mathematics

Article

A Multi-Information Dissemination Model Based on Cellular Automata

Changheng Shao [1], Fengjing Shao [1], Xin Liu [2], Dawei Yang [2], Rencheng Sun [1,*], Lili Zhang [1] and Kaiwen Jiang [1]

1 College of Computer Science and Technology, Qingdao University, Qingdao 266071, China; sch_mail@163.com (C.S.); sfj@qdu.edu.cn (F.S.); zhanglili117@yeah.net (L.Z.); 2021023831@qdu.edu.cn (K.J.)
2 College of Computer Science and Technology, China University of Petroleum, Qingdao 266580, China; lx@upc.edu.cn (X.L.); 18351805913@163.com (D.Y.)
* Correspondence: src@qdu.edu.cn

Abstract: Significant public opinion events often trigger pronounced fluctuations in online discourse. While existing models have been extensively employed to analyze the propagation of public opinion, they frequently overlook the intricacies of information dissemination among heterogeneous users. To comprehensively address the implications of public opinion outbreaks, it is crucial to accurately predict the evolutionary trajectories of such events, considering the dynamic interplay of multiple information streams. In this study, we propose a SEInR model based on cellular automata to simulate the propagation dynamics of multi-information. By delineating information dissemination rules that govern the diverse modes of information propagation within the network, we achieve precise forecasts of public opinion trends. Through the concurrent simulation and prediction of multi-information game and evolution processes, employing Weibo users as nodes to construct a public opinion cellular automaton, our experimental analysis reveals a significant similarity exceeding 98% between the proposed model and the actual user participation curve observed on the Weibo platform.

Keywords: public opinion; cellular automata; multi-information game; propagation dynamics

MSC: 94-10

check for updates

Citation: Shao, C.; Shao, F.; Liu, X.; Yang, D.; Sun, R.; Zhang, L.; Jiang, K. A Multi-Information Dissemination Model Based on Cellular Automata. *Mathematics* **2024**, *12*, 914. https://doi.org/10.3390/math12060914

Academic Editors: Mariano Torrisi and Liguo Fei

Received: 22 December 2023
Revised: 24 February 2024
Accepted: 16 March 2024
Published: 20 March 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Numerous components within complex systems exhibit intricate connectivity patterns and dynamic processes, often characterized by mutual dependencies among heterogeneous units or the coupling of multiple layers, resulting in the coexistence of diverse relationships among homogeneous units [1,2]. In the realm of a complex and interconnected network, public opinion is influenced by a myriad of factors encompassing unique individual values, beliefs, cultural contexts, and social affiliations. Consequently, a multitude of information dissemination phenomena coalesce within this intricate network, where various pieces of information interact dynamically throughout the dissemination process, giving rise to instances of either reinforced or attenuated dissemination which reflect the intricate dynamics of public opinion propagation.

Data from the China Internet Network Information Center [3] reveals that the number of internet users and the penetration rate in China has reached 1.079 billion and 76.4%, respectively, underscoring the widespread adoption of the Internet. However, users navigating social networks often lack a comprehensive understanding of public opinion and are frequently exposed to rumors and misinformation, posing threats to public safety and the social order [4]. Consequently, this phenomenon often leads to a cascading and secondary spread of public opinion.

Hence, public opinion dissemination analysis has become a pressing concern among scholars, necessitating an exploration of the mechanisms governing public opinion dissemination, especially with the coexistence of multiple pieces of information. In the broader context of research on the evolution of online public opinion abroad, a multidisciplinary approach prevails. Scholars commonly employ three established strategies to delve into the intricacies of public opinion dissemination: complex networks, propagation dynamics and cellular automata. In detail, complex networks provide insights into network structures and connectivity patterns facilitating the understanding of topology and evolution [5–7]; propagation dynamics modeling offers insights into consensus emergence and information cascades, but it may suffer from limited predictive power and sensitivity to initial conditions [8–12]; cellular automata offers a simplified and interpretable strategy for simulating complex system dynamics, although they may oversimplify real-world dynamics and require refinement for accurate modeling [13–15]. Notably, research focusing on public opinion dissemination based on cellular automata has emerged as a prominent area in public opinion analysis.

Consequently, there is a need to develop more sophisticated propagation mechanisms based on the cellular automata model to enhance the accuracy of public opinion trend prediction. In this paper, we propose the SEInR model that incorporates different states to capture the complexities of public opinion dynamics based on the classical SEIR (Susceptible-Exposed-Infectious-Recovered) model. By implementing the SEInR model within a cellular automaton framework, we capture the nuanced interactions between multi-information dissemination, thereby enhancing the reliability and precision of our predictions. Finally, the reliability and accuracy of the model are validated through experiments.

2. Related Works

2.1. Propagation Dynamics

The utilization of social networks facilitates the exchange of diverse information, leading to multiple emotional expressions. Users heavily rely on friendships within online social networks to express and propagate information opinions. The dissemination of online public opinion information follows a systematic, dynamic and complex life cycle, necessitating a comprehensive study of intricate mechanisms. With complex topology feature, online social networks contain diverse collection of various popular online sentiments [16–18].

Propagation dynamics are a crucial facet of information dissemination which explore how information spreads and evolves across diverse social contexts, encompassing vehicular transportation systems, urban structures, viral infections and public opinion dynamics. Understanding user behavior and psychology dynamics is essential for comprehending public opinion formation. This field primarily explores how information dissemination occurs, how messages are interpreted, and how communication influences network dynamics. Research in transmission dynamics often adopts a multidisciplinary approach, drawing insights from communication theory, sociology, psychology, anthropology and related disciplines.

In the realm of information dissemination, several prevalent models exist, including the following:

(1) The Susceptible–Infectious (SI) model [19] represents the simplest model that can capture the transition from susceptible to infectious states. Shao et al. introduced a SIn model for multi-information spreading and demonstrated its ability to imitate and predict dynamic behaviors [20].
(2) The Susceptible–Infectious–Susceptible (SIS) model [21] extends the SI model, allowing for repeating or recurring infections with infected individuals returning to the susceptible state. Xuan et al. proposed a network continuous-time SIS model coupled with individual opinion dynamics [22].
(3) The Susceptible–Infectious–Recovered (SIR) model [23,24] is a generic epidemiological model that describes the transmission of infectious diseases through individuals who transit between susceptible, infectious, and recovered states. Han et al. analyzed the impact of human activity patterns on information diffusion using the SIR model [25].
(4) The Susceptible–Knowledgeable–Infectious–Recovered (SKIR) model extends the traditional SIR model by introducing a "knowledgeable" state, where individuals

have been exposed to both the disease or information and counteracting knowledge. Xiao et al. proposed an SKIR rumor propagation model to describe the propagation of rumors and the dynamic changes in the influence of anti-rumor information [26].

(5) The Susceptible–Exposed–Infectious–Recovered (SEIR) model [26,27] introduces an "exposed" state between being susceptible and infectious, representing individuals who have been exposed to the disease or information but are not yet infectious. This model is particularly relevant for diseases with an incubation period. Li et al. proposed a public opinion evolution HK–SEIR model which combines the opinion fusion HK and the epidemic transmission SEIR models [28].

2.2. Cellular Automata

A cellular automaton (CA) is a discrete grid-based dynamic model capable of simulating the spatiotemporal evolution processes of complex systems, wherein spatial interactions and temporal causality are localized. By iteratively applying predefined rules based on the states of neighboring cells, the CA drives the evolution of these systems [29–31]. Its versatility spans diverse fields, including image processing [32–34] and traffic management [35–37], underscoring its profound influence on scientific inquiry and technological innovation.

Meanwhile, a CA is utilized in the context of information dissemination [38,39] to model the spread of information through a network of interconnected cells or nodes. Each cell represents an individual or entity within the network, and the state of each cell evolves over discrete time steps and states of neighboring cells according to predefined rules. By simulating the interactions and dynamics between cells, it provides insights into the propagation of information, the interpretation of messages and the influence of communication on networks. The model helps researchers to understand the complex mechanisms underlying information dissemination and predict trends in public opinion dynamics. As such, public opinion prediction based on CA has been a hot topic recently. Liu et al. [40,41] proposed public opinion cellular automata for situation deduction to predict the possible trending of public events. It has been proved that the cellular automata-based prediction of the number of participants and emotional trending are more accurate than other methods.

While cellular automata have demonstrated promise in modeling the dissemination of public opinion, existing research has predominantly concentrated on two-dimensional relationships, thus constraining their ability to capture the intricacies of opinion dissemination within networked environments. Addressing this gap, our paper introduces the SEInR model, a novel application of a CA to the multi-state of public opinion dissemination. By employing meticulous analysis and empirical validation, our goal is to enhance our understanding of public opinion dynamics and advance computational modeling methodologies in the social sciences.

3. Detailed Model

3.1. Model Characteristic

3.1.1. Model Definition

Definition 1. *Node State. Although existing models have been widely used in the spread of public opinion, they usually do not consider the reality of information dissemination among multiple infected users. Here, an infected user is defined as a user who has been exposed to public opinion and influenced by it, forming a certain attitude or opinion that is resistant to further influence. In reality, the interaction of multiple pieces of information within a network is complex, leading to various states such as negative and positive. However, there is a lack of comprehensive models to describe the propagation process and interactions between multiple pieces of information with different infected states.*

Given the premise assumption of the diffusion process for multiple pieces of public opinion information and the transfer process for individual states, it becomes necessary to construct a multi-information dissemination mechanism based on cA. This paper proposes an SEInR model *with features for multi-information public opinion propagation, aiming to analyze the propagation process of public opinion in crowds with high accuracy and objectivity.*

In the enhanced cellular automata model, node states are categorized into $n + 3$ *types:* $\{S, E, I_1, \ldots, I_n, R\}$. *This classification is illustrated by Figure 1 and Table 1:*

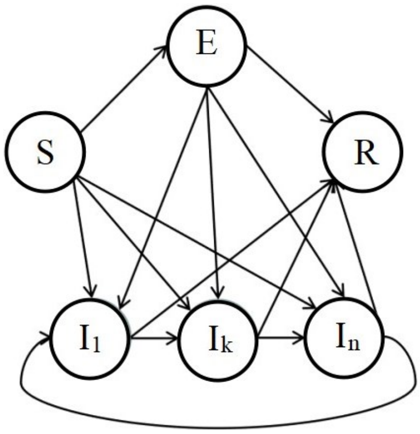


Figure 1. The transition of S-E-I1-I2...-In-R.

Table 1. Cell node state types.

Numbers	States	Notations
1	Susceptible State (S)	Nodes not involved in public opinion topics
2	Exposed State (E)	Nodes involved in public opinion topics
3	Infectious State of Information 1 (I1)	Nodes that have been exposed to information of state 1 and have been affected
…	…	…
$m-1$	Infectious State of Information n (In)	Nodes that have been exposed to information of state n and have been affected
m	Recovered State (R)	Nodes previously participated in a topic have withdrawn from the topic and no longer participate in discussions or dissemination

Definition 2. *State Transition. In the cellular automata model, a state is usually transited to a different state. When the state of a node transits from the* $\Gamma^{t-1}(node)$ *to the* $\Gamma^{t}(node)$, *the* $\Gamma^{t-1}(node)$ *stands as the source state of the node in* $t-1$ *time and the* $\Gamma^{t}(node)$ *stands as the current state of the node in t time.*

Definition 3. *Node model. Within the cellular automata model, nodes serve as representations of users within the network's public opinion space. These nodes undergo state transitions based on varying conditions. For example, nodes initially uninvolved in a topic may transit to either the "Exposed State" or the "Infectious State" upon encountering public opinion information. Similarly, nodes in an Engaged state may shift to the "Exposed State" when influenced by neighboring nodes in the propagating state. Nodes in the "Exposed State" are susceptible to influence from neighboring nodes in different states, potentially leading to state transformation into other types. Over time, nodes in the "Exposed State" or "Infectious State" may recover to the "Recovered State" due to various factors. This dynamic framework encapsulates the evolving feature of public opinion dissemination within the network, shedding light on the intricate interplay between different states* *and the impact of neighboring nodes on state transitions. The differential equation of node model is defined as follows:*

$$\frac{ds(t)}{dt} = -(\beta_{12}s(t)e(t) + \beta_{13}s(t)i_1(t) + \ldots + \beta_{1(m-1)}s(t)i_n(t)) \tag{1}$$

$$\frac{de(t)}{dt} = \beta_{12}s(t)e(t) - (\beta_{23}e(t)i_1(t) + \ldots + \beta_{2(m-1)}e(t)i_n(t) + \beta_{2m}e(t)r(t)) \tag{2}$$

$$\frac{di_1(t)}{dt} = \beta_{13}s(t)i_1(t) + \ldots + \beta_{(m-1)3}i_n(t)i_1(t) - \beta_{3m}i_1(t)r(t) - \beta_{32}i_1(t)i_n(t) \tag{3}$$

$$\frac{dr(t)}{dt} = \beta_{2m}e(t)r(t) + \beta_{3m}r(t)i_1(t) + \ldots + \beta_{(m-1)3}i_n(t)i_1(t) \tag{4}$$

Here, β corresponds to transitions of probability from the $\Gamma^{t-1}(node)$ to the $\Gamma^{t}(node)$. And $s(t)$, $i_n(t)$, $e(t)$, $r(t)$ represent the proportions of nodes in state S, In, E and R, respectively.

Definition 4. *Inference rule. In the cellular automata model, the state transition of user nodes and propagation is expressed by the following:*

$$\Gamma(node)^t \in S, E, I_1, \ldots, I_n, R$$

(1) S→E Transition:

$$\Gamma(node)^t = E, if(N_p - N_n\delta_1) + (W_p - W_n\delta_2) + (N_s - N_n\delta_3) > 1 and \Gamma(node)^{t-1} = S. \tag{5}$$

where N_p represents the number of user nodes in the neighborhood space of a user node expressing a certain opinion, N_n represents the total number of user nodes in the neighborhood space of a user node, W_p represents the total number of user nodes in the entire space expressing a certain opinion, W_n represents the total number of user nodes in the entire space and $\beta_N s$ represents the total number of user nodes in the neighborhood expressing a certain opinion; δ_i represents the transition factor.

(2) S,E→I Transition: When users engage in discussions on public opinion topics, they may exhibit a proactive or reactive behavior when expressing their viewpoints. These behaviors must be processed with different conditions. Active Propagation: If a cellular node is actively participating, there exists a probability that it will contribute relevant commentary in the subsequent time step. This probability is determined by I_v the "Independent Opinion Index" of the node.

Passive Propagation: If a cellular node is in the participating state, neighboring nodes exhibit a propagation influence greater than that of the node at the previous time step; then, it may passively transition to a propagation state (I) with a certain probability.

Continuation of Participation: If the node remains in a participating state without entering either the active or passive propagation states, it will persist in this state until it meets the conditions for exiting.

(3) E,I→R Transition: When both the permanent exit time limit and the temporary exit time limit are less than 0, the user will transit to the exit state (R).

3.1.2. Model Properties

3.1.2.1. Balance State. In the process dissemination of information, the network is in equilibrium when the values of $s(t), e(t), i_1(t), \ldots, i_n(t), r(t)$ remain unchanged.

3.1.2.2. Dissemination influence. In order to enable user nodes to reflect the situation of the entire network's public opinion space, each node contains eight attributes. These are the user's node influence I, the independent opinion index Iv, the forwarding index Sc, opinion firmness Sp, opinion interest Ip , emotional inclination Mu, frequency of idle remarks Is and interest list Il. The value of the attributes will affect the user's response to different public opinion information, such as whether they are interested in the topic and which emotional attitude they hold. The attribute values of each user node are determined based on their own historical comments. In addition to attributes, user nodes will also be in different states during the deduction process. The changes to cellular state nodes are introduced as follows:

The dissemination influence I of node V reflects the importance of the opinion subject in the online public opinion space. The more influential the node is, the more users will accept the opinion published by the node, which can even affect the emotional and interest tendencies of its fan nodes. In the model proposed in this article, the node influence of the user node is based on the user's activity degree W_1 and the dissemination degree W_2. Here, the activity degree W_1 is computed based on the total number of user comments X_1 and the total number of original comments X_2. While dissemination degree is comprehensively computed based on the number of reposts X_3, the number of responses X_4, reposts of original content X_5, responses to original content X_6 and likes received.

Based on the weight ratio of each part in Table 2, the formula for calculating node dissemination influence is shown in the following equation:

$$I = (CW_1 + (1-C)W_2) \tag{6}$$

The calculation formula for user activity W_1 and the calculation formula for the activity of user comments W_2 are shown in the following equations, respectively.

$$\begin{gathered} W_1 = \lambda_1 \ln(X_1 + \mathbf{1}) + \lambda_\mathbf{2} \ln(X_\mathbf{2} + \mathbf{1}) \\ \lambda_1 + \lambda_\mathbf{2} = \mathbf{1} \\ 0 \le \lambda_\mathbf{i} \le \mathbf{1} \end{gathered} \tag{7}$$

$$\begin{aligned} W_\mathbf{2} = {} & \alpha_1 \ln(\mathrm{X}_3 + \mathbf{1}) + \alpha_\mathbf{2} \ln(\mathrm{X}_4 + \mathbf{1}) + \alpha_3 \ln(\mathrm{X}_5 + \mathbf{1}) \\ & + \alpha_\mathbf{4} \ln(\mathrm{X}_\mathbf{6} + \mathbf{1}) + \alpha_5 \ln(\mathrm{X}_7 + \mathbf{1}) \end{aligned} \tag{8}$$

$$\begin{gathered} \alpha_\mathbf{1} + \alpha_\mathbf{2} + \alpha_\mathbf{3} + \alpha_\mathbf{4} + \alpha_5 = \mathbf{1} \\ \mathbf{0} \le \alpha_\mathbf{i} \le \mathbf{1} \end{gathered} \tag{9}$$

By collecting historical comments from users, and distinguishing between forwarded and original comments, we can obtain the number of reposts, responses and likes per post. Based on the above formula, the user's node influence can be calculated. The larger the value of the dissemination influence, the greater the user's influence in the online public opinion space.

Table 2. Weight ratio for influence calculation.

Activity degree W_1	Total number of comments X_1 (λ_1) Total number of original comments X_2 (λ_2)
Dissemination degree W_2	Total number of reposts X_3 (α_1) Total number of responses X_4 (α_2) Total number of original reposts X_5 (α_3) Total number of original responses X_6 (α_4) Total number of likes X_7 (α_5)

3.1.2.3. Independent Opinion Index. In the online public opinion space, some users enjoy exploring others' comments and forwarding them, but rarely actively express certain opinions themselves. At the same time, some users are the main publishers of online public opinions. Furthermore, these users will actively express their opinions to participate in opinion dissemination when public opinion events occur. The user's independent opinion index "Iv" is used to measure the probability of users actively expressing their own opinions when participating in online public opinion events. In the cellular automata modeling, the independent opinion index will be used to determine whether a user node will make comments about public opinion events or not. The calculation formula is shown as follows:

$$Iv = \frac{X_2}{X_1} \tag{10}$$

where X_2 represents the number of original user comments and X_1 represents the number of all comments made by the user. The higher the user's independent opinion index, the more frequently the user will make original comments when participating in public opinion events.

3.1.2.4. Forwarding index. The user's forwarding index Sc is used to measure the likelihood of users actively disseminate information when participating in online public opinion events. In the network public opinion inference model, the forwarding index will be used to determine whether user nodes should disseminate their opinions during public opinion events or not. The calculation formula is shown as follows:

$$Sc = 1 - Iv \tag{11}$$

And the sum of the user's forwarding index and the user's independent opinion index is one. The higher the user's forwarding index is, the greater the probability that the user will forward relevant opinions when participating in public opinion events.

3.1.2.5. Opinion firmness. In the online public opinion space, users have a certain level of judgment ability regarding the comments they receive, making it difficult for the received comments to affect their original views. The firmness of opinion Sp signifies the extent to which individuals' personal opinions are susceptible to external influence from neighboring users and the surrounding environment. The following steps are designed to calculate opinion firmness. We denote this process in Algorithm 1.

In the described algorithm, the initial step involves the extraction of keywords from all user comments and forwards within a specified public opinion event, leading to the creation of an opinion keyword matrix. Subsequently, the Principal Component Analysis (PCA) dimensionality reduction algorithm is applied to compress the vectors within this matrix into two-dimensional representations. During this reduction process, two-dimensional vectors that meet predefined criteria for local density and relative distance between centers are identified, effectively serving as clustering centers among the reduced vectors. A higher count of such clustering centers indicates a more dispersed distribution of user opinions during the event, reflecting a decreased level of conviction or firmness in their expressions. This analytical methodology yields valuable insights into the dynamics of public opinion dissemination and user engagement within the investigated context.

3.1.2.6. Topic Initiation Ability. Certain users exhibit a keen interest in exploring a diverse array of public opinion topics and demonstrate a high sensitivity to trending internet discussions. Typically, these users engage in public opinion events not by relying on interactions with neighboring nodes but rather by monitoring hot search topic rankings to stay abreast with events and actively participate in relevant discussions. The frequency of idle comments made by users serves as a metric to gauge their responsiveness to public opinion hotspots within the broader online environment. The subsequent procedure outlined herein is tailored to compute the frequency of idle comments made by users. We denote this process in Algorithm 2.

Algorithm 1: Firmness of opinion

Input: All the comments made and forwarded by a user in a certain public opinion event as the user's opinion dataset $\{Ds\}_{i=1}^{n}$, new word discovery algorithm Ns;
Output: The firmness of the user's opinion Sp;
for $i = 1 \rightarrow n$ **do**
  Obtain segmentation results $w_i = \{t_1, t_2, ..., t_l\}$ of comment $di \in Ds$ by Ns;
  for $j = 1 \rightarrow len(Wi)$ **do**
    Set $c(t, w)$ to represent the frequency of token appearing in comment w;
    get the weight $TF - IDF$ of each word:
    $TF - IDF(t_j, w_i) = \frac{c(t_j, w_i)}{\sum n_{k=0}(c(t_j, w_k))} \times log(\frac{n}{c(t_j, Ds)})$
  Sort the $TF - IDF$ values of all tokens in w_i in descending order to obtain the top 10 weighted tokens $K_i = \{k_{i1}, ..., k_{i10}\}$ as keywords, if the number of tokens in w_i is less than 10, use all tokens in w_i as keywords.
Get the keyword dataset $D_k = \{K_1, ..., K_n\}$ of the user's opinion dataset Ds.
Get the keyword set $S_k = \{s_1, ..., s_m\}$, m is equal to the number of non repeating keywords present in D_k.
for $i = 1 \rightarrow n$ **do**
  for $o = 1 \rightarrow m$ **do**
    if s_o *in* K_i **then**
      $a_{io} = 1$;
    else
      $a_{io} = 0$;
    Get comment vector $V_i = [a_{i1}, ..., a_{im}]$.
  Get vector matrix $W = [V_1, ..., V_n]$ of user's opinion.
  Get the dimensionality reduction representation of the vector matrix: $W^p = PCA(W)$.
for $i = 1 \rightarrow n$ **do**
  set $Ed(w_i, w_j)$ represents the Euclidean distance between vector w_i and vector w_j.
  the local density $Ld_i = 0$.
  for $j = 1 \rightarrow n$ **do**
    if $E_d(w_i^p, w_j^p) < e^l$ **then**
      $Ld_i = Ld_i + 1$;
for $i = 1 \rightarrow n$ **do**
  if $Ld_i = Ld^{max}$ **then**
    the relative distance $Rd = E_d^{max} w_i^p, w^p$;
  else
    $Rd = E_d^{min} w_i^p, w^p$, where $E_d(w_i^p, w^p) = E_d^{min} w_i^p, w^p$
for $i = 1 \rightarrow n$ **do**
  $Cn = 0$; **if** $Ld_i > e^{ldt}$ *and* $Rd_i > e^{rdt}$ **then**
    $Cn = Cn + 1$;
  $Sp = \frac{1}{Cn}$;
Return (Sp)

Algorithm 2: Topic Initiation Ability

Input: All comments posted and forwarded by the user as the user's opinion dataset Ds

Output: The frequency of idle comments by this user Is;

1 Obtain all comments posted and forwarded by the user as the user's opinion dataset Ds;

2 Perform steps 2 to 8 in the above Algorithm 1;

3 Set the discrete local density threshold Ld_t' and the discrete relative distance threshold Rd_t', and count the number of vectors Cn in W_{PCA} with local density less than Ld_t' and relative distance greater than Rd_t';

4 Return($\frac{Cn}{X_1}$)

Based on Algorithm 2, the vectors in W_{PCA} with a local density less than the discrete local density threshold Ld_t' and relative distance greater than the discrete relative distance threshold Rd_t' which can be considered as discrete points in all user comment data. The greater the number of discrete points, the broader the range of public opinion topics that users are engaged with, indicating a heightened sensitivity to events within the online public opinion space. Such users are predisposed to readily participate in topic events without relying on information dissemination from neighboring nodes.

In the online public opinion space, each user maintains their own set of interests. Despite the topic's overall popularity, users tend to be more receptive to comments within their interest areas, while showing less engagement with topics outside those areas. Consequently, user participation becomes challenging when topics fall beyond their interests.

The interest list, denoted as Il, comprises a q-dimensional vector. Each dimension represents a distinct field of interest, with values ranging between 0 and 1. A higher value indicates greater interest in the corresponding field, while a lower value reflects lesser interest.

To establish the interest list, we compute the occurrences of keywords from different fields across all user comments. The frequency of keywords in a particular field directly correlates with the user's interest level in that field. The value of the i-th dimension in the interest list is calculated as follows:

$$Il_i = \frac{Kw_i}{Kw_{max}} \tag{12}$$

where Kw_i represents the number of times a keyword in the i-th interest field appears in the user's entire opinion, and Kw_{max} represents the maximum number of occurrences of keywords in each of the q domains across all user comments. Therefore, the dimension corresponding to the domain of interest that users are most interested in is set to one. Throughout the inference process, an interest list serves as an effective way of demonstrating the user's level of interest in comments across various fields.

3.1.3. Sentiment Orientation

The sentiment orientation refers to the degree of intensity with which a subject expresses positive or negative emotions towards an object. These varying degrees of emotion are typically conveyed through different emotional words or tones. To accurately capture this phenomenon, it is common practice to assign different weights to each emotional word.

Furthermore, user sentiment orientation refers to the inclination of users to align themselves more closely with specific emotional orientations when interacting with particular public opinion topics. It also encompasses the emotional orientations they are prone to express when providing comments or opinions.

To improve the accuracy of computing user sentiment orientation, this study extends the existing sentiment lexicon. By leveraging this enhanced sentiment lexicon, following is proposed in this paper to accurately compute user sentiment orientation, thereby offering a deeper insight into how users express and engage with emotions in the context of public opinion discussions. It is computed as following:

1. Collect a vocabulary of positive words, negative words, negative words and degree adverbs;
2. Obtain all comments posted and forwarded by the user, and initialize the emotional value $M = 0$ for each comment;
3. Use a new word-discovery algorithm based on the association confidence of the word segmentation of each user's opinion data;
4. Traverse through the word sequence obtained from step 3 for each statement. If a keyword appears in the positive word library, determine whether the previous word is a definite or degree adverb. If it is a negative word, reduce the value of M by one; If it is a degree adverb, increase the value of M by two; If it is not a negative word or a degree adverb, increase the value of "M" by one; If a keyword appears in the negative word vocabulary, determine whether the previous word is a definite or degree adverb; If it is a negative word, increase the value of M by one; If it is a degree adverb, decrease the value of M by two; If it is not a negative word or degree adverb, decrease the value of M by one.
5. Based on step 4, calculate the emotional value of each comment from the user, and the user's sentiment orientation is $Mu = \frac{1}{n}\sum_{i=1}^{n} M_i$, where n is the total number of comments made by the user, M_i is the sentiment orientation value of the user's i-th comment.

Compute the average emotional value of a user's opinion using the above procedure to determine the user's sentiment orientation. This orientation influences the emotional bias of users' comments during the deduction process. A higher sentiment orientation indicates a predominantly positive emotional stance, whereas a lower orientation suggests a more negative emotional disposition.

3.2. Model Definition

3.2.1. Preliminary Segmentation

To better address the segmentation and computation of short texts in public opinion and enable an accurate assessment of the firmness of user node comments, this paper proposes a novel word discovery algorithm based on association confidence. In order to prevent the occurrence of accidental word units and their left and right adjacent words with a 100% association confidence, it is necessary to merge the short text before preliminary word segmentation to increase the length of the text to be segmented. Assuming a total of n text data to be segmented, merge each m text datum to obtain the merged text segment $Text' = Text_1 + Text_2 + \ldots + Text_m$, resulting in a total of $\lceil nm \rceil$ merged text segments. For each merged text segment, use the precise mode from the Jieba tool library to segment, and obtain multiple segmentation results Tc'. For each segmentation result Tc', proceed to the next step of processing.

3.2.2. Correlation Confidence

The correlation confidence level of the association between each word unit and its left and right adjacent word units is determined based on each segmentation result, denoted as Tc. Candidate new words are obtained by merging multiple word units that satisfy the correlation confidence threshold Th, forming the candidate new word set W.

There are two scenarios with candidate new words in the candidate new word set: one involves consolidating overly fragmented words to obtain correct new word results, while the other pertains to phrases formed from excessive merging. If combinations of multiple word units appear multiple times and repetitively in the text, they are merged due to meeting the correlation confidence threshold. Some phrases may represent longer text segments formed by the combination of multiple correct words, while others may denote lengthier named entities, such as network terminologies or foreign names. Consequently, it is necessary to filter the candidate new words in the candidate new word set, retaining only those with practical significance and smaller.

3.2.3. Splitting Conjunctions

To refine the granularity of newly identified words, it is crucial to address cases of excessive merging within phrases. This necessitates further segmentation of phrases, especially when conjunctions are detected within candidate new words. Specific measures are taken to ensure optimal segmentation when candidate new words contain conjunctions. It is computed as follows:

1. Compute the average correlation confidence for each connecting word w_i and its adjacent word w_j to the left or right; this is the average value of $RConf(w_i \rightarrow w_j)$ and $RConf(w_j \rightarrow w_i)$;
2. If the average correlation confidence values differ between a connecting word and its adjacent word units, the candidate new word undergoes splitting. The split point is determined between the connecting word and the adjacent word units with lower average correlation confidence values.
3. When the average correlation confidence value is consistent among a connecting word and its adjacent word units, maintain the merging state of the two word units. Proceed to identifying the next connecting word in the candidate new word.
4. Following the aforementioned steps of connecting word splitting, the resulting word unit sequence represents the final word segmentation outcome of the text segment. This sequence encompasses both newly formed words by combining word units and separated connecting words.
   By implementing the splitting of connecting words within candidate new words, the phrase blocks formed from merging multiple word units can be dismantled. This process effectively reduces the granularity of the final new word result while ensuring semantic coherence. Consequently, the accuracy of the new word result is enhanced.

Drawing upon the newly proposed new word discovery algorithm, a comprehensive word segmentation process can be executed on both pre-existing historical data and recently acquired public opinion texts, facilitating the precise identification of words imbued with distinct semantic nuances within Chinese sentences. Subsequently, established keyword extraction techniques such as TF-IDF can be harnessed to distill pivotal terms from the textual corpus. Furthermore, the utilization of an incremental association rule mining model enables the exploration of association patterns between keywords entrenched in the historical dataset and those introduced in the newly acquired corpus. This integrated approach fosters a nuanced comprehension of semantic structures and evolving trends within the corpus, thereby augmenting the efficacy of text analysis and information retrieval methodologies.

3.2.4. Emotional Calculation

The calculation rule for emotional value information is as follows:

1. When there are no neighboring users posting comments, the user's emotional value and information entropy are updated to the emotional value and information entropy of the received comments.
2. When neighboring users make comments, the user's emotional value is represented by the product of the impact of their idle comments and the average emotional value of all users neighboring comments; the impact of idle user comments is represented by the product of the average information entropy of all user neighbor comments.

The exit topic deduction rule is the following:

1. Establish a fixed duration for users' participation in an event, setting a permanent exit time limit. This limit decreases by 1 after each round of user engagement in the event.

When the permanent exit time limit reaches or falls below 0, the user discontinues their involvement in the event discussion, opting out permanently.

2. Upon a user's initial participation in an event, assign a temporary exit time limit. With each successive round of user activity in the event, this limit decreases by 1 in the absence of any comments from the user. If the temporary exit time limit drops to or below 0, the user's departure from the event discussion is subject to reconsideration based on the participation discussion rule.

4. Experiments and Results

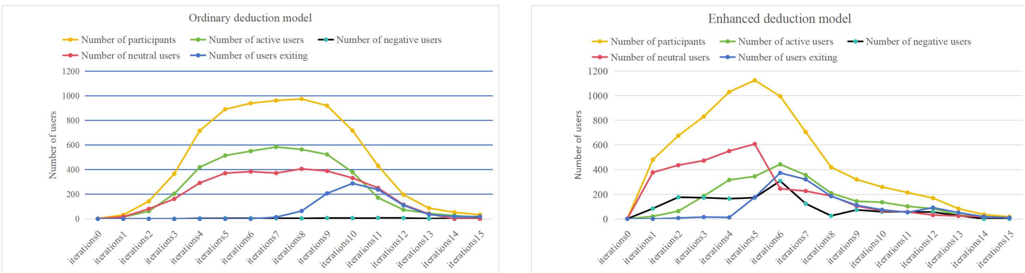
The experiments were conducted on a 64 bit Windows 10 operating system, employing Python 3.7 as the programming language within the development environment PyCharm 2020.

Before conducting the experiments, we initialized the inferred public opinion event information, representing domains that are relevant to public opinion events using topic vectors. Each dimension of the topic vector corresponds to a distinct domain, with values ranging from 0 to 1 indicating the relevance of the public opinion event to the represented domain content. The initial public opinion event comprises emotional values, entropy of multiple information pieces and the set of nodes initially employed for information dissemination.

To evaluate the impact of different factors on the inference process and results, we conducted comparative experiments for the selection of the initial propagation nodes and performed inference simulations. Through systematic experimentation and analysis, we aim to provide insights into the effectiveness and versatility in deducing public opinion dynamics.

4.1. SEInR Model

In this paper, we used a game inference cellular automaton model. We conducted an experimental comparison between the ordinary inference cellular automaton model and the game inference cellular automaton model (SEInR) constructed in this article, shown in the following Figure 2:



(**a**) Ordinary Deduction Model (**b**) Enhanced Deduction Model

Figure 2. Comparison between ordinary inference cellular automaton model and the enhanced cellular automaton model.

The conventional cellular automaton inference model relies solely on basic inference rules. In contrast, the SEInR model concurrently processes positive, negative, and neutral information, providing a comprehensive view of the information dissemination dynamics. In the diagram above, the "Number of participants" depicted by the red line represents the total number of individuals engaging in the topic. The "Number of active users" illustrated by the dark blue line signifies those involved in discussions pertaining to positive information, while the "Number of passive users" represented by the light green line denotes individuals participating in discussions regarding negative information. Additionally, the "Number of neutral users" indicated by the orange line reflects participation in discussions centered on neutral information.

Empirical analysis of the comparative experiment depicted in Figure 2 reveals a significantly higher number of users engaged in discussions on negative and neutral information topics in the right figure compared to the left figure. This observation underscores the enhanced effectiveness of the game-based cellular automaton model (SEInR) in information dissemination, thereby enabling a more realistic simulation of information spread. These findings suggest that the SEInR model presented in this study outperforms the conventional cellular automaton model in multi-information dissemination scenarios.

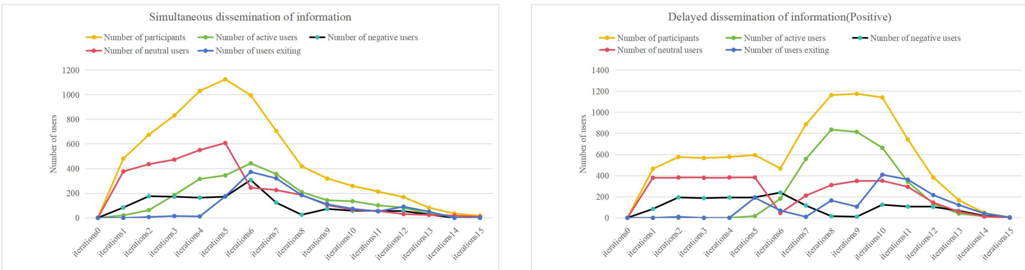
4.2. Information Dissemination with Different Strategy

In order to explore the impact of information placement at different time points on the dissemination of user node information, this article conducted a comparative experimental analysis based on the SEInR model. In the experiment, this article used user nodes at different time points as initial propagation nodes, and used different information dissemination strategies to compare the impact of different time points and information dissemination strategies to verify the method.

4.3. Dissemination with Different Time

A comparative experiment was conducted to assess the impact of information placement at different time points on information dissemination. It was designed as follows: (1) Simultaneous dissemination of Information: Positive, negative, and neutral types of information are concurrently disseminated. (2) Delayed dissemination of information (Positive): Initially, negative and neutral information is disseminated, with the dissemination of positive information being delayed.

The left portion of Figure 3a illustrates the results of the experiment where positive, negative, and neutral types of information were simultaneously delivered to observe information dissemination. Conversely, Figure 3b demonstrates the scenario where negative and neutral information were delivered first, followed by a three-round delay in the placement of positive information, thereby facilitating the observation of changes in information dissemination.



(**a**) Simultaneous dissemination (**b**) Delayed dissemination (Positive)

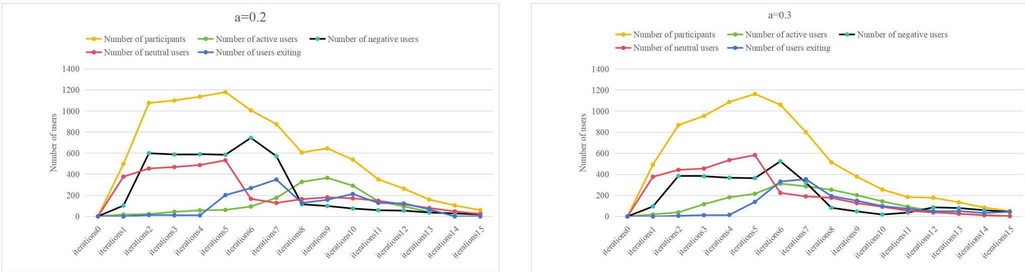
Figure 3. Comparison between simultaneous and delayed dissemination of information (Positive).

Through our comparative experiments and empirical analysis, it was found that the timing of information delivery plays a regulatory role in the game-like propagation of information. There is a certain correlation between the timing of delivering different types of information and their propagation speed. Additionally, mutual interactions among different types of information were observed. Our experimental results vividly demonstrate the complexity of the information game process, providing valuable insights for a deeper understanding of information dissemination mechanisms.

4.4. Dissemination with Different Thresholds

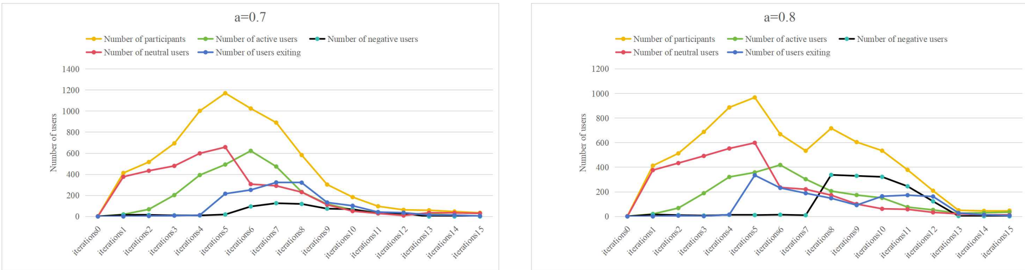
A comparative experiment was conducted to assess the impact of different thresholds on information dissemination. It was designed as follows:

The transition threshold of state transition was set to 70%. In order to observe the impact of threshold on information dissemination in SEInR model for users transferring from a state of not participating in a topic to a state of participating in a topic, two comparative experiments were conducted by adjusting the values of transition threshold. Figures 4 and 5 were created by simultaneously introducing both positive, negative, and neutral information with transition thresholds of 0.2, 0.3, 0.7 and 0.8.



(**a**) threshold is set 0.2 (**b**) threshold is set 0.3

Figure 4. Comparison between threshold (0.2) and threshold (0.3).



(**a**) threshold is set 0.7 (**b**) threshold is set 0.8

Figure 5. Comparison between threshold (0.7) and threshold (0.8).

As shown by Figure 4, it was found that the adjustment of the transition threshold had a moderating effect on the speed and duration of information dissemination, so a lower transition threshold can lead to more people participating in the topic and a longer duration of information dissemination. Moreover, as shown by Figure 5, it was found that adjusting parameter a has a certain moderating effect on the speed of information dissemination, but has no effect on the duration of information dissemination. In conclusion, combining the two experiments, it is shown that transition threshold has significant impact on the speed of information dissemination at lower values and also has a certain impact on the duration of information dissemination. While a higher transition threshold has a slight impact on the speed of information dissemination.

4.5. Inference

This article verifies the reliability of the deduction model through an empirical analysis of a public opinion event to verify the performance of the model. A comparative experiment was conducted between the evolution of positive and negative information user participation change curves using SEInR and the positive and negative information user participation change curves from real events. As shown by Figure 6, the cosine similarity between the curve of the true positive user participation and the evolution of positive user participation is 99.56%, while the cosine similarity between the curve of the real negative user participation curve and the evolution of negative user participation curve is 98.35%. It is proved that the game propagation model has a high accuracy in predicting the evolution trend of positive and negative participants.

In all, it is found that the principal factors influencing information dissemination in the online public opinion space included the domain of opinion, the node influence of the initial propagation node and the timing of information placement. These findings align with the propagation patterns observed in real online public opinion spaces, validating the reliability of the model proposed in this study through comparisons with historical event data. The model presented herein facilitates real-time monitoring of the popularity of topic events in the public opinion space, user emotional tendencies, user engagement in topic discussions and opinion dissemination status. By inferring the model's evolution process, predictions regarding the impact of public opinion topic events on cyberspace can be made, including assessments of the event's scope and its effects on user emotions.

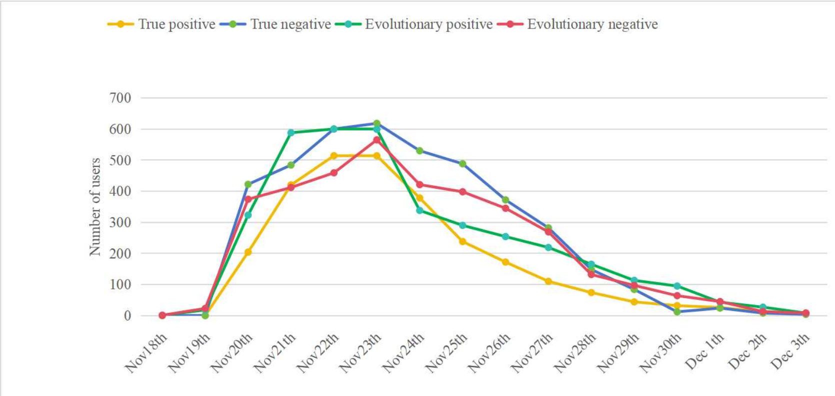


Figure 6. Comparison between true user participation and the evolution of user participation.

5. Conclusions

This paper is grounded in the principles of cellular automata and successfully constructs a sophisticated network model. It comprehensively addresses the dynamics of public opinion information dissemination and individual emotional tendencies. Through experimental verification using real user data from Weibo, it is demonstrated that the inference model (SEInR) proposed herein effectively captures changes in the state of user nodes within the cellular space and the dynamics between information during the inference process. The proposed model accurately mirrors real shifts in user behavior during public opinion event dissemination and offers a degree of predictability regarding event development, thereby creating an effective solution for complex network research.

In future endeavors, there is a potential to broaden the scope of the model, considering a more comprehensive array of factors both within and beyond the cellular space. Aiming to enhance the fidelity of simulations to real-world network public opinion dissemination scenarios. Moreover, there is room for enriching and refining the intricacies of information deduction rules, thereby shedding light on the intricate interactions among different types of information. Ultimately, this research endeavored to provide deeper insights into the field of public opinion analysis and to furnish a more scientifically grounded basis for formulating information dissemination strategies in cyberspace.

Author Contributions: C.S., F.S. and R.S. developed the initial idea for the study. C.S., X.L. and D.Y. designed the research methodology, including data collection, experimental procedures. L.Z. and K.J. created figures, tables of the presentation of results. C.S., F.S. and R.S. wrote the initial version of the manuscript. R.S. provided oversight and guidance throughout the research. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data generated or analyzed during this study are included in this published article.

Conflicts of Interest: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

References

1. De Domenico, M. More is different in real-world multilayer networks. *Nat. Phys.* **2023**, *19*, 1247–1262. [CrossRef]
2. Jin, C.; Yin, C.; Jin, X.; Min, Y.; Li, Y.; Chen, N.; Huang, J. Group-based rewiring rules of binary opinion competition dynamics. *Sci. Rep.* **2018**, *8*, 14423. [CrossRef] [PubMed]
3. China Internet Network Information Center. *Statistical Report on the Development of China Internet Network*; China Internet Network Information Center: Beijing, China, 2023.
4. Geng, L.; Zheng, H.; Qiao, G.; Geng, L.; Wang, K. Online public opinion dissemination model and simulation under media intervention from different perspectives. *Chaos Solitons Fractals* **2023**, *166*. [CrossRef]
5. Luo, Y.; Ma, J. The influence of positive news on rumor spreading in social networks with scale-free characteristics. *Mod. Phys. C* **2018**, emph29. [CrossRef]
6. Dai, L.; Shi, L.; Xie, G. Public opinion analysis of complex network information of local similarity clustering based on intelligent fuzzy system. *Intell. Fuzzy Syst.* **2020**, *39*, 1693–1700.
7. Jin, N. The Research on the Public Opinion Dissemination in Universities based on the Two-layer Coupling Network. In Proceedings of the International Conference on Frontiers of Electronics, Information and Computation Technologies (ICFEICT 2021), Changsha, China, 21–23 May 2021; p. 109.
8. Xiao, Y.; Chen, D.; Wei, S.; Li, Q.; Wang, H.; Xu, M. Rumor propagation dynamic model based on evolutionary game and anti-rumor. *Nonlinear Dyn.* **2018**, *95*, 523–539. [CrossRef]
9. Wu, B.; Yuan, T.; Qi, Y.; Dong, M. Public Opinion Dissemination with Incomplete Information on Social Network: A Study Based on the Infectious Diseases Model and Game Theory. *Complex Syst. Model. Simul.* **2021**, *1*, 109–121. [CrossRef]
10. Morita, S. Six Susceptible-Infected-Susceptible Models on Scale-free Networks. *Sci. Rep.* **2016**, *6*, 22506. [CrossRef] [PubMed]
11. Xiao, Y.; Zhang, L.; Li, Q.; Liu, L. MM-SIS: Model for multiple information spreading in multiplex network. *Phys. A Stat. Mech. Its Appl.* **2019**, *513*, 135–146. [CrossRef]
12. Heng, K.; Althaus, C.L. The approximately universal shapes of epidemic curves in the Susceptible-Exposed-Infectious-Recovered (SEIR) model. *Sci. Rep.* **2020**, *10*, 19365. [CrossRef] [PubMed]
13. Chen, H.; Song, Y.; Dan, L. Research on Cellular Automata Network Public Opinion Transmission Model Based on Combustion Theory. *J. Phys. Conf. Ser.* **2020**, *95*, 20–22. [CrossRef]
14. Liu, X.; Zhao, Q.; Wang, X.; Dong, X.; Li, Y.; Tian, Y. Iteratively Tracking Hot Topics on Public Opinion Based on Parallel Intelligence. *IEEE J. Radio Freq. Identif.* **2023**, *7*, 158–162. [CrossRef]
15. Liu, S.; Zhang, X. The Research of Public Opinion Inversion Based on Swarm Intelligence and Emergence. In Proceedings of the 2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Taizhou, China, 28–30 October 2023; pp. 1–6.
16. Barabási, A.L.; Albert, R. Emergence of scaling in random networks. *Science* **1999**, *286*, 509–512. [CrossRef]
17. Barabási, A.L.; Oltvai, Z.N. Network biology: Understanding the cell's functional organization. *Nat. Rev. Genet.* **2004**, *5*, 101–113. [CrossRef]
18. Watts, D.J.; Strogatz, S.H. Collective dynamics of 'small-world' networks. *Nature* **1998**, *393*, 440–442. [CrossRef]
19. Pastor Satorras, R.; Vespignani, A. Epidemic Spreading in Scale-Free Networks. *Phys. Rev. Lett.* **2001**, *86*, 3200–3203. [CrossRef] [PubMed]
20. Shao, F.; Sun, R.; Li, S. A model of Multi-Information Dissemination with suppressed action. *Complex Syst. Complex. Sci.* **2010**, *7*, 47–51.
21. Allen, L.J.S.; Bolker, B.M.; Lou, Y.; Nevai, A.L. Asymptotic profiles of the steady states for an SIS epidemic reaction-diffusion model. *Disc. Cont. Dyn. Syst.* **2008**, *21*, 1–20. [CrossRef]
22. Xuan, W.; Ren, R.; Pare, P.E.; Ye, M.; Ruf, S.; Liu, J. On a Network SIS Model with Opinion Dynamics. *IFAC Pap. Online* **2020**, *53*, 2582–2587. [CrossRef]
23. Lv, X.; Fan, D.; Li, Q.; Wang, J.; Zhou, L. Simplicial SIR rumor propagation models with delay in both homogeneous and heterogeneous networks. *Phys. A Stat. Mech. Its Appl.* **2023**, *627*, 129131. [CrossRef]
24. Zhang, L.; Su, C.; Jin, Y.; Goh, M.; Wu, Z. Cross-network dissemination model of public opinion in coupled networks. *Inf. Sci.* **2018**, *451*, 240–252. [CrossRef]
25. Han, S.C.; Liu, Y.; Chen, H.L.; Zhang, Z.J. Influence Model of User Behavior Characteristics on Information Dissemination. *Int. J. Comput. Commun. Control* **2016**, *11*, 209–223. [CrossRef]
26. He, S.; Peng, Y.; Sun, K. SEIR modeling of the COVID-19 and its dynamics. *Nonlinear Dyn.* **2020**, *101*, 1667–1680. [CrossRef] [PubMed]
27. Bjørnstad, O.N.; Shea, K.; Krzywinski, M.; Altman, N. The SEIRS model for infectious disease dynamics. *Nat. Methods* **2020**, *17*, 557–558. [CrossRef] [PubMed]
28. Li, Q.; Du, Y.; Li, Z.; Yan, H.; Hu, J.; Hu, R.; Lv, B.; Jia, P. HK–SEIR model of public opinion evolution based on communication factors. *Eng. Appl. Artif. Intell.* **2021**, *100*, 104192. [CrossRef]
29. Wolfram, S. Cellular Automata as Simple Self-Organizing Systems. *Comput. Sci. Phys.* **1982**.
30. Ilachinski, A. *Cellular Automata: A Discrete Universe*; World Scientific: Singapore, 2001.
31. Mitchell, M. *The Genesis of Cellular Automata*; Oxford University Press: Oxford, UK, 1993.

32. Su, Y.; Wo, Y.; Han, G. Reversible cellular automata image encryption for similarity search. *Signal Process. Image Commun.* **2019**, *72*, 134–147. [CrossRef]
33. Ping, P.; Wu, J.; Mao, Y.; Xu, F.; Fan, J. Design of image cipher using life-like cellular automata and chaotic map. *Signal Process.* **2018**, *150*, 233–247. [CrossRef]
34. Xu, X.; Fan, C.; Wang, L. A deep analysis of the image and video processing techniques using nanoscale quantum-dots cellular automata. *Optik* **2022**, *260*, 169036. [CrossRef]
35. Sun, Z.; Chen, Z.; Zheng, J. Ship interaction in narrow water channels: A two-lane cellular automata approach. *Phys. A.* **2015**, *431*, 46–51. [CrossRef]
36. Qi, L.; Zheng, Z.; Gang, L. Marine traffic model based on cellular automation: Considering the change of the ship's velocity under the influence of the weather and sea. *Phys. A* **2017**, *483*, 480–494. [CrossRef]
37. Qi, L.; Ji, L.; Balling, R.; Xu, W. A cellular automaton-based model of ship traffic flow in busy waterways. *J. Navig.* **2021**, *74*, 605–618. [CrossRef]
38. Hegselmann, R.; Krause, U. Opinion dynamics and bounded confidence models, analysis, and simulation. *J. Artif. Soc. Soc. Simul.* **2002**, *5*, 1–33.
39. Krause, U. A discrete nonlinear and non-autonomous model of consensus formation. *Commun. Differ. Differ. Equ.* **2002**, *7*, 227–236.
40. Liu, X.; Cao, S.; Zheng, L.; Gong, F.; Wang, X.; Zhou, J. POCA4SD: A Public Opinion Cellular Automata for Situation Deduction. *IEEE Trans. Comput. Soc. Syst.* **2020**, *8*, 201–213. [CrossRef]
41. Liu, X.; Cao, S.; Cao, Y.; He, J.; Zhang, W.; Wang, X.; Zheng, L. Online Public Opinion Deduction Based on an Innovative Cellular Automata. In Proceedings of the Cyberspace Data and Intelligence, and Cyber-Living, Syndrome, and Health, Beijing, China, 16–18 December 2019; Volume 1137, pp. 141–160.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.