of j is converted to:

$$\frac{\delta E_j(t)}{\delta t} = \varepsilon_{ij}^{\alpha} \Delta E_{ij}(t) w_{ij}^{\alpha}(t) \tag{4}$$

where $\varepsilon_{ij}^{\alpha}$ is the transmissibility between i and j on the retweeting layer as defined in Definition 9, whose value depends on statistical results in the past and treated as time-invariant. $\Delta E_{ij}(t)$ represents the emotional difference between i and j at step t. Eq. (4) means that j would imitate the emotion of i, in other words, the emotion spreads from i to j. Taking all the neighbors of j into consideration, the emotion of j is changed to:

$$\frac{\delta E_j(t)}{\delta t} = \sum_{i \in B(j)} \varepsilon_{ij}^{\alpha} \Delta E_{ij}(t) w_{ij}^{\alpha}(t) \tag{5}$$

where $B(j)$ is the set of j's prior neighbors.

Then we take Eq. (3) to Eq. (5) and obtain the following result:

$$\frac{\delta E_j(t)}{\delta t} = \sum_{i \in B(j)} \varepsilon_{ij}^{\alpha} \Delta E_{ij}(t) \frac{n_{ij}^{\alpha}(t, \Delta t)}{D_j^{\alpha}(t, \Delta t)} \tag{6}$$

$B(j)$ are divided into two subsets: $B_{in}(j)$ and $B_{out}(j)$ which are j's prior neighbors existing in the same community and across different communities, respectively. Given the transmissibilities in the same community and across different communities on layer α are represented by $\varepsilon_{in}^{\alpha}$ and $\varepsilon_{out}^{\alpha}$, respectively. Then we have:

$$\frac{\delta E_j(t)}{\delta t} = \frac{\varepsilon_{in}^{\alpha}}{D_j^{\alpha}(t, \Delta t)} \sum_{i \in B_{in}(j)} \Delta E_{ij}(t) n_{ij}^{\alpha}(t, \Delta t) + \frac{\varepsilon_{out}^{\alpha}}{D_j^{\alpha}(t, \Delta t)} \sum_{i \in B_{out}(j)} \Delta E_{ij}(t) n_{ij}^{\alpha}(t, \Delta t) \tag{7}$$

Now we discuss a special circumstance in which a user only interacts with close friends in the same community, and is hardly influenced by users outside the community. That is to say $\varepsilon_{in}^{\alpha} \gg \varepsilon_{out}^{\alpha}$, and the equation is converted to:

$$\frac{\delta E_j(t)}{\delta t} = \frac{\varepsilon_{in}^{\alpha}}{D_j^{\alpha}(t, \Delta t)} \sum_{i \in B_{in}(j)} \Delta E_{ij}(t) n_{ij}^{\alpha}(t, \Delta t) = \varepsilon_{in}^{\alpha} * \overline{\Delta E_{j,in}(t)} \tag{8}$$

where $\overline{\Delta E_{j,in}(t)}$ is the weighted average of the emotion of j's prior neighbors in the same community at step t. We compute the integral for both sides of Formula (8) and obtain the following equation:

$$E_j(t) = E_j(t_0) + \varepsilon_{in}^{\alpha} \sum_{\tau = t_0}^{t} \overline{\Delta E_{j,in}(\tau)} \tag{9}$$

Finally, we take all the three layers into consideration and obtain the following equation:

$$E_j(t) = E_j(t_0) + \sum_{\alpha} \varepsilon_{in}^{\alpha} \sum_{\tau = t_0}^{t} \overline{\Delta E_{j,in}(\tau)} \tag{10}$$

Eq. (10) shows that the increment of user j's emotion equals to the accumulative weighted average of the difference between j's prior friends and j over all layers.

The process of the proposed eIC model is described as follows:

Algorithm 1 eIC model

Input: S: The n-dimensional vector storing users' emotion (including initial emotion).
R_{init}: The $3 \times n \times n$ matrix storing the occurrence of behaviors before the initial time point.
R: The $3 \times n \times n \times sn$ matrix storing the occurrence of behaviors every step.
H: The $4 \times n \times n$ adjacent matrix of the four layers with the following layer as the first layer.
sn: The number of evolutionary steps.
Δt_{init}: The sum of evolutionary steps before the initial time point.
ε: The 3×2 matrix storing six constant elements, each of which represents the transmissibility on one of the three behavioral layers in the same community or across different communities.
Output: S: The n-dimensional vector storing users' emotion (including final emotion at step sn).

```
1:  for each t ∈ [1, sn] do
2:      initialize T (a 3 × n × n matrix);
3:      for each α ∈ {L_R, L_M, L_S} do
4:          V ← GetVertex(H[α]);
5:          for each j ∈ V do
6:              priors ← GetPriorNeighbors(H[α], j);
7:              for each i ∈ priors do
8:                  if t ⩾ Δt then
9:                      T[α, i, j] ← SubCount(R[α], t − Δt + 1, t, i, j);
10:                 else
11:                     q ← (Δt − t)/Δt_init;
12:                     T[α, i, j] ← ⌊R_init[α] ∗ q⌋+SubCount(R[α], 1, t, i, j);
13:                 end if
14:             end for
15:         end for
16:     end for
17:     for each α ∈ {L_R, L_M, L_S} do
18:         V ← GetVertex(H[α]);
19:         for each j ∈ V do
20:             priors ← GetPriorNeighbors(H[α], j);
21:             behavior_sum ← GetBehaviors(T, α, all, j);
22:             for each i ∈ priors do
23:                 behavior_number ← GetBehaviors(T, α, i, j);
24:                 w ← behavior_number/behavior_sum;
25:                 if IsOneCommunity(H(1), i, j)=true then
26:                     ε ← ε_{α,in};
27:                 else
28:                     ε ← ε_{α,out};
29:                 end if
30:                 S[j] ← S[j]+(S[i]-S[j])*w*ε;
31:             end for
32:         end for
33:     end for
34:     t ← t + 1;
35: end for
```

The working mechanism of Algorithm 1 is given as follows,

1. Repeat the evolutionary process sn times (line 1). For each time step, the process is partitioned into two subsections: the first one is used for calculating the number of each behavior during $[t - \Delta t, t]$ (lines 3–16) and the second one is used for calculating the new emotion of each user (lines 17–33).
2. Initialize T which is a $3 \times n \times n$ matrix (line 2).
3. Repeat the process on three different behavioral layers (line 3).
4. Obtain the vertex set from the adjacent matrix on layer α (line 4).
5. Repeat the process for all users on this behavioral layer(line 5).
6. Obtain j's prior users on this behavioral layer (line 6).
7. Repeat the process for j's prior neighbors (line 7).

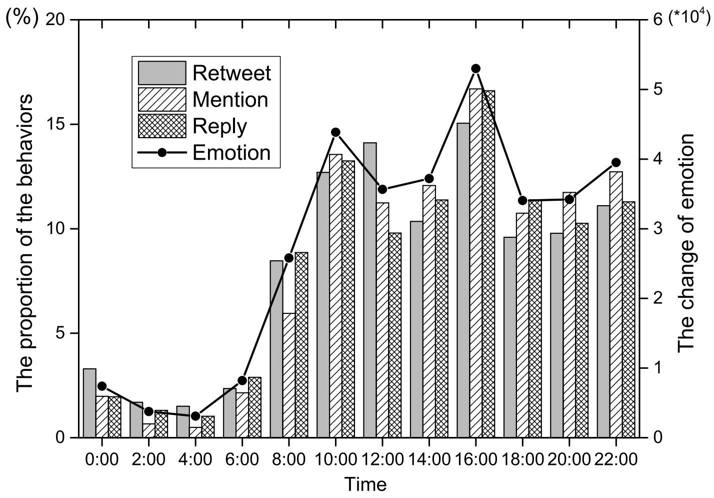


Fig. 3. The time distribution of the three behaviors and the change of emotion in a whole day. All the data is the average value of inconsecutive three days.

Table 4
The notations in the proposed model.

Symbol	Description
i, j	User i and user j
L_R, L_M, L_S	The retweeting layer, the mention layer and the replying layer
α	One of the three behaviors (layers)
$n_{ij}^{\alpha}(t, \Delta t)$	The occurrence number of interaction from i to j within $[t-\Delta t, t]$ on layer α
$D_j^{\alpha}(t, \Delta t)$	The total occurrence number of interaction from j's neighbors to j within $[t-\Delta t, t]$ on layer α
$w_{ij}^{\alpha}(t)$	The interaction weight from i to j on layer α
$\varepsilon_{ij}^{\alpha}$	The transmissibility between i and j on layer α
$\Delta E_{ij}(t)$	The emotional difference between i and j at step t
$B(j)$	The set of j's prior neighbors
$B_{in}(j), B_{out}(j)$	j's prior neighbors existing in the same community and across different communities

8. Count the occurrences of behaviors within the given time window (line 8–13). The function SubCount($R[\alpha], t_1, t_2, i, j$) is used for calculating the occurrence of behavior α from i to j during $[t_1, t_2]$.
9. lines 17–22 have the same meaning as lines 3–7.
10. The functions GetBehaviors(T, α, all, j) (line 21) and GetBehaviors(T, α, i, j) (line 23) are used for summing the occurrence of behavior α from j's priors to j and i to j, respectively.
11. Calculate the percentage of behavior α from i to j (line 24).
12. Obtain the transmissibility between i and j from the matrix ε. The value depends on the mutual positions of the two users and the behavioral type. The function IsOneCommunity($H(1), i, j$) is used for obtaining the truth whether the two users are in the same community on the following layer (lines 25–29).
13. Update the emotion of j according to Eq. (5) (line 30).
14. Update the time step (line 34).

Fig. 4 illustrates the process of the emotional contagion and the shift of the time window in the eIC model in a simple way. On one of the behavioral layers, user 1 is superior to other four users since a higher index. At the initial state, the time window is $[t, t+\Delta t]$. The emotional tendencies of users vary and user 1 has the negative emotion. First, user 3 spreads his/her positive emotion to user 1 and user 2, and changes their polarity scores. The tendencies of user 1 and user 2 become negative and neutral respectively. Second, user 4 spreads his/her emotion to user 3 and decreases his/her polarity score. However, the polarity score of user 3 is still larger than 1, and the tendency of user 3 maintains positive. Finally, user 5 spread the negative emotion to his/her neighbors. User 3 and user 4 decrease their polarity score, and then become neutral and negative respectively. At the following time steps, the time window is shifted forward, and all the users update their polarity scores and emotional tendencies sequentially. The negative emotion gradually spreads from user 5 to other users in this sub-network.

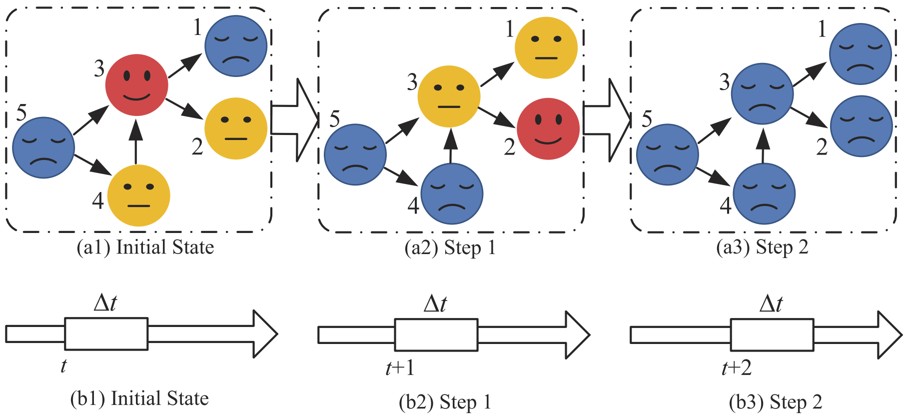


Fig. 4. The process of the emotional contagion and the shift of the time window.

5. Experimental results

5.1. Dataset description

Twitter, a typical online social network, has become one of the most popular social media on the Internet. With 271 million monthly active users in 2015, Twitter has demonstrated its strength for information propagation and emotional contagion. A large volume of data were collected from Twitter by our designed web crawler, and the main information include users' profiles and their posts. After one hour of collection from Twitter, 33,070 user profiles and approximately 180,000 tweets from March 2016 to May 2016 were obtained. We can calculate that the average degree without direction was 45.4 and the clustering coefficient was 0.18. The degree distribution agreeing with the power law is a typical scale-free network [34].

5.2. The relation between emotional contagion and homophily

Homophily is the principle that a contact between similar people occurs at a higher rate than among dissimilar people [35]. The pervasive fact of homophily means that cultural, behavioral, genetic or material information that flows through networks tends to be localized. Homophily is translated into a network distance, i.e., the number of segments between two individuals that a piece of information travels through. As shown in Fig. 5, a tweet posted by a user has the capability of affecting his/her friends subsequent emotion. The emotion of a user is relevant to those of users within the distance of three segments, and the correlation becomes smaller when the distance becomes larger. When the distance of two users is larger than three segments, it is hard for them to affect others. Moreover, it is easier for the negative emotion such as depression, loneliness and anger to propagate than the positive one like happiness. Meanwhile, the positive emotion of a user is suppressed by his/her friends negative expressions. With respect to FHS dataset, we can see the feature of three degrees of separation, and this conclusion was also verified in Twitter [16].

According to Fig. 5, we can see several similar patterns: when a user initiates a behavior, the words they choose influence the words chosen later by their friends. This effect is consistent with the previous research in emotional contagion [36], that is, ones friends who express emotional language end up expressing more language with the same tendency as well. In other words, users emotion will be affected by the words of themselves and their friends on the prior days.

5.3. The evolution of the emotional distribution

The information of vertices and relations were retrieved from the dataset. The information of vertices contains their emotion at each timestep. The interval between any two timesteps is two hours. The distribution of vertices depends on the directed interaction relations and the communities formed on the following topology.

Fig. 6 shows the initial and final emotion distributions of the evolution. We only illustrate the vertices with degrees larger than 100. The size of a vertex depends on its degree, including the indegree and the outdegree. The edge color agree with its source vertex. The polarity scores of the initial emotion of vertices are assigned a continuous random values between -4 and 4. The red and blue vertices represent the users with positive and negative emotion, respectively. Darker colors imply stronger emotion, and a light color indicates a weak emotion. The white vertex represents a neutral emotion.

The emotional distribution in Fig. 6(a) is random because of the differences between individuals in the usual circumstances. Only several small communities are composed of vertices with similar emotion. The positive hub vertices are a little bit more than the negative ones.

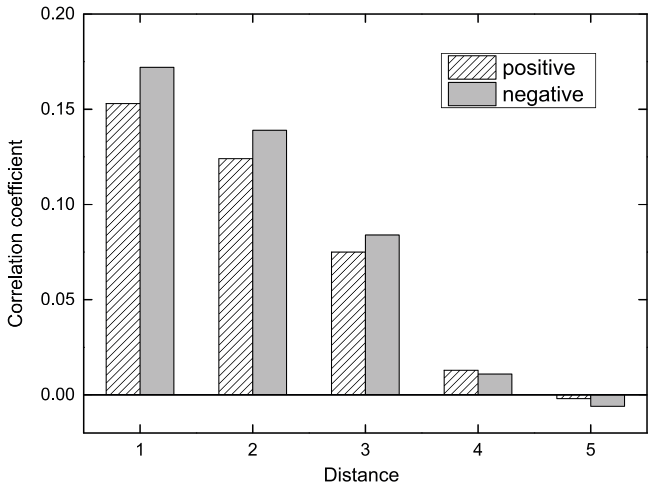
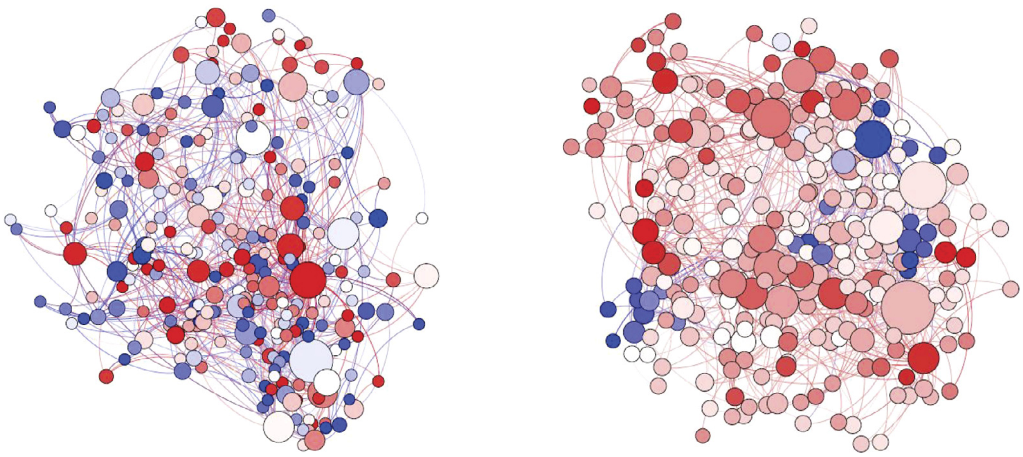


Fig. 5. A tweet posted by a user has the ability to affect his/her friends' subsequent emotion. The emotion of a user is relevant to those of users within the distance of three, and the correlation becomes smaller when the distance becomes larger.



(a) The initial distribution.

(b) The final distribution.

Fig. 6. The initial and final distributions of the model evolution. A vertex with a red, white or blue color represents a user with a positive, neutral or negative emotion, respectively. Only the vertices with a degree larger than 100 are illustrated. The size of a vertex depends on its degree on the following layer, including the indegree and the outdegree.

Then we analyze the emotional influence of hub vertices. We changed the emotion of three hub vertices to positive manually to simulate the condition that the community leaders held positive emotion. These users were viewed as optimistic, happy and kept active. After the 100-step evolution, the emotional distribution is shown in Fig. 6(b). Most of the vertices became positive. Only the users in several small communities were left negative because they were affected by the leaders of the small communities. These community leaders are known to have defects in their characters or even have mental diseases, and they usually feel unhappy and are never affected by others.

In order to see the evolutionary process of the eIC model clearly, rather than the emotional distribution at a specific timestep, as shown in Fig. 7, the increment of positive vertices mainly comes from the negative ones, not the neutral ones. In the beginning of the process, both the positive and negative emotion spread in the network, and neither predominate. Some of the positive or negative users near the boundaries become neutral, which leads to the decrease of the positive and negative ones. Then the three artificial community leaders of great influence play a key role in the evolution of the network, and change the emotion of others to positive gradually. Consequently, the positive and neutral lines rise and the negative line keeps descending. Most of the negative vertices disappear, except the defective users. Then other neutral vertices become positive because of the influence of positive ones.

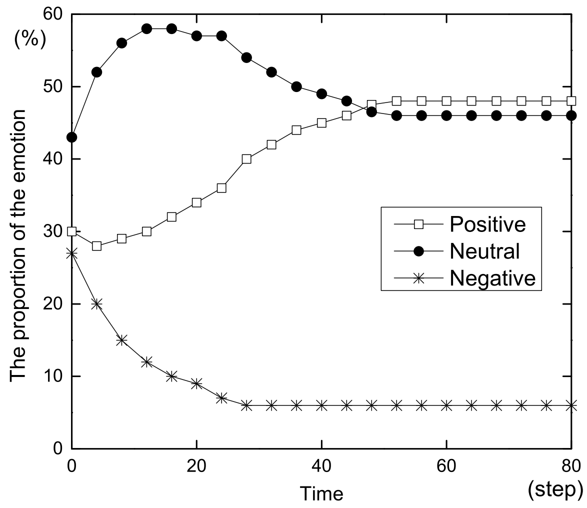


Fig. 7. The evolution process of the model at each step.

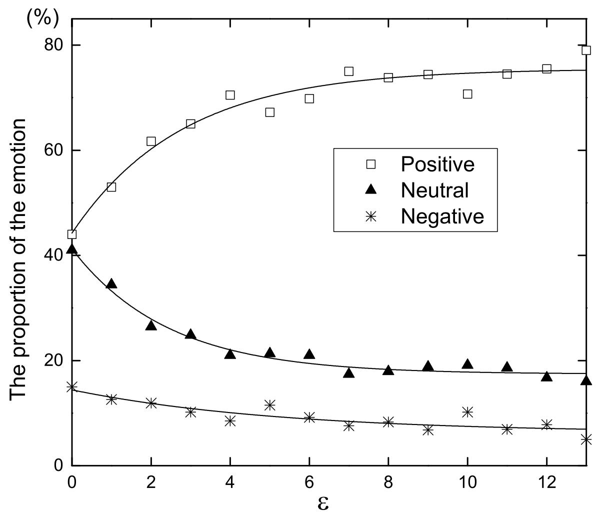


Fig. 8. The final emotion distributions with different ε.

5.4. The influence of the heterogeneity

The heterogeneity of networks is indicated by transmissibilities, a parameter formalized in Definition 7 and represented by ε. It characterizes the emotion spreading between any two vertices, and influences the evolution process and the final emotional distribution of the network. Fig. 8 shows the final emotion distributions with different values of ε, which should be adapted to these three layers. It is straightforward that when ε becomes larger, the proportion of the positive vertices is increased, and the proportion of the neutral ones is decreased. In any case, the negative vertices disappear gradually, only left the defective ones. It comes to a conclusion that the heterogeneity only affects the proportions of the positive and neutral vertices, not that of the negative ones.

To further explore the factors affecting the final emotional distribution, we take vertices' degrees, initial emotion and the transmissibility into consideration. Fig. 9 shows the average emotion of the network under different conditions. The x-axis represents the average value of each vertex's emotion multiplying its degree. It is revealed from Fig. 9 that for a large ε, the emotion is easy to be spread. When most of vertices, especially the hub vertices, are neutral at the initial moment, neither the positive nor negative emotion can spread in the network, as shown in the central area. The reason is that users tend to imitate the emotion of community leaders, not common users.

We use L_{in} to represent the number of the links between two users in the same community and L_{ext} to represent the number of the links across two communities. Thus L_{in}/L_{ext} can be used to measure the closeness and the connectivity in a

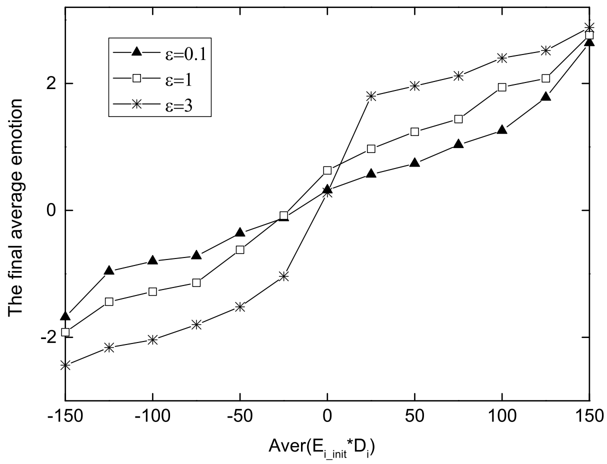


Fig. 9. The relation between the final average emotion, the degrees and the initial average emotion under the different ε.

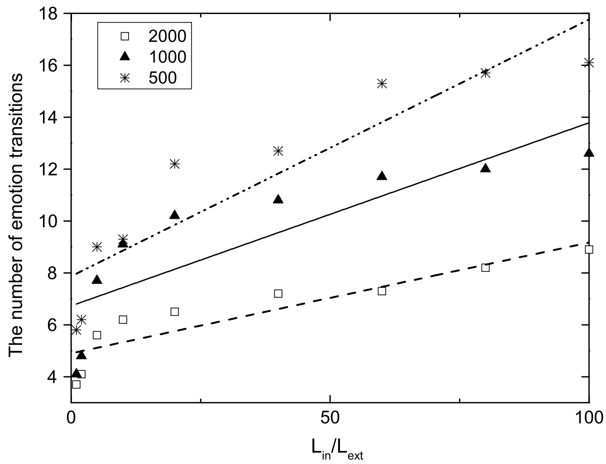


Fig. 10. The number of individual emotional tendency changes as a function of L_{in}/L_{ext}. The two variables have an approximately linear relation.

community like modularity [37]. L_{in}/L_{ext} is an optimal choice for its simple definition rather than modularity. Fig. 10 shows the number of individual emotional tendency changes as a function of L_{in}/L_{ext}. These two variables have an approximately linear relation. When more users are involved in the model, the complexity of the network is increased. A user is influenced by more friends, and is less likely to change his/her emotional tendency. When the number of users is a specific value, the emotional tendency is more likely to change with the increased L_{in}/L_{ext}. The user in the same community can easily influence his/her friends because of the high connectivity.

The extreme cases are also taken into consideration. When the L_{in}/L_{ext} is small, the values deviate from the linear tendency. In some cases, it is difficult to divide the network into multiple communities. Although the network is divided into communities by the *GN* algorithm, actually the division result may not be the optimal choice, and may be totally different in terms of different algorithms. A mature social network has an obvious community structure. Therefore, a small L_{in}/L_{ext} means users in the network are relatively inactive. They can hardly affected by others. Then we extracted a sub-network with only one community of 461 vertices, and disconnected all the edges with other communities. When L_{ext} is zero and L_{in}/L_{ext} is infinite, we found that the number of emotion tendency changes was 5, smaller than those in multi-community networks. This phenomenon can be explained as follows: if a user only communicates with the friends in the same community, his/her emotion is mainly influenced by the leader of the community and keeps stable.

Another factor related to the emotion changes is the average degree of the vertices in a community. Fig. 11 shows the relation between the number of individual emotional tendency changes and the average degree in the community on the following layer. Fewer vertices in a community often result in the frequent interactions and the frequent changes of the emotion, as shown in Fig. 10. When the number of community vertices is a specific value, the emotion changes more frequently in the beginning and then becomes less. This rising tendency roots in the users' aim to use social media. The

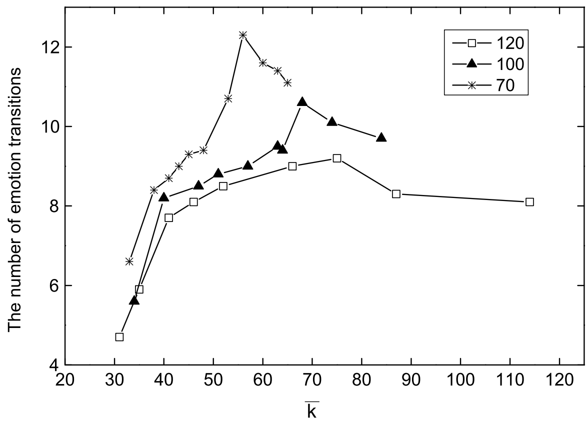


Fig. 11. The relation between the number of individual emotional tendency changes and the average degree in the community on the following layer.

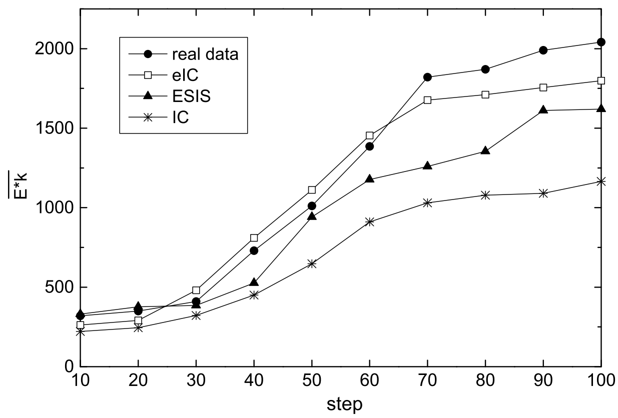


Fig. 12. The comparison of the three models and the real data. The proposed eIC model shows better effect than the ESIS model and the standard IC model.

users with fewer friends often obtain information online, and rarely express their emotion. The line descends after the peak value because the communication with the users of other communities becomes more frequent. The friends in the same community can hardly affect his/her emotion.

6. The comparison of the models

In order to demonstrate the matching effect of the proposed model with the real-world situation, we compared our model, ESIS model and the IC model with the collected real data. Because these models have different base conditions, we needed to adjust them to the same benchmark:

(1) The ESIS model categorizes emotion into fine-grained classes. We first calculated fine-grained emotion w.r.t. an individual based on the number of re-posted tweets containing this type of emotion. Then we classified emotion in the ESIS model into three types. Happiness was viewed as positive; surprise was viewed as neutral; anger, sadness, fear and disgust were all regarded as negative. Finally we added the values of all kinds of emotion in an individual's tweets to obtain his/her emotion at a specific time step.

(2) In the IC model, the edge weights were modified to represent degrees of influence among users, not only a probability of informing inactive neighbors. Thus, a vertex continued evolving even when it was affected by the emotion. Moreover, the emotional value was calculated by the same method as in the eIC model. When an individual received a tweet, the current emotional value of the individual equaled to the emotional value of the tweet plus the individual's former emotional value.

(3) The ESIS model and the IC model were considered to be appropriate to multilayer networks, whose topology on each layer was the same.

(4) A evolutionary step of these three models was fixed to two hours.

Table 5
The confusion matrix.

	P_p	N_p
P_o	TP	FN
N_o	FP	TN

6.1. Time complexity analysis

The computational complexity of the eIC is $O(m * n^2)$, where m and n denotes the total steps and the number of vertices, respectively. We can obtain the complexity of the ESIS model and the IC model in the similar way and find that all of them are the same.

6.2. Comparison for evolutionary data-fitting

As shown in Fig. 12, the distributions of the real data and the three different models are represented by different marks, respectively. The results obtained by the proposed model demonstrate better performance than the ESIS model and the standard IC model. The IC model is simple and deviates from the real data, because the effect of multilayer is not considered. The ESIS model is based on the SIS model, which is used to explain the process of epidemic or information diffusion, but some detailed features such as the three degree of separation are not involved.

The aforementioned comparison results reveal that our eIC model has good time performance comparable to those of two other models, and is also fits the real data better.

6.3. Comparison of prediction accuracy

In order to evaluate the performance of those models, we use the following common measures: *precision*, *recall* and *F-measure* [38]. In a binary classification problem, P_o and N_o represent the observed positive and negative instances; P_p and N_p represent the predicted positive and negative instances. We can represent the classification performance by a confusion matrix in Table 5, where *TP* stands for true positives, *FP* for false positives, *FN* for false negatives, and *TN* for true negatives. In this study, the positive class [39] is defined as the positive emotion, and the negative class includes the neutral and the negative emotion.

Precision, *recall* and *F-measure* are defined as follows [40]:

$$precision = \frac{TP}{TP + FP} = \frac{TP}{p_p} \tag{11}$$

$$recall = \frac{TP}{TP + FN} = \frac{TP}{p_o} \tag{12}$$

$$F\text{-}measure = \frac{(1 + \beta^2) * precision * recall}{\beta^2 (precision + recall)} \tag{13}$$

where β determines the relative weight given to precision in comparison to recall. When β=1, Eq. (13) becomes the most common form as below:

$$F\text{-}1 = \frac{2 * precision * recall}{precision + recall} \tag{14}$$

The parameter θ_1 is the boundary between the positive emotion and the neutral emotion. Table 6 shows the common measures when θ_1 has a typical value (θ_1=1.3). We can see that the eIC model has a better performance for these three measures. Fig. 13 depicts the change of F-1 with θ_1 and the number of users for the three models. The eIC model has a larger value than the other models when parameters vary. When θ_1 is near the value 1.5, The F-1 of all the three models has the largest value because the real emotion is more likely to follow the uniform distribution. A relative large or small θ_1 can easily affect the distribution of the emotion. Meanwhile, The F-1 of all the three models is increased with the number of users because a larger training set can provide a better classification performance. The eIC model describes the characteristics of emotional contagion adequately, thus it is increased more quickly than the other models.

The receiver operating characteristics (ROC) graph is a technique for visualizing, organizing and selecting classifiers based on their performance [41]. The true positive rate of a classifier is estimated as:

$$true\ positive\ rate(TPR) = \frac{TP}{TP + FP} = recall \tag{15}$$

$$false\ positive\ rate(FPR) = \frac{FP}{FP + TN} \tag{16}$$

Table 6
The common measures of the three models.

	Precision	Recall	F-1
eIC	0.839	0.762	0.798
ESIS	0.681	0.627	0.652
IC	0.644	0.568	0.604

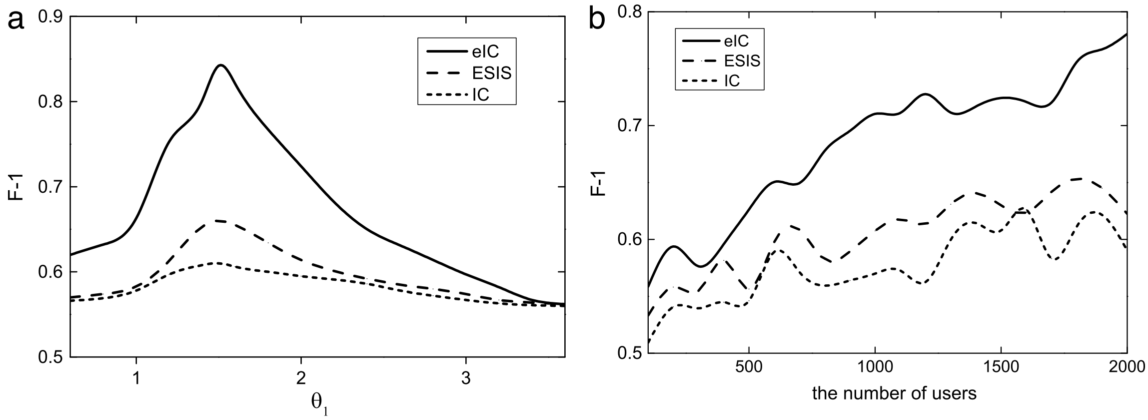


Fig. 13. F-1 changes with θ_1 and the number of users for the three models.

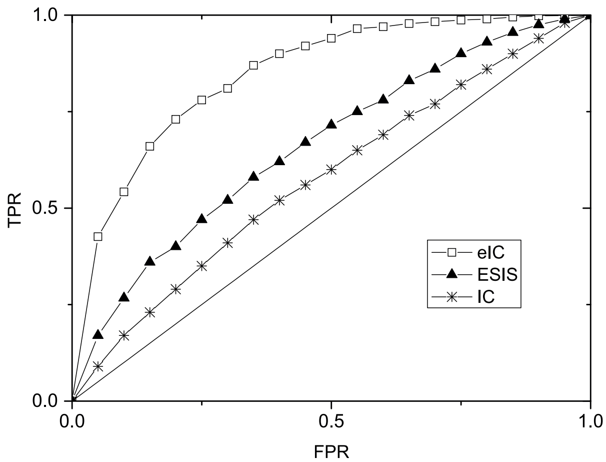


Fig. 14. The ROC curves for the three models.

The ROC graph is a two-dimensional graph in which *TPR* is plotted on the *y*-axis and *FPR* rate is plotted on the *x*-axis. A ROC curve depicts the relative trade-off between the benefit (*TPR*) and the cost (*FPR*). We use the ROC graph to compare those models after the evolution. Fig. 14 shows the ROC graph in which only the discrete values under different emotion threshold are obtained. The eIC model appears at the northwest of the graph and shows a better performance than the other models. The ESIS model and the IC model appear on the left-hand side of an ROC graph near the *x*-axis and seems to be more "conservative", which implies few false positive errors as well as low true positive rates.

7. Conclusion and future work

In this study, we firstly propose an emotional independent cascade model in which the individual emotion can affect his/her friends' emotion. The transmissibility is introduced to measure a user's ability to propagate emotion. By analyzing the patterns of emotional contagion with respect to a Twitter dataset, we find that the value of the transmissibility varies on different layers and on different community structures. Secondly, the simulation results carried out by us reveal that the polar emotion of several hub users causes the disappearance of the opposite emotion, irrelevant to the value of the transmissibility. The final emotional distribution depends on the initial emotional distribution and the transmissibilities.

Users in a small community are more likely to change the mood for the influence of community leader. Finally, we compared our model with two other models, the emotional SIS model and the standard independent cascade model. The comparison results demonstrate that the computational complexities of all the three models are similar and our model is suitable for describing real data.

Due to the observational nature of our experiment, our study is certainly not immune of possible shortcomings and these problems are required to handle in the future:

(1) Emotional contagion is a complicated process and might co-occur with other network effects such as the dynamic evolution of the network structure. Users with opposite emotion may have weak relation and disconnect the former users; users with close emotion may create new connections. Analyzing multiple effects on emotional contagion is expected to be a valuable work.
(2) Other fundamental limits arise from the current state of the art in sentimental analysis algorithms. Modern approaches, like SentiStrength here employed, although more robust and precise than ever before, still produce crude heuristics and hardly capture the many nuanced expressions that human language is able to convey. The inability to capture complex contexts triggering expressions like sarcasm or irony, the suppression of multiple emotion and the presence of ambiguity, etc., are only few examples of potential noisy outputs of such methods. The above problems are required to handle in the future.

Acknowledgments

The paper is supported by the National Natural Science Foundation of China (Grant No. 61602389, Grant No. 61602064 and Grant No. 61772091), the Scientific Research Foundation of Chengdu University of Information Technology (Grant No. KYTZ201637), the Planning Foundation for Humanities and Social Sciences of Ministry of Education of China (Grant No. 15YJAZH058), and the Youth Foundation for Humanities and Social Sciences of Ministry of Education of China (Grant No. 17YJCZH202 and Grant No. 14YJCZH046).

References

[1] L. Coviello, Y. Sohn, A.D. Kramer, C. Marlow, M. Franceschetti, N.A. Christakis, J.H. Fowler, Detecting emotional contagion in massive social networks, PLoS One 9 (3) (2014) e90315.
[2] E. Hatfreld, J.T. Cacioppo, R.L. Rapson, Emotional contagion, Curr. Dir. Psychol. Sci. 2 (3) (1993) 96–99.
[3] S.D. Pugh, Service with a smile: emotional contagion in the service encounter, Acad. Manag. J. 44 (44) (2001) 1018–1027.
[4] Q. Wang, Z. Lin, Y. Jin, S. Cheng, T. Yang, Esis: emotion-based spreaderignorantstifler model for information diffusion, Knowl.-Based Syst. 81 (2015) 46–55.
[5] A. Guille, H. Hacid, C. Favre, D.A. Zighed, Information diffusion in online social networks: a survey, ACM SIGMOD Rec. 42 (2) (2013) 17–28.
[6] J. Goldenberg, B. Libai, E. Muller, Talk of the network: a complex systems look at the underlying process of word-of-mouth, Mark. Lett. 12 (3) (2001) 211–223.
[7] W.D. Cole, H. Liu, C. Hessam, S.K.S. Candan, An Information Diffusion Approach for Detecting Emotional Contagion in Online Social Networks (Dissertations and Theses - Gradworks).
[8] M. Granovetter, Threshold models of collective behavior, Amer. J. Sociol. 83 (6) (1978) 1420–1443.
[9] S. Abdullah, X. Wu, An epidemic model for news spreading on twitter, in: IEEE International Conference on TOOLS with Artificial Intelligence, vol. 32, IEEE, 2011, pp. 163–169.
[10] X. Zhang, D. Jiang, A. Alsaedi, T. Hayat, Stationary distribution of stochastic sis epidemic model with vaccination under regime switching, Appl. Math. Lett. 59 (2016) 87–93.
[11] F. Xiong, Z.Y. Li, Effective methods of restraining diffusion in terms of epidemic dynamics, Sci. Rep. 7 (1) (2017) 6013.
[12] F. Xiong, Y. Liu, J. Cheng, Modeling and predicting opinion formation with trust propagation in online social networks, Commun. Nonlinear Sci. Numer. Simul. 44 (2017) 513–524.
[13] F. Xiong, Y. Liu, Z.J. Zhang, J. Zhu, Y. Zhang, An information diffusion model based on retweeting mechanism for online social media, Phys. Lett. A 376 (30–31) (2012) 2103–2108.
[14] J.H. Fowler, N.A. Christakis, Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the framingham heart study, BMJ 337 (49) (2008) a2338.
[15] J.N. Rosenquist, J.H. Fowler, N.A. Christakis, Social network determinants of depression, Mol. Psychiatry 16 (3) (2011) 273–281.
[16] J.T. Cacioppo, J.H. Fowler, N.A. Christakis, Alone in the crowd: the structure and spread of loneliness in a large social network, J. Pers. Soc. Psychol. 97 (6) (2009) 977–991.
[17] W.V. Brown, Framingham heart study, J. Clin. Lipidol. 5 (5) (2011) 335.
[18] M.D. Vicario, G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, W. Quattrociocchi, Echo chambers: Emotional contagion and group polarization on facebook, Sci. Rep. 6 (2016) 37825.
[19] A.D.I. Kramer, J.E. Guillory, J.T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, Proc. Natl. Acad. Sci. 111 (24) (2014) 8788–8790.
[20] E. Ferrara, Z. Yang, Measuring emotional contagion in social media, PLoS One 10 (11) (2015).
[21] X. Wang, L. Zhang, Y. Lin, Y. Zhao, X. Hu, Computational models and optimal control strategies for emotion contagion in the human population in emergencies, Knowl.-Based Syst. 109 (2016) 35–47.
[22] S. Boccaletti, G. Bianconi, R. Criado, C.I.D. Genio, J. Gómez-Gardees, M. Romance, I. Sendia-Nadal, Z. Wang, M. Zanin, The structure and dynamics of multilayer networks, Phys. Rep. 544 (1) (2014) 1–122.
[23] J. Borondo, A.J. Morales, R.M. Benito, J.C. Losada, Multiple leaders on a multilayer social media, Chaos Solitons and Fractals 72 (2015) 90–98.
[24] O. Yagan, D. Qian, J. Zhang, D. Cochran, Information diffusion in overlaying social-physical networks, Inf. Sci. Syst. 31 (6) (2012) 1038–1048.
[25] E. Ferrara, Z. Yang, Quantifying the effect of sentiment on information diffusion in social media, PeerJ Comput. Sci. 1 (51) (2015) 1e26.
[26] L. Le, E. Ferrara, A. Flammini, On predictability of rare events leveraging social media: A machine learning perspective, 2015.

[27] J. Wiebe, Subjectivity word sense disambiguation, in: Workshop in Computational Approaches To Subjectivity and Sentiment Analysis, 2012, pp. 2–2.
[28] C.J. Hutto, Vader: A parsimonious rule-based model for sentiment analysis of social media text, 2014.
[29] M. Thelwall, K. Buckley, G. Paltoglou, Sentiment strength detection for the social web, J. Assoc. Inf. Sci. Technol. 63 (1) (2012) 163–173.
[30] M. Thelwall, K. Buckley, Topic-based sentiment analysis for the social web: The role of mood and issue-related words, J. Assoc. Inf. Sci. Technol. 64 (8) (2013) 1608–1617.
[31] H. Kwak, C. Lee, H. Park, S. Moon, What is twitter a social network or a news media?, in: International Conference on World Wide Web, 2010, pp. 591–600.
[32] D. Boyd, S. Golder, G. Lotan, Tweet, tweet, retweet: Conversational aspects of retweeting on twitter, in: Hawaii International Conference on System Sciences, 2010, pp. 1–10.
[33] M.E. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E 69 (2 Pt 2) (2004) 026113.
[34] X. Xiong, Y. Hu, Research on the dynamics of opinion spread based on social network services, Acta Phys. Sin. 61 (15) (2012) 150509.
[35] M. Mcpherson, L. Smithlovin, J.M. Cook, Birds of a feather: Homophily in social networks, Sociology 27 (27) (2001) 415–444.
[36] A.D.I. Kramer, The spread of emotion via facebook, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2012, pp. 767–770.
[37] L.N. Ferreira, L. Zhao, Time series clustering via community detection in networks, Inform. Sci. 326 (C) (2016) 227–242.
[38] M. Bogaert, M. Ballings, D.V.D. Poel, Evaluating the importance of different communication types in romantic tie prediction on social media, Ann. Oper. Res. (2016) 1–27.
[39] Hern, J. Ndez-Orallo, P. Flach, C. Ferri, A unified view of performance metrics: translating threshold choice into expected classification loss, J. Mach. Learn. Res. 13 (1) (2012) 2813–2869.
[40] J.F. Diez-Pastor, J.J. Rodriguez, C. Garcia-Osorio, L.I. Kuncheva, Random balance: Ensembles of variable priors classifiers for imbalanced data, Knowl.-Based Syst. 85 (2015) 96–111.
[41] T. Fawcett, The introduction to roc analysis, Pattern Recognit. Lett. 27 (8) (2006) 861–874.