The Journal of Supercomputing (2023) 79:7717–7735
https://doi.org/10.1007/s11227-022-04968-5

Research on information dissemination model based on heat transfer in online social network

Chen Jing[1,4] · Huang Jincheng[1] · Xin Chen[1,2,3] · Liu Mingxin[4]

Accepted: 18 November 2022 / Published online: 6 December 2022

Abstract
At present, the information dissemination model based on the structure modeling of online social network can no longer simulate the increasingly complex information dissemination process. Therefore, a heat transfer (HT) information dissemination model based on heat transfer idea is proposed in this paper. Firstly, based on the physical characteristics of heat transfer due to temperature difference between objects, the model introduces the temperature value attribute of nodes and defines the dissemination node and observer node. Secondly, the nodes in the online social network were described as heat receivers and dispersers. The diffusion heat and the receiving heat process of information dissemination were defined, and a two-stage greedy diffusion (TGD) algorithm based on HT model was proposed for information dissemination in online social network. Finally, the HT model is verified by experiments in real data sets. The experimental results show that the HT model proposed has high accuracy in the process of information dissemination of social networks of different scales and has great advantages in the prediction of dissemination trend and node variation trend.

✉ Xin Chen
dddream@stumail.ysu.edu.cn

✉ Liu Mingxin
mingxinliu001@sina.com

1 School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, Hebei, China

2 Hebei Key Laboratory of Virtual Technology and System Integration, Yanshan University, Qinhuangdao 066004, Hebei, China

3 Hebei Key Laboratory of Software Engineering, Yanshan University, Qinhuangdao 066004, Hebei, China

4 College of Electronic and Information Engineering, Guangdong Ocean University, Zhanjiang 524088, Guangdong, China

Keywords Online social network · Temperature value attribute · Heat transfer · Information dissemination model

1 Introduction

With the explosive growth of the number of network users, online social network has become the main medium of information dissemination. Users can easily exchange information and share ideas with each other by online social networks. Because online social network is characterized by fast information dissemination and independent of geographical influence, its dissemination process can affect the social, political, economic, and cultural fields of modern life. An in-depth understanding of the process of information dissemination and the application of dissemination mechanism can effectively have a positive impact on the society.

Most information dissemination models are improved based on the SIR (Susceptible Infected Recovered) model [1], and other models such as IBU (Ignorant Believed and Unbelieved) model [2] and SEIR (Susceptible Exposed Infectious and Recovered) model [3] are derived from them by discussing the precise division of the three types of nodes in the SIR model. In addition, some information dissemination models study the interaction between network structure and users to simulate the process of information diffusion [4, 5]. The research shows that the traditional information dissemination model cannot meet the needs of the increasingly complex information dissemination process [6]. For this reason, in recent years, researchers have begun to systematically study the effects of information dissemination over time using mathematical models [7] or physical phenomena [8]. Simulation of information dissemination on online social networks based on physical [9] or natural phenomena [10] to reflect how information dissemination in real life is transformed into a computer-based modeling process.

In order to simulate the increasingly complex information dissemination process online social networks, the relationship between the heat transfer process and the information dissemination process is simulated and verified by means of the research results of physics principles in complex networks, and the application scope of the information dissemination model is extended. The contributions of this paper are as follows:

(1) The phenomenon of heat transfer due to temperature difference between objects is applied to the information dissemination process between nodes. The behavior of nodes in social network is analyzed, and the characteristic attributes of nodes are introduced to comprehensively discuss the process of diffusion heat and receiving heat of information dissemination.
(2) By means of the idea of heat transfer process in physics and Ebbinghaus Forgetting Curve, a HT information dissemination model based on heat transfer process is proposed, which fully considers the problem of information dissemination in online social networks.

(3) The experiment proves that HT information dissemination model can adapt to different scales of online social networks and conforms to the law of information dissemination of real online social networks.

The structure of this paper is as follows: In Sect. 2, the related work is analyzed and discussed; In Sect. 3, the HT information dissemination model is described, its mathematical representation is defined and information dissemination process is explained based on heat transfer; In Sect. 4, the algorithm of HT model is presented; In Sect. 5, the model is verified on real data set, and compared with SI model and Particle model; Finally, the thesis is summarized and the future work is discussed.

2 Related work

Information dissemination belongs to the research area of online social network analysis and has become a focus of research in recent years because of its more practical implications in advertising and marketing, rumor dissemination, and opinion management. At present, most information transmission models are the improved models of SIR, IC (Independent Cascading Model) [11], and LT (Linear Threshold Model) [12].

One of the very classical research perspectives is the discussion of the division of the nodes of the SIR model, deriving a very large number of models. Fan et al. [2] proposed the IBU model, which studied the influence of the authenticity of the information content and the credibility of users on the dynamic process of information transmission. Wei et al. [13] introduced the super communicator in disease transmission and established the SIER (Super-spreader Infectious Exposed and Recovered) transmission model based on the traditional SIR model. Zhang et al. [14] studied the characteristics of rumor dissemination and established a SITR (Susceptible Infective True Removed) rumor dissemination model in consideration of the factors that users forgot about rumors, then discussed the equilibrium point and local stability under various circumstances. Wang et al. [15] added the roles of observer and rumor debunker based on the classic SIR model to reduce the explicit and implicit negative effects of rumor dissemination on the society. Fu et al. [16] studied competitive dissemination among groups on social networks and divided the crowd into three groups. They found that the innovator group can improve the range of information dissemination, and the convergence of the network depends more on the topology of the network. Sun et al. [17] added uncertainty theory to the SIR model, gave the derivation process of the model, and proved that the uncertain SIR model is measured stable under certain conditions. This kind of model is derived from the network group behavior, or individuals are classified into different node types for discussion, so it does not consider the impact of individuals and network structure on information transmission.

In addition, it is also a research perspective to study the laws of information dissemination in terms of the user interaction and network structure. Zhang et al. [18] established the SETQR (Susceptible Exposed Trusted Questioned Recovered) model, derived the law of information transmission by using the probability theorem, and further obtained the equilibrium point and basic reproduction number of the SETQR model. Mathew et al. [19] analyzed and discussed the process of hate speech dissemination and found that the relationship between hate users is stronger and hate speech spreads more rapidly compared to ordinary speech. Yin et al. [4] considered the emotional factor in the process of information dissemination, and for this purpose, they developed an MNE-SFI (multiple negative emotional susceptible forwarding immune) model to investigate the important role of users' subjective emotions on information dissemination. Hui et al. [20] considered the spreading scale of rumors from the perspective of both registration times and education level and concluded that the education level of the population is an important factor affecting the final spread scale of rumors. Nian et al. [5] developed a new model HFS (Human Flesh Search) and analytically demonstrated the effect of dissemination effects in dissemination and the effectiveness of the new model. The kind of information dissemination model is based on individuals or network structures. However, these models contain more artificial assumptions that cannot be adapted to social networks with diverse and rapidly updating structures.

In recent years, many scholars have begun to introduce knowledge from physics, mathematics, biology, and other disciplines into information dissemination models and to extend their research. Chai et al. [7]. propose a stochastic perturbation model that considers both population perturbation and connectivity variation and demonstrates that both perturbations have a large impact on information dissemination. Ulyana et al. [10]. compared the main data of microbial process with the information diffusion process in social network based on the conceptual model of Monod equation, and elaborated the main factors affecting the speed of information transmission. Wang et al. [21]. used the mean-field theory to derive the information diffusion system and established a new model of information dissemination called in-out-Unacquired-Acquired-Rejected to demonstrate the enhanced effect of mobile devices on information dissemination. Xia et al. [9] proposed an information diffusion model based on particle dynamics, and verified the information transmission process with real data sets. Liu [22] analyzed the information communication of public emergencies, and constructed the mathematical model of information communication of public emergencies. Zhang et al. [23] proposed an information transmission model based on individual behavior by combining random time generation rules and an information visibility prediction method. Chou et al. [24] proposed an information transmission model based on multiple factors, which could adapt to more communication scenarios under different factors. Yin et al. [25]. improved the traditional SIR model from a kinetic perspective by introducing an "uncertain" mental state that allows nodes to repeatedly spread infections, providing new insights into information dissemination. Jiang et al. [8] proposed the OBM (Operator-Based Model) model based on the thermal diffusion model and the agent model to improve the performance of simulated dissemination.

It can be found from the above research that: (1) The traditional modeling based on SIR model requires certain artificial assumptions, which is difficult to adapt to the complex and diverse online social networks. (2) Some scholars have introduced the knowledge of other disciplines into social networks and extended the research on information transmission models, but most research method is still in the exploratory stage.

In online social networks, information itself will gradually lose attention over time, and the influence of information will also decrease. This is similar phenomena to how heat is lost over time in physics. Therefore, in the process of dissemination, the heat transfer equation in physics is introduced to describe the information dissemination process, and the information dissemination model HT is proposed in this paper.

3 Description and mathematical representation of HT information dissemination model

Heat transfer is a physical phenomenon. In essence, as long as there is a temperature difference in an object or between multiple objects, heat will spontaneously transfer from high temperature to low temperature. This process is called heat transfer. Inspired by this physical phenomenon and combined with the information transmission law of online social networks, a HT dissemination information model based on heat transfer is proposed in this paper.

The detailed descriptions of the notations could be found in Table 1

3.1 Overview of HT model

In the real world, as long as there is a temperature difference between or within an object, heat must flow from a hotter place to a cooler place. In terms of heat transfer, an object starts at any heat source and gradually diffuses to all parts of the object. The process of heat dissemination is similar to the process of information dissemination in online social networks. Therefore, the information transmission in online

Table 1 Notations used in this paper

Notions	Descriptions
T_i^t	Temperature of node i at time t
T_c	The room temperature constant, and is set to zero
N	The comment time, day or hour
c_{time}	The comment time, day or hour
k	The cooling coefficient
S^{date}	The collection of comment dates
S^{out}	The out-degree neighbor node set of the Node v_i
α	The diffusion coefficient
DH(v_i)	The temperature change caused by the total heat diffusion
RH(v_i)	All of the receiving heat in the node v_i
N	The collection of diffuse nodes
D	The number of diffusion nodes set

social networks of different application scales with different parameter settings is simulated based on the heat transfer process. HT model is described as follows:

At initial time, given two adjacent objects A and B, the temperature of A is T_1, and the temperature of B is zero. Supposed that the two objects transfer heat to each other and reach an equilibrium temperature T_2. The relationship between temperature and heat in the process of heat transfer is shown in formula (1).

$$\begin{cases} Q_{endothermic} = Q_{exothermic} \\ Q = cm\Delta t \end{cases} \tag{1}$$

where Q is the heat, c is the specific heat capacity, m is the mass of the object, and Δt is the temperature change value.

In combination with formula (1), assuming that the specific heat capacities of A and B are the same value c, the heat transfer progress can be expressed as

$$cm_A(T_2 - T_1) = -cm_B T_2 \tag{2}$$

where m_A and m_B are the masses of the object, respectively. The equilibrium temperatures of A and B are expressed as T_2 can be obtained by deducing formula (2), namely

$$T_2 = \frac{m_A}{m_A + m_B} T_1 \tag{3}$$

The result of heat transfer between objects can be expressed by the change of temperature according to formula (3). Therefore, the process of heat transfer can be reflected by the change of temperature in the model. For example, heat transfer between A and B can be expressed as a change in the temperature value of the object. When the specific heat capacity c is the same value, the temperature change is related to the mass difference between objects. For the whole information transmission process, as long as the node has a temperature difference, it inevitably carries out the process of heat transfer, and the quality of the node has little influence on this process. In order to describe the mass relationship between nodes in the network and simplify the subsequent calculation process in the HT model, the following definitions is described:

Definition 1 (Node mass relation) In an online social network G, its node set is $V=\{v_1,v_2,...,v_n\}$, where the mass of each node is, respectively, set as $m_1,m_2,...,m_n$, then the mass relation between nodes is$m_1 = m_2 = ... = m_n$.

Considering that the neighbor nodes of most nodes are not unique in online social network, the information transmission process is not limited to two nodes, that is, HT model is carried out between multiple nodes. If node v affects n neighbor nodes, heat transfer will occur between the $n+1$ nodes. Combined with Definition 1 and formula (3), the equilibrium temperature between nodes is defined as:

$$T_2 = \frac{1}{n} T_1 \tag{4}$$

In the real social network, not all nodes are active nodes. In this paper, the active nodes are nodes that are in the online state and have the ability to transmit or receive information. For example, in microblog, users can browse, post or receive information when they are online, while offline users cannot participate in these processes. The online time of a node reflects its activity to some extent, which is directly related to the ability of a node to transmit and receive information. If the node is offline, it cannot enter the information transmission process. In order to simulate the node individuals in the real social network, the online characteristic attributes of social network nodes are introduced, that is, the simulated node individuals are not necessarily online all the time.

In this paper, a day is divided into 24 periods, and the number of comments in each period is set as $c_{hour}(hour = 1, 2, 3, ..., 24)$, and the total number of comments in a day is set as $c_{day}\left(day \in S^{date}\right)$, S^{date} is the collection of comment dates. The online probability of node v_i in 24 time periods of a given day is shown as:

$$p\left(v_i, day\right) = \frac{c_{hour}}{c_{day}} \tag{5}$$

where c_{hour} and c_{day} must belong to the same dataset. The online probability of each day is obtained, and the average value is the average online probability of nodes v_i. The probability reflects the situation that all nodes in the online social network are online, and is shown as in formula (6).

$$p(v_i) = \frac{\sum_{day=1}^{days} p(v_i, day)}{days} \tag{6}$$

In the whole process of information transmission, there are two kinds of nodes in the network: (1) Diffuser with temperature (Diffuser), which represent the ability of information transmission and spread information to other nodes; (2) The temperature-free Watcher (Watcher) represents that it has no ability to transmit information and only passively receive information. When the information is transmitted in the network, the node carrying the information has an initial temperature value, and the node acts as the heat source. Because there is a temperature difference between the heat source node and the neighbor node, heat transfer is inevitable. In the online social network, during the process of heat transfer, the temperature value of the neighbor nodes change until there is no node with a temperature value of zero, and end of the transmission process.

3.2 Mathematical representation of HT model

The mathematical presentation of HT models and parameter settings in online social networks are defined in this section. HT model takes the nodes as heat disseminators or receivers, the initial information carrying nodes as heat sources, and the edges between nodes as heat transfer relations to transmit information. HT model divided the information transmission process into two steps : heat diffusion and heat reception. It can be seen from the above that the heat change caused by heat transfer between nodes, which is represented by the temperature change. Therefore, the calculation of temperature value of node attributes is involved in HT model.

3.2.1 Diffusion heat

In this step, the initial transmission node carrying information is regarded as a heat source to disseminate information to the surrounding neighbor nodes. You can think of it as a heat source transfers heat around. Newton's cooling law represents the process of heat transfer from the heat source to the surrounding medium, which is applied to the node diffusion heat. Therefore, Newton's cooling law is shown as follow:

$$\frac{dT}{dt} = -k\left(T - T_c\right) \tag{7}$$

where T is the function of node temperature changing with time, t is time, k is the cooling coefficient, T_c is the room temperature constant, and is set to zero. Formula (8) is obtained from formula (7).

$$T = T_0 e^{-kt} \tag{8}$$

where T_0 is the node temperature at the time $t = 0$, from which the temperature change function with time is obtained. The cooling coefficient k decreases gradually as the temperature approaches the indoor temperature constant according to Newton's cooling law, so the cooling coefficient is a function $k(t)$ that changes with time. In the process of information transmission, after the information is received by the individual, the forgetting begins, that is, the diffused heat corresponds to the state where the individual gradually forgets the information. The Ebbinghaus Forgetting Curve function is introduced here, which describes the law of human brain's forgetting of new things, and is expressed as follows:

$$k(t) = 1 - 0.56t^{0.06} \tag{9}$$

Formula (7) and (8) describe Newton's cooling law and Formula (9) is used to replace the cooling coefficient K in the law of cooling. Make it closer to the process of information dissemination in online social networks. Since temperature change is involved, a temperature value attribute should be added to reflect the nodes temperature change in the process of information transmission. In HT model, the process of diffusion heat is defined as follows by combining with the relative description.

Definition 2 (Diffused heat) Consider a graph of a social network $G = (V, E)$. V is the node set, $V = \{v_1, v_2, v_3, ..., v_n\}$, E is an edge set, $E = \{(v_i, v_j) \mid v_i, v_j \in V\}$. At time t, the initial temperature of the node v_i is set as T_i^{inital} and the node v_i has an out-degree neighbor node set S^{out}. Node v_i will transfer heat to their out-degree neighbor nodes $v_j \in S^{out}$ in time intervals Δt, resulting in a temperature change of $DH(v_i)$.

HT model should adopt the following rules when diffusing heat: (1) The temperature of node v_i is higher than that of neighboring node v_j; (2) Heat flows from node v_i to node v_j; (3) Nodes in the online state. It can be known that the temperature change is defined as follows according to the Newton's cooling law:

$$DH(v_i) = T_i^{inital}(1 - e^{-k(t)t}) \tag{10}$$

The description of $k(t)$ is shown in formula (9).

$DH(v_i)$ is the temperature change caused by the total heat diffusion. After heat transfer, the changed temperature value of a single node in the neighborhood is $\alpha DH(v_i)$ according to formula (4). where α is the diffusion coefficient, and its value is shown in formula (11).

$$\alpha = \begin{cases} 0, & S^{out} = \phi, no \quad neighbors \quad can \quad spread \\ 1/n, & S^{out} \neq \phi, n \quad is \quad the \quad number \quad of \\ & out - degree \quad neighbor \quad nodes \end{cases} \tag{11}$$

3.2.2 Receive heat

In this step, the node that does not receive information is the spectator node, which corresponds to the medium around the heat source. It receive information from the dissemination nodes, that is, the node receive heat from the heat sources. When its own temperature value changes, it means that the heat transfer process is over, that is, the node receives information and changes from the spectator node to the propagating node. The heat receiving process of HT model is defined as follows by combining with the social network graph $G = (V, E)$ in Definition 2.

Definition 3 (Receive heat) Given the current temperature of the node v_i be $T_i^{current}$, and the node v_i has in-degree neighbor node set S^{in}. Node v_i receive heat from their in-degree neighbors sets $v_i \epsilon S^{in}$ over a period of time Δt, and the change in temperature caused by receiving heat $RH(v_i)$ is added to the current temperature $T_i^{current}$. HT model should follow the rules when receiving heat:(1) The temperature of node v_i is lower than that of neighboring node v_j; (2) Heat flows from node v_j to node v_i; (3) Nodes are online state. From the above, it can be seen that:

$$RH(v_i) = \sum_{v_i \epsilon S^{in}} \alpha_j DH(v_j) \tag{12}$$

where α_j is the diffusion coefficient corresponding to in-degree neighboring node v_j. The current node temperature becomes $T_i^{current} + RH(v_i)$.

In conclusion, in the process of information transmission at time t, the temperature change of node v_i is T_i^{change}, it is shown in formula (13).

$$T_i^{change} = -DH(v_i) + RH(v_j) \tag{13}$$

In this section, the mathematical representation of the two steps is given in the HT model. Diffused heat describes the process of information transmission from propagating nodes to spectator nodes in HT model. Every time heat transfer occurs between propagating nodes and spectator nodes, information is transferred between nodes. Receiving heat describes the process of the spectator node receiving information from the propagating node in the HT model. The spectator node receives heat from the propagating node and changes its own temperature value, corresponding to the node receiving information and transforming itself into the propagating node.

4 TGD algorithm

4.1 Algorithm description

To clearly describe the information transmission process of HT model, an online social network information transmission algorithm TGD based on HT model is proposed. The basic idea of the algorithm is as follows: firstly, select the seed node carrying information and initialize the temperature value of the seed node. The total degree of the current seed node (the sum of the in-degree and out-degree) is taken as the initial temperature value. Then, the seed nodes carry out heat transfer and influence the surrounding neighbor nodes.

The algorithm is divided into two stages: heat diffusion and heat reception. Heat diffusion needs to traverse all the transmission nodes and calculate the temperature change value generated by the heat diffused at each dissemination node. When receiving heat, all nodes need to be traversed to determine whether the nodes are affected. The algorithm is described as follows:

Algorithm 1 TGD algorithm

```
Require: Social network graph G(V, E), the collection of diffuse nodes N
Ensure: The number of diffusion nodes set D
1: Initialize node attributes in the graph G and assign initial temperature to
   seed nodes
2: N ← seed Nodes D ← φ
3: while N ≠ 0 do
4:     D ← D ⋃ (Number of nodes in n)
5:     for all v in N do
6:         for all v^out in S^out do
7:             i ← 1
8:             // Count the number of qualified nodes
9:             if v_out is online & meets the receiving heat conditions then
10:                i ← i + 1
11:            end if
12:            //Compute the α by number of i
13:            if i ≠ 0 then
14:                α ← 1/i
15:            else
16:                α ← 0
17:            end if
18:        end for
19:        //Compute the DH(v) of V by formula (10)
20:        if α ≠ 0 then
21:            for all v^out in S^out do
22:                //Upate the Node status
23:                transfer αDH(v) to v^out
24:            end for
25:        end if
26:        Compute RH(v) by formula (12)
27:    end for
28:    Sort out the temperature of each node in the Graph G
29: end while
30: return D
```

4.2 Algorithm complexity analysis

The time complexity of the TGD algorithm is analyzed as follows. Let the total number of nodes in the social network $G(V, E)$ be n, the number of nodes in the set N of record spreading nodes be m, and the number of neighboring nodes owned by the nodes in the set N be v. In Algorithm 1, the line 1 initializes the node attributes in the social network, which requires traversing the $G(V, E)$ of n nodes, which has a time complexity of $O(n)$. From lines 6 to 25, there are two parallel for loops, both of which have a time complexity of $O(v)$, however, the if condition in line 20 changes the overall number of runs of the second for loop. If the number of loops in the for loop from line 5 to line 11 is m, the time complexity is $O(m)$. Let the overall run count of the second for loop be $m'(m' \leq m)$ and time complexity of $O(m')$, so the time complexity from lines 5 to 25 is $O(mv + m'v)$. The while loop in line 3 controls the number of disseminations in the overall dissemination process, and let the number be $k(k > 0 andkismuchsmallerthann)$, then the time complexity of lines 3 to 28 is $O(k(mv + m'v + n))$. In summary, the time complexity of the TGD algorithm is $O(n) + O(k(mv + mv + n))$.

5 Experiment and result analysis

5.1 Experimental data set and numerical simulation

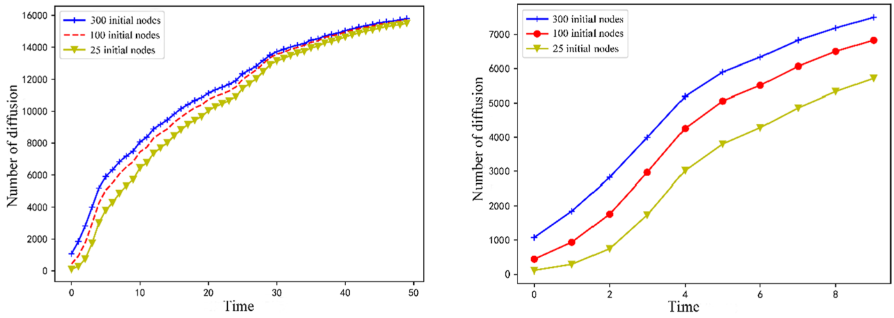
Two data sets rt-pol [26] and rt-bahrain [26] are selected and taken from a web site, http://networkrepository.com/index.php, which describes the process of information transmission in the real social network environment. Data sets are the forwarding network data collected from various social and political topic tags on Twitter, and the data format is (USER_A, USER_B, TIME_STAMP). Each data represents user USER_A to influence user USER_B at the time TIME_STAMP. The Dataset description is shown in Table 2.

Using the network structure of rt-pol and rt-bahrain, different initial influence nodes were selected as heat sources to simulate the process of information transmission, and the influence on HT information transmission model was discussed. Figures 1 and 2 show HT information dissemination model curves with different initial nodes. The data set rt-pol is used in Fig. 1, and the data set rt-bahrain is used in Fig. 2. The two figures are composed of the overall data curve of Fig. (a) and the data curve of the first ten rounds of Fig. (b). And, Fig. (b) is used to explain the law of information transmission in the early stage of Fig. (a).

Figures 1b and 2b are shown: When the initial impact nodes are different, the early stage of dissemination process differ greatly. The more initial influence nodes, the faster the iteration of information dissemination process. When it reaches a certain number, fewer and fewer nodes can be propagated in the network, the increase process of propagating nodes will be slower and slower, and it will tend to converge to a stable value, as shown in the latter part of Figs. 1a and 2a. Therefore, it can be seen that the number of initial influencing nodes is related

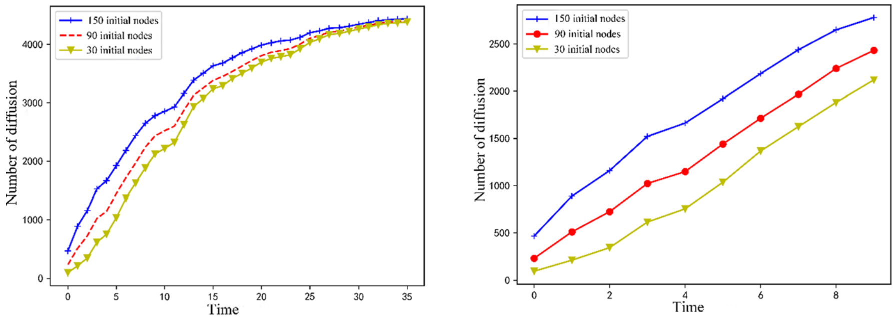
Table 2 Dataset description

Data set	Number of nodes	Number of edges	Average clustering coefficient	The time span
rt-pol	18470	48365	0.059	50 days
rt-bahrain	4676	8007	0.017	36 hours



(a) Overall data **(b)** The first ten rounds of data curve

Fig. 1 Information dissemination process curve on rt-pol dataset



(a) Overall data **(b)** The first ten rounds of data curve

Fig. 2 Information dissemination curve on rt-bahrain dataset

to the convergence of HT information dissemination model. The larger the number of initial influencing nodes is, the faster the convergence reaches.

5.2 Experimental comparison

5.2.1 Measurement indicators and comparison models

In this paper, the dissemination scale is used as an indicator to compare the advantages and disadvantages of the algorithm, and the same initial impact nodes are selected to experience the same time step to derive the experimental results. In order to measure the accuracy of the model, three indicators are selected to measure the error between the predicted value and the true value, including accuracy, R2 coefficient (R2_score) and average absolute percentage error (MAPE).

The formula of Accuracy is as follows:

$$accuracy = 1 - \frac{actualvalue - predicate \quad value}{actual \quad value} \tag{14}$$

SI model [1]: one of the classic infectious disease models. The nodes in the social network are divided into two categories: susceptible to infection and infected. Select the initial influencing nodes, and these nodes will affect the surrounding neighbor nodes with a certain infection rate. The infected node is regarded as the node receiving information and also has the ability to transmit information. The dissemination process ends when all nodes in the social network are transformed into infected nodes or a finite step is reached.

Particle model [9]: information dissemination model based on particle elastic collision. Inspired by the physical phenomenon of particle collision, the model takes the collision between particles as the information dissemination behavior between nodes, and defines three types of node states: dissemination state (active and capable of dissemination), contact state (active but not capable of dissemination) and observation state (inactive). Through the "collision" of the initial influence node to the neighbor node, the information dissemination process is simulated, and the observation state node after the collision will change into the observation state or the contact state node.

5.2.2 Analysis of experimental result

The relative error fitting of different information dissemination models is shown in Table 3.

Based on the dataset rt-pol, the accuracy of the HT model is 95.32%, which is 3.07% higher than Particle model and 6.57% higher than SI model. The MAPE of the HT model is only 4.68%, which is exceeded by the Particle model by 3.07%, and by the SI model by 6.57%. The R2_score of the HT model is 0.9642, which is better than Particle model and Si model. Based on the dataset rt-bahrain, the accuracy of the HT model is 95.41%, which is 2.77% higher than that of the Particle model and 3.25% higher than that of the SI model; the MAPE of the HT model is only 4.59%, which is exceeded by the Particle model by 2.76% and by the SI model by 3.26%; in terms of R2_score, the HT model is also better than both the Particle model and the SI model. It can be seen from the above that in the data sets rt-pol and rt-bahrain, the experimental results of HT model are the best among the three measurement indexes, followed by Particle model, and SI model is the worst.

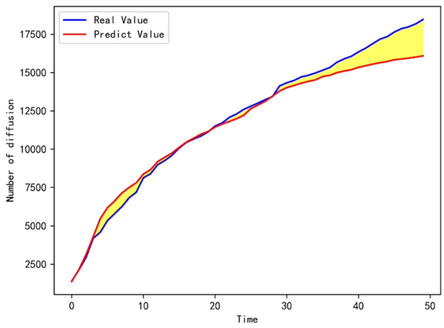
Table 3 Experimental results

Dissemination model	accuracy		MAPE		R2_score	
	rt-pol	rt-bahrain	rt-pol	rt-bahrain	rt-pol	rt-bahrain
SI model	88.75%	92.16%	11.25%	7.85%	0.927	0.9572
Particle model	92.25%	92.64%	7.75%	7.35%	0.916	0.9487
HT model	95.32%	95.41%	4.68%	4.59%	0.9642	0.9822

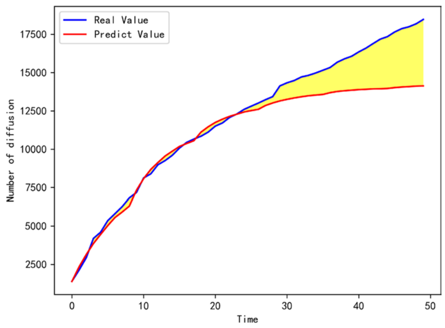
Compare the experimental results obtained for each model with the rt-pol and rt-bahrain data sets, the difference of accuracy index of the SI model was 3.41%, and the difference of MAPE index was 3.4%. The difference value of accuracy index of Particle model is 0.39%, and the difference value of MAPE index is 0.4%. The difference value of accuracy index of HT model is 0.09%, and that of MAPE index is 0.09%. According to the above data, the difference between the variation data of SI model is 9 times that of Particle model and 37 times that of HT model. In terms of data, the data obtained from HT model and Particle model experiments are relatively stable, which have good adaptability to data sets of different scales and will not affect the results too much to produce errors. And the stability and adaptability of HT model are better than Particle.

In order to better verify the dissemination process of the three models on different data sets, the three models were separately represented on different data sets, and the proximity between the real data of the data set and the predicted data of the model was represented by the filled area. The smaller the filling area is, the closer the information dissemination process simulated by the model is to the real situation.

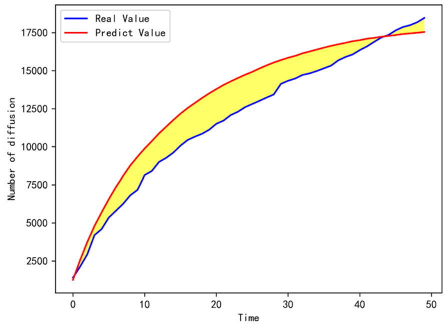
Figure 3 shows the dissemination curves of the three models on the data set rt-pol. It can be seen from the above figure that, in terms of the filled area, the SI



(a) HT model curve



(b) Particle model curve



(c) SI model curve

Fig. 3 Three model curve in rt-pol

model has the largest area, followed by the Particle model, and the HT model has the smallest area in the early information dissemination process. It shows that the HT model is closest to the real situation on rt-pol data set. However, in terms of the nodes affected by the model, the SI model affected the most nodes. HT model was close to the SI model, but there was still a certain distance, and the Particle model had a large gap.

Figure 4 shows the dissemination curves of the three models in the rt-barain data set. It can be seen that the growth trend of node dissemination number of the three models are similar. Before the 12th round of dissemination, the Particle model area is the smallest and the effect is the best. As the dissemination process goes on, the dissemination effect of Particle model becomes worse and worse, and the error gradually increases. In terms of the overall filled area, it can be clearly found that the HT model and the SI model have similar areas, and the overall dissemination process is better than the Particle model, in which the HT model is slightly better than the SI model.

In combination with Figs. 3 and 4 and Table 3, it can be found that the filled area in Figs. 3 and 4 has a certain relationship with the experimental results in Table 3. In Figs. 3 and 4, the smaller the filling area, the better the data performance in Table 3.

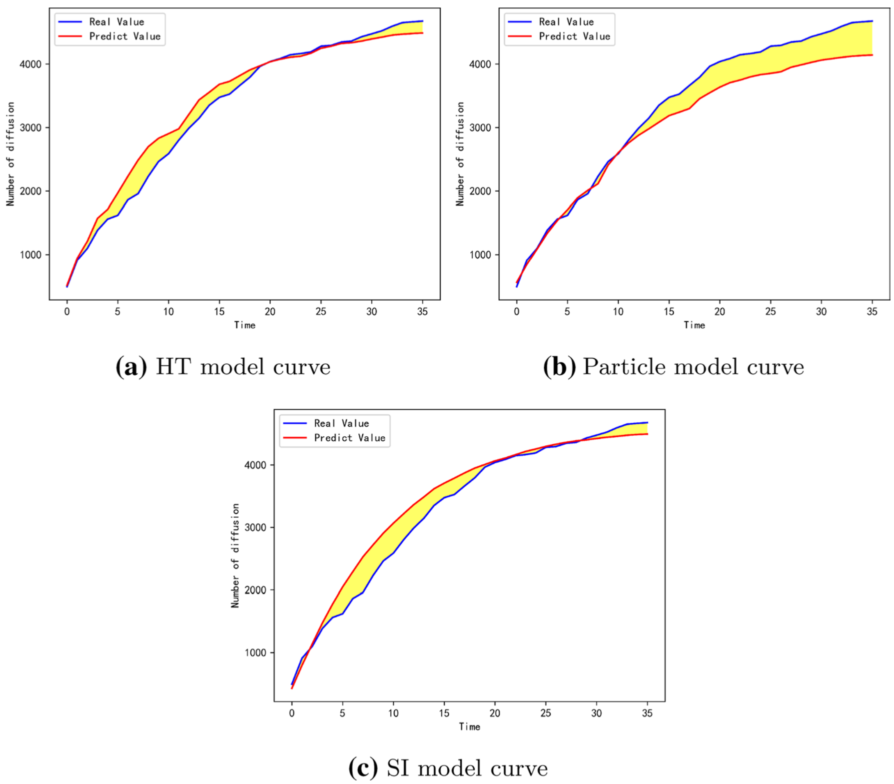


Fig. 4 Three model curve in rt-bahrain

It also explains why the performance of Particle model in Table 3 is better than that of the SI model, but the final total number of affected nodes is not as high as that of the SI model. It can be seen from Table 3 that the experimental results of HT model are better than those of the other two models. In Figs. 3 and 4, HT model has the smallest filled area, which is closest to the real information transmission process of online social network. Therefore, it can be proved that HT model is the best among the three models. Compared with the other two models, HT model can predict the real information dissemination and reveal the information dissemination law.

5.3 The application of HT model

Online social networks are full of advertising messages and spread among users. How to get the maximum advertising return with the lowest advertising cost in online social networks has become a hot problem. Therefore, taking the rt-bar as the experimental data set, and simulating the advertising application. Set the initial number of dissemination for 30, 90, 150, the number of dissemination rounds for 30, and the number of dissemination for all three kinds of simulation ten times to take the results, the experimental results are shown in Table 4.

From Table 4, it can be seen that in the case of the initial spread of 30, the number of spreaders was 2026.4 at the spread round of 10, and the average spreaders were 4174.1 at the spread round of 30. In the case of 90 initial spreaders, the number of spreaders was 2560.9 for 10 rounds, and the average number of spreaders was 4280.1 for 30 rounds. With an initial spread of 150, the average number of spreaders is 2747.2 for a spread of 10 rounds and 4171.1 for a spread of 30 rounds.

It can be seen that in the initial stage of transmission, the greater the initial number, the greater the spread. Advertising is easier to spread than advertising in its early stages. In the middle and later stages of communication, the influence of dissemination decreases significantly. The reason is that most users already know the message, and the impact and return on investment of continued dissemination will be reduced.

6 Conclusion

A heat transfer model HT for the information dissemination of online social network is proposed in this paper. The model takes the initial influence nodes in the social network as the heat source and the information dissemination between nodes as the heat transfer process between objects. Combined with the node online characteristics

Table 4 The scope of information dissemination under different number of disseminators

Initial number	Rounds for 10	Rounds for 20	Rounds for 30
30	2026.4	3552.5	4171.1
90	2560.9	3785.4	4280.1
150	2747.2	3860.1	4309.3

and the information dissemination process of heat diffusion and heat receivers in HT model, the real information dissemination situation in online social network can be well predicted. The experimental results show that HT model has high accuracy and adaptability in simulating the information dissemination process of real social network.

The future work is as follows: (1) Nodes in online social networks correspond to individuals in real life, so there are many aspects worth exploring for node characteristics; (2) The information dissemination model proposed simulates the dissemination process of a single piece of information in social network, so it cannot be applied to the dissemination process of different information in social network. Therefore, the game between multiple sources of information will be further discussed.

Funding The project was supported by the National Natural Science Foundation of China (Grant Nos.: 62172352, 61871465) and The Science and Technology Research Project of Higher Education of Hebei Province (Grant Nos.: ZD2019004).

Data availability Metadata for model testing is available through the networkrepository.(rt-pot can be found at: https://networkrepository.com/rt-pol.php and rt-bahrain can be found at: https://networkrepository.com/rt-bahrain.php) Derived data supporting the findings of this study are available from the corresponding author on request.

Declarations

Conflict of interest The authors declare no competing financial interests.

References

1. Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. Proc R Soc Lon Ser A Contain Pap Math Phys Character 115(772):700–721
2. Fan Haibo XX, Liu Wanwan (2020) Research and simulation of UVFR rumor propagation model based on microblog. Comput Appl Res 37(5):1332–1335. https://doi.org/10.19734/j.issn.1001-3695.2018.10.0813
3. Harko T, Lobo FSN, Mak MK (2014) Exact analytical solutions of the susceptible-infected-recovered (sir) epidemic model and of the sir model with equal death and birth rates. Appl Math Comput 236:184–194. https://doi.org/10.1016/j.amc.2014.03.030
4. Yin F, Xia X, Pan Y, She Y, Feng X, Wu J (2022) Sentiment mutation and negative emotion contagion dynamics in social media: A case study on the Chinese Sina Microblog. Inf Sci 594:118–135. https://doi.org/10.1016/j.ins.2022.02.029
5. Nian F, Diao H (2020) A human flesh search model based on multiple effects. IEEE Trans Netw Sci Eng 7(3):1394–1405. https://doi.org/10.1109/TNSE.2019.2931943
6. Meel P, Vishwakarma DK (2020) Fake news, rumor, information pollution in social media and web: a contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst Appl 153:112986. https://doi.org/10.1016/j.eswa.2019.112986
7. Chai Y, Wang Y, Zhu L (2019) A stochastic information diffusion model in complex social networks. IEEE Access 7:175897–175906. https://doi.org/10.1109/ACCESS.2019.2957764
8. Jiang C, D'Arienzo A, Li W, Wu S, Bai Q (2021) An operator-based approach for modeling influence diffusion in complex social networks. J Soc Comput 2(2):166–182. https://doi.org/10.23919/JSC.2021.0007
9. Xia Z, Tan Z, Zhang Y, Zhang S, Ma Y (2019) A novel information diffusion model inspired by particle-collision dynamics for online social networks. In: 2019 IEEE Intl Conf on Parallel Distributed

Processing with Applications, Big Data Cloud Computing, Sustainable Computing Communications, Social Computing Networking (ISPA/BDCloud/SocialCom/SustainCom), pp 1629–1634. https://doi.org/10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00240
10. Ulyana Y, Andrij P (2017) Conceptual model the speed of information dissemination in social networks on the basis of the Monod equation. In: 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), vol. 1, pp 318–321. https://doi.org/10.1109/STC-CSIT.2017.8098795
11. Goldenberg J, Libai B, Muller E (2001) Talk of the network: a complex systems look at the underlying process of word-of-mouth. Mark Lett 12(3):211–223
12. Granovetter M (1978) Threshold models of collective behavior. Am J Sociol 83(6):1420–1443
13. Wei Z, He Y, He W, Liu X (2017) Information dissemination model based on clustering analysis of information network development. In: 2017 4th International Conference on Information Science and Control Engineering (ICISCE), pp 403–407. https://doi.org/10.1109/ICISCE.2017.91
14. Ju-Ping Z, Hao-Ming G, Wen-Jun J, Zhen J (2019) Dynamic analysis of rumor propagation model based on true information spreader. Acta Physica Sinica 68(15): . Publisher: CHINESE PHYSICAL SOC PO BOX 603, BEIJING 100080. PEOPLES R CHINA. https://doi.org/10.7498/aps.68.20190191
15. Yujia W, Heyin H (2019) Study of an improved rumor-propagation model on small world networks. J Intell 38(4):138
16. Fu G, Chen F, Liu J, Han J (2019) Analysis of competitive information diffusion in a group-based population over social networks. Physica A 525:409–419. https://doi.org/10.1016/j.physa.2019.03.035
17. Sun H, Sheng Y, Cui Q (2021) An uncertain SIR rumor spreading model. Adv Differ Equ 2021(1):286. https://doi.org/10.1186/s13662-021-03386-w
18. Zhang Y, Chen Z (2019) SETQR propagation model for social networks. IEEE Access 7:127533–127543. https://doi.org/10.1109/ACCESS.2019.2939150
19. Mathew B, Dutt R, Goyal P, Mukherjee A (2019) Spread of hate speech in online social media. In: Proceedings of the 10th ACM Conference on Web Science. WebSci '19, pp 173–182. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3292522.3326034. event-place: Boston, Massachusetts, USA
20. Hui H, Zhou C, Lü X, Li J (2020) Spread mechanism and control strategy of social network rumors under the influence of COVID-19. Nonlinear Dyn 101(3):1933–1949. https://doi.org/10.1007/s11071-020-05842-w
21. Wang Y, Wang J, Wang H, Zhang R, Li M (2021) Users' mobility enhances information diffusion in online social networks. Inf Sci 546:329–348. https://doi.org/10.1016/j.ins.2020.07.061
22. Liu X, He D, Liu C (2019) Information diffusion nonlinear dynamics modeling and evolution analysis in online social network based on emergency events. IEEE Trans Comput Soc Syst 6(1):8–19. https://doi.org/10.1109/TCSS.2018.2885127
23. Zhang L, Li H, Zhao C, Lei X (2017) Social network information propagation model based on individual behavior. China Commun 14(7):78 (**Publisher: China Communications**)
24. Chou C-K, Chen M-S (2018) Learning multiple factors-aware diffusion models in social networks. IEEE Trans Knowl Data Eng 30(7):1268–1281. https://doi.org/10.1109/TKDE.2017.2786209
25. Yi Y, Zhang Z, Yang LT, Gan C, Deng X, Yi L (2021) Reemergence modeling of intelligent information diffusion in heterogeneous social networks: the dynamics perspective. IEEE Trans Netw Sci Eng 8(2):828–840. https://doi.org/10.1109/TNSE.2020.2975112
26. Rossi RA, Ahmed NK (2015) The network data repository with interactive graph analytics and visualization. In: AAAI. https://networkrepository.com