mathematics

MDPI

Article

Information Dissemination Model Based on Social Networks Characteristics

Jianwei Ding [1], Zehan Li [2], Xia Wu [3], Rong Liu [3] and Hangyu Hu [3,*]

[1] 30th Research Institute of China Electronics Technology Group Corporation, Chengdu 610041, China
[2] School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China
[3] School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
* Correspondence: huhy@uestc.edu.cn

Abstract: As a crucial platform, online social networks provide individuals with avenues to exchange ideas and access information, exerting profound impacts on society and nations. In social networks, key users, serving as edge nodes in the process of information dissemination, play a pivotal role because they directly connect users and can process and forward information in real-time. Furthermore, edge nodes enable personalized information dissemination based on users' social relationships and behavioral characteristics, more accurately reflecting the pathways and influence of information spread. Early research primarily focused on the dynamics of information dissemination in complex networks, aiming to develop general predictive models to understand the overall mechanisms of information spread. However, there is still a lack of research on how the unique social relationships and attributes in social networks affect information dissemination. To address this gap, we conducted an in-depth study of the characteristics of information dissemination in social networks and improved the classic independent cascade model, proposing a novel predictive model for information spread. This enhancement not only improves the accuracy of simulating the information dissemination process in social networks but also demonstrates that our proposed model significantly outperforms other models in terms of accuracy. The findings provide a more effective tool for understanding and predicting information dissemination in social networks.

Keywords: social network; information diffusion; complex networks

MSC: 91D30; 68T01

check for updates

Academic Editor: Rossana Mastrandrea

Received: 25 February 2025
Revised: 4 April 2025
Accepted: 8 April 2025
Published: 10 April 2025

Citation: Ding, J.; Li, Z.; Wu, X.; Liu, R.; Hu, H. Information Dissemination Model Based on Social Networks Characteristics. *Mathematics* **2025**, *13*, 1254. https://doi.org/10.3390/math13081254

Copyright: © 2025 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

The advent of the Internet, coupled with the growth of various social media platforms, has cemented social networks as a key component of contemporary society. These platforms serve not only as conduits for sharing and expressing opinions, but also deeply permeate real-world societal dynamics. Scholarly research on the mechanisms of information diffusion within social networks has gained paramount importance, particularly in the context of preserving societal stability.

In social networks, key users act as edge nodes and play a crucial role in information dissemination. Serving as bridges within the network, these users, through their extensive social connections and frequent interactions, directly influence the pathways and speed of information spread. Because they can reach users across different social circles, information

can rapidly diffuse through these key users to a broader network, significantly enhancing the efficiency and coverage of information dissemination. In addition, the personalized dissemination strategies of these edge nodes allow the information to reach the target user groups more accurately, further improving the effectiveness of the spread of information.

However, existing information diffusion models face several limitations when applied to real-world social networks. Traditional models often assume idealized and static network structures, neglecting the dynamic and evolving nature of social networks. These models typically fail to account for the intricate interplay between user behaviors, network topologies, and external factors, leading to inaccurate predictions and limited applicability. Moreover, many models focus on centralized processing, which struggles to meet the demands of real-time information dissemination in large-scale networks. The computational complexity and resource consumption of these models hinder their deployment in edge devices, where low latency and efficient processing are essential. Furthermore, the structural characteristics of information cascades, critical for the identification and management of sensitive information, are often oversimplified or overlooked, reducing the effectiveness of existing models in capturing the true nature of information diffusion.

To address these challenges, this article explores various aspects of user behavior within social networks and delves into the study of information diffusion processes. By examining these processes, we aim to discern the underlying laws governing information spread. A key aspect of this research involves analyzing the structures of information cascades which emerge from different types of information diffusion. Such an analysis is pivotal for identifying sensitive information.

To ensure that the model can operate efficiently on resource-constrained edge devices, we have imposed limitations on the model's size. By streamlining the model architecture and optimizing parameter configurations, we have effectively reduced the computational and storage requirements while maintaining the model's predictive accuracy. This approach not only enables the model to run smoothly on edge devices, but also enhances real-time processing capabilities, meeting the demands for low latency. Despite the size limitations, the method has demonstrated excellent performance in edge computing environments, proving its feasibility and effectiveness in practical applications.

In this study, we explore the dynamics of information diffusion within social networks. Through an in-depth analysis of information dissemination patterns, we have developed a refined model that more accurately mirrors the inherent laws of information spread. Central to our investigation is the examination of how a user's position within a social network and the utilization of edge devices influence the scope of information diffusion. Our empirical results reveal a notable correlation between the centrality of the initiating user's entry and the extent of dissemination of information. Specifically, we discovered that when a node with minimal entry centrality in a social network serves as the starting point for our model, the dissemination of information is limited. In contrast, initiating diffusion from a node with slightly higher entry centrality and leveraging the computational power of edge devices leads to a more expansive spread of information across various timelines.

2. Related Work

Recent scholarly efforts in the field of information diffusion within social networks have mainly been grouped into two distinct but interconnected domains. Firstly, there is a growing focus on constructing and scrutinizing the processes underpinning information dissemination. This involves a detailed examination of how information propagates through various channels within a network [1]. Secondly, there is an increasing emphasis on analyzing the architecture of social networks themselves, delving into how their inherent structural attributes influence the patterns and efficacy of information diffusion. This dual

approach offers a comprehensive understanding of the dynamics at play in social networks and other digital platforms where information dissemination is crucial.

Research in the field of information dissemination is multifaceted and includes several key aspects. A significant contribution is the construction of information diffusion cascade networks. Vosoughi et al. [2] introduced a time-informed method for constructing these networks on Twitter, a technique that has gained traction in academic circles. Another vital aspect is the analysis of parameters within existing propagation cascade networks, such as their depth and breadth.

Furthermore, the structure of these networks is analyzed to simulate the information transmission process using various mathematical models. Zaman et al. [3] employed a Bayesian approach to model this process, establishing a probabilistic framework to predict the eventual reach of the information. Similarly, Ruchjana et al. [4] and Lobel et al. [5] utilized Markov processes for simulation purposes. Shen et al. [6] proposed an improved Poisson process-based generative probability framework to model the popularity of information, significantly improving the prediction capabilities of volumes of information dissemination.

Moreover, Vu, Zhao, and Rogers [7–9] have all applied Hawkes processes to simulate the growth of information cascades. Beyond these traditional methodologies, neural networks have emerged as a powerful tool for modeling information diffusion. Hernandez et al. [10] posited that the voluminous datasets generated by user activities on social networks could facilitate various predictive analyses, including experiments predicting social network platform user behaviors through deep neural network models.

In the realm of social network structure and its impact on information dissemination, research is conducted from various angles. Firstly, the parameters and architecture of social networks themselves are analyzed. Dey et al. [11] introduced a motif-based approach for assessing the reliability of complex networks, deducing that specific motifs can serve as metrics for network reliability assessment.

Secondly, the link between the characteristics of social network nodes and information diffusion is explored. Namtitha et al. [12] developed a technique to pinpoint the most pivotal propagators in a network by utilizing its global structural features. They postulated that the influence of nodes varies with the network's connection patterns, particularly the size and density of network clusters.

Furthermore, the development of models for information propagation in social networks is a key focus. This includes models akin to those in infectious disease research [13–15], influence models [2,16], and their derivatives, such as the SVFR model [17] and the SEI model [18]. Although these models vary in their simulation methodologies for information dissemination, they are instrumental in scenarios such as simulating real-world information spread and validating network node influence outcomes. The application of the SIR model to validate node rankings based on their influence in information dissemination is highlighted in [19].

Existing research in the field of information diffusion law primarily concentrates on the dynamics of diffusion within complex networks, proposing a universal model for information dissemination and summarizing the diffusion laws. However, this approach often overlooks the significant influences of distinct social relationships and the inherent attributes of social networks on the process of information diffusion. The homogenization of network models in the analysis of social media data can lead to the omission of critical information, adversely affecting the accuracy of the resulting classifications. In response to these limitations, our study proposes the construction of a heterogeneous network that integrates user information, user interaction data, and textual content. This approach

aims to enrich the network information and enhance the precision of account classification, addressing the complexities inherent in social media data.

3. Developing and Analyzing a Social Network-Based Information Diffusion Model for Transmissibility

Based on the analysis of social networks, we focus on the critical role of edge nodes in these networks. Key attributes were extracted from real-world information dissemination processes, which aid in constructing an information diffusion model designed to closely simulate these processes. Additionally, we introduce a novel method for calculating the impact of dissemination, which is used to define the spread ability within social networks. Experimental results demonstrate that our model effectively replicates the dynamics of actual information dissemination. In particular, compared to existing methods, our approach to calculating the impact of dissemination more accurately identifies the user nodes with the greatest influence on the spread of information within social networks. Figure 1 shows the general framework of the proposed method.

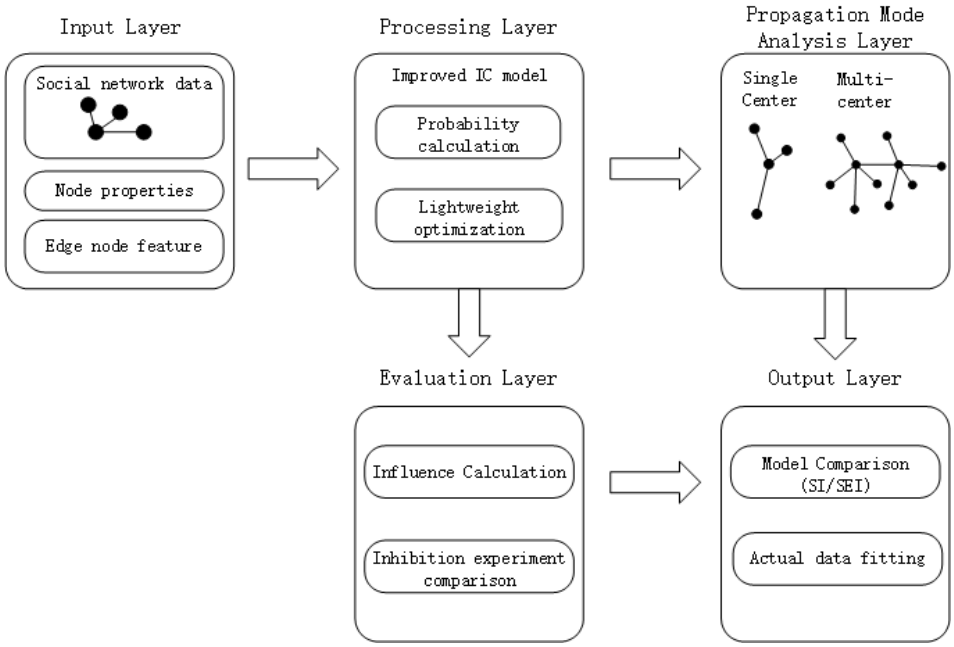
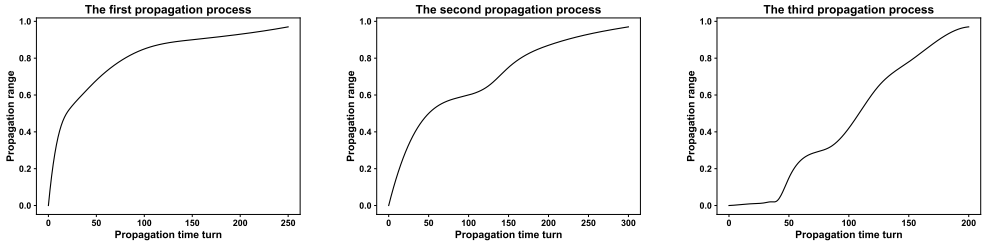


Figure 1. Overall framework of the proposed method.

3.1. Extraction and Analysis of Information Diffusion Patterns in Social Networks

In the realm of social networks, the transmission of information is subject to the influence of various network characteristics. These characteristics encompass the size and structure of the social network, the attributes associated with network nodes, and the dynamic nature of information propagation. It is imperative to consider these diverse characteristics when constructing a model for information dissemination and to ground this analysis in empirical data pertaining to information transmission.

To gain a deeper understanding of the impact of these characteristics on information diffusion, it is essential to leverage real-world data. By scrutinizing the cumulative distribution curve of actual information dissemination under different circumstances, we discern three distinct types of information propagation processes. Figure 2 illustrates the cumulative distribution curves that correspond to these three distinct propagation phenomena.



(**a**) (**b**) (**c**)

Figure 2. Three types of information propagation processes: (**a**) The first propagation process. (**b**) The second propagation process. (**c**) The third propagation process.

The three figures in Figure 2 correspond to the cumulative distribution curves of the nodes of the three kinds of information diffusion. The horizontal axis in the figure represents the time elapsed during the information propagation, with each unit on this axis corresponding to a six-minute interval, effectively delineating propagation rounds. The vertical axis depicts the proportion of nodes within the largest connected component of the social network.

Our study investigates three unique information diffusion scenarios, each reflective of varying dynamics in the communication process:

Uniform Diffusion: In this scenario, we observe a consistent and uninterrupted flow of information without distinct stages of promotion or hindrance. The propagation maintains a steady pace throughout.

$$P_{\text{cum}}(t) = 1 - (1 - p_0)^t \tag{1}$$

Accelerated Diffusion: Here, the information propagation experiences sudden acceleration at a specific point in time. Certain nodes play a prominent role in promoting the information, resulting in an abrupt increase in the propagation rate.

$$P_{\text{cum}}(t) = 1 - \prod_{k=1}^{t} (1 - \min(p_0 + \alpha k, 1)) \tag{2}$$

Hindered Diffusion: In this case, the communication process encounters an initial stage of stagnation with minimal growth. Unlike the first scenario, there is an evident obstruction that impedes the early phases of communication.

$$P_{\text{cum}}(t) = \int_0^t p_0 e^{-\lambda\tau}\, d\tau = \frac{p_0}{\lambda}\left(1 - e^{-\lambda t}\right) \tag{3}$$

The representation of these three communication processes provides valuable insights into the diverse dynamics at play during information diffusion within social networks.

The disparities observed in the cumulative distribution curves of propagation in Figure 2 can be attributed to distinct underlying factors.

In the initial phase of the first propagation process, information originates from user nodes with a substantial following within the actual social network. Consequently, it experiences rapid dissemination through its initial followers. However, as time elapses, the temporal relevance of the information diminishes, user interest wanes, and the pace of information transmission decelerates.

In the second propagation process, a similar pattern emerges, with information initially emanating from user nodes with a substantial following. However, during the early stages of dissemination, user engagement with the information is relatively low, leading to a

slower rate of propagation. At a critical inflection point, nodes with a significant number of subscribers begin disseminating the information. This action effectively revitalizes the information's temporal relevance, piquing the interest of other users and accelerating the dissemination process.

Conversely, in the third propagation process, information initially originates from a node with only a few followers. Consequently, the dissemination rate remains sluggish during the initial phase. It is only at the first inflection point that a node with a substantial following takes charge of propagating the information, resulting in an acceleration of the dissemination process.

According to the analysis conducted, several key conclusions can be drawn:

Effect of Information Transmission Speed: It has been observed that the rate of information transmission within a network has a significant impact on the probability of nodes successfully transmitting information over time. A slower information transmission speed is associated with a decrease in the likelihood of successful information propagation.

Influence of Node Subscribers: Nodes with a substantial number of subscribers play a crucial role in the information dissemination process. When such nodes initiate the propagation of information, it significantly increases the probability of subsequent nodes successfully transmitting the same information.

Building upon the above findings, we can further delve into the impact of various factors on the probability of nodes disseminating information within social networks. These factors can be categorized into three main groups:

a. Network Characteristics: This category encompasses factors related to the size and structure of the social network. These include the topology of the network, its connectivity, and the density of connections between nodes.
b. Node Attributes: Factors in this category pertain to the characteristics of individual nodes within the network. These may include the duration of exposure to information, the current volume of information being spread, the concentration of information in the proximity of nodes, the spatial positioning of nodes within the network, and the level of activity exhibited by nodes.
c. Information Transmission Status: This category addresses the overall status of information transmission within the network. It involves factors such as the current state of information dissemination, the pace of information flow, and the patterns of interaction among nodes.

By considering these three broad categories, we can comprehensively assess the multitude of factors influencing the likelihood of nodes transmitting information in social networks. This structured approach allows for a more systematic analysis of the intricate dynamics governing information dissemination.

indegree: This parameter quantifies the degree of node centrality within the current network of interest. It represents the number of incoming connections or followers that a node possesses when the scope of the social network is restricted. In essence, in-degree reflects the significance of a node within the designated network and directly influences the likelihood of information dissemination by the node;

follower_count: This parameter corresponds to the user attribute of the node and denotes the total number of followers that the user has in the real-world social network. In practical social networks, the follower count plays a pivotal role in determining the reach of information dissemination initiated by the user node. When one or more nodes (where 'x' is a constant) disseminate information at any given time, it elevates the probability of other nodes participating in the information diffusion process; activitiness: The activity level of user nodes associated with a particular node is derived from a statistical analysis of user behaviors within a specified time frame in the actual social network. In social

networks, higher user activity levels directly correlate with an increased probability of information propagation; t_expose: This parameter quantifies the duration of time a node remains exposed to a specific piece of information. The probability of the node engaging in information dissemination is inversely proportional to this exposure time; active_node_count: The current volume of ongoing information dissemination and the duration of the dissemination collectively influence the probability of node participation in information diffusion. The ratio of active nodes to time active_node_count/ t can be regarded as the rate of information propagation within a unit of time. The probability of node engagement in dissemination is directly proportional to this rate.

3.2. Calculation of Propagation Probability in Social Networks

The methodology section of this paper presents enhancements to the Information Cascades (IC) model. Our improved information propagation model is designed to closely mimic real-world social networks and can accurately simulate the cumulative distribution curve observed in actual network data. Consequently, we select the same initial node for information propagation simulation as that observed in real-world information dissemination scenarios.

Building upon the insights gained from our earlier discussion on the factors influencing information transmission probabilities, this section forms the foundation for calculating these probabilities through their integration. The computation of information propagation probability is predicated on the assumption that only when node v directs its attention towards node u, does node v possess the potential to propagate information originating from node u. The probability of node v transmitting information from node u can be accurately determined through the utilization of relevant characteristic parameters.

In accordance with the previously defined characteristics that impact information transmission probabilities, this section synthesizes these factors to compute the probability of information transmission. We operate under the assumption that information can only propagate from node u to node v when node v actively engages with node u. The precise calculation of the probability of node v disseminating information from node u is contingent upon the incorporation of pertinent characteristic parameters.

$$p_{uv} = \frac{C}{t_\text{expose}} \times \text{user_status} \times \text{info_status} \times \text{factor}_t \tag{4}$$

In Equation (4), C represents constant, and t_expose denotes the time at which the node becomes exposed to the information. This time corresponds to the current round of the node and the moment when the concerned node initially disseminates this message. user_status signifies the impact of the node's intrinsic attributes and its position within the network of interest (e.g., activity level, degree of centrality within the network of interest, etc.) on the propagation probability. info_status represents the influence of the information's status (comprising the total propagation count and the current propagation round) on the propagation of the node. factor_t encapsulates the effect of nodes that have disseminated the information prior to the current time on other nodes.

$$\text{user_status} = \log\left(\frac{\text{indegree}}{\sum \text{indegree}} + 1\right) \times f(\text{activitiness}) \tag{5}$$

Equation (5) delineates the quantity of individuals who exhibit an interest in the focal user within the contemporary social network. To ensure that the logarithmic function yields a positive value, constant terms are introduced. This value is indicative of the node's level of activity within the real-world social network, specifically denoting the dynamic count within a unitary time frame. The formulation presented above is the

outcome of amalgamating prior analyses regarding the attributes of nodes in the context of information dissemination.

$$\text{info_status} = f\left(\frac{\text{active_node_count}}{t}\right) \times \text{info_density} \tag{6}$$

In Equation (6), active_node_count denotes the cumulative propagation quantity of the current information, while t signifies the current propagation iteration. info_density represents the ratio of nodes that have disseminated information among the nodes pertinent to the current node, signifying the concentration of information.

$$factor_t = factor_{t-1}(1+\alpha)^{n_{t-1}}, 0 < \alpha < 0.01 \tag{7}$$

$$\text{factor}_{\,0} = 1 \tag{8}$$

In the context of real-world social networks, when a total of $\sum_0^{t-1} n_i$ users with a significant $follower_count$ simultaneously propagate the same message before time t, these nodes collectively influence the information dissemination process beyond time t. To model the impact of nodes that have participated in information propagation at time t on subsequent nodes, we introduce the parameter $factor_t$, where α represents the growth coefficient of influence.

Given that the experimental network is generated to emulate the actual information propagation process, it can be assumed that all nodes within the network are interested in the information being disseminated. For the purpose of this discussion, we temporarily omit consideration of the specific level of interest that nodes have in the information.

This section has provided a comprehensive explanation of the information dissemination model and has outlined the probability calculation associated with information propagation. Building upon the foundation of the information dissemination model, we can proceed to simulate information propagation using real-world social network data.

3.3. Analysis of Transmissibility in Social Networks

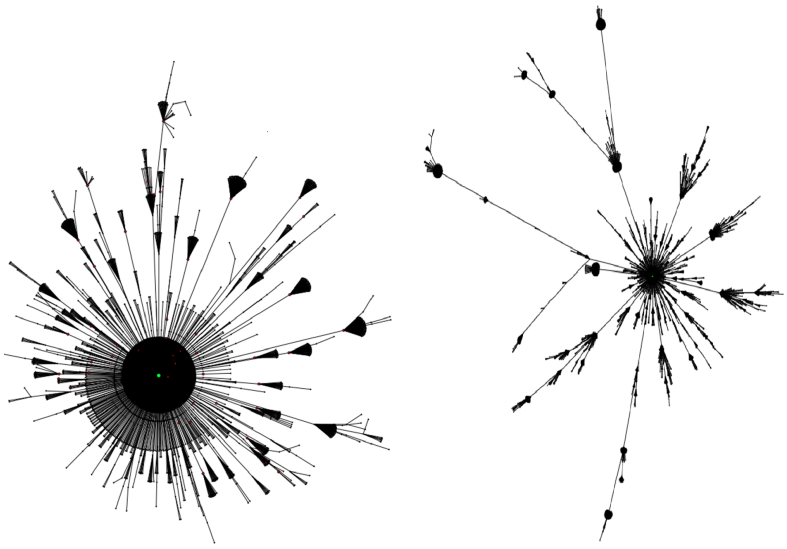
3.3.1. Propagation Modes Generated by Information Diffusion Models

In accordance with the information propagation model as outlined in this research paper, when conducting propagation simulation experiments with various initial nodes, numerous propagation cascades are concurrently generated. These cascades exhibit substantial disparities in their structural characteristics and statistical parameters. These distinctions serve as indicators of diverse propagation modes and the varying propagation capabilities associated with different nodes selected as starting points. This section will exemplify these principles using empirical data.

In the utilization of the information propagation model for propagation experiments, nodes with distinct centrality of penetration are chosen as the starting nodes for experimentation, thereby generating multiple propagation cascades. Although these cascades adopt a tree-like structure, discernible disparities persist among them. As illustrated in Figure 3, the depicted propagation cascades originate from two different initial nodes within the same social network. In the left figure, the entry centrality value of the information propagation starting node stands at 0.591, while in the right figure, it is 0.033.

In Figure 3, the green nodes symbolize the information propagation starting nodes, while the remaining black nodes represent nodes propagated in subsequent simulation processes. The connecting edges denote the propagation paths generated during the simulated information propagation. Notably, when selecting the initial node, the entry centrality of the initial node in the left figure significantly surpasses that in the right figure.

Furthermore, it is discernible that the structural characteristics of the two cascaded networks markedly differ. Specifically, in the left figure, the information within the propagation cascade predominantly propagates from the initial node as its focal point, with most nodes in the entire cascade concentrated around this initial node, resulting in a shallow cascade depth. Conversely, in the right figure, the information within the propagation cascade disseminates from multiple centers, with the starting node being just one among several central nodes, leading to a greater cascade depth.



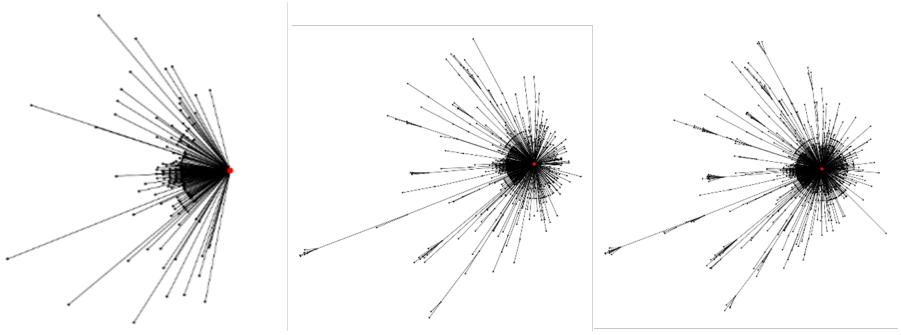
(**a**) (**b**)

Figure 3. Two different propagation cascade: (**a**) the entry centrality value of the information propagation starting node stands at 0.59. (**b**) The entry centrality value of the information propagation starting node stands at 0.033.

Figures 4 and 5 depict the temporal evolution of two distinct communication modes, which we refer to as 'single center divergence' and 'multi-center divergence' throughout this paper. These graphical representations elucidate the progression of these modes over time with greater clarity.

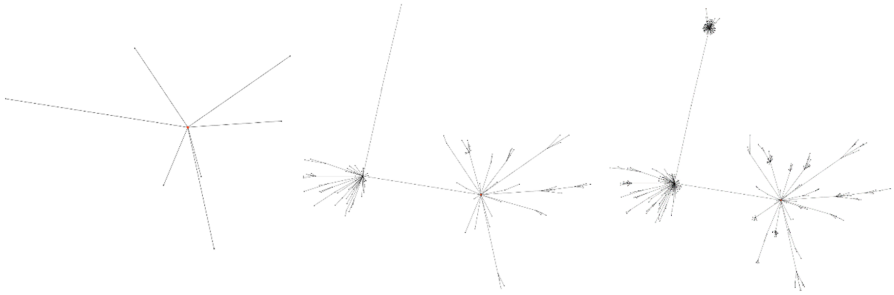
In the context of the first communication mode, 'single center divergence', the majority of nodes that can be reached by the initial node are concentrated in close proximity to the source node. Conversely, the second communication mode, 'multi-center divergence', exhibits a more dispersed distribution of nodes.

The observed patterns of information communication cascades align closely with these two communication modes. Our analysis leads to a significant conclusion: When an influential node at the core of the social network serves as the point of origin for information dissemination, it exerts a direct influence on its immediate neighbors. Even in the absence of other highly communicative nodes in its vicinity, this central node can independently propagate information. In contrast, nodes with limited communication capabilities and situated away from the social network's core lack the capacity to significantly impact their surroundings. They must rely on other more communicative nodes to facilitate the dissemination process. This fundamental disparity forms the basis for the divergence between the two communication modes observed within cascading networks.



(**a**) time slice 1 (**b**) time slice 2 (**c**) time slice 3

Figure 4. Process of a "single-center divergent" communication model in different time slices.



(**a**) time slice 1 (**b**) time slice 2 (**c**) time slice 3

Figure 5. Process of a "multi-center divergent" communication model in different time slices.

When examining the communicative impact of nodes within social networks, it becomes imperative to consider the influence of two distinct modes of communication. Specifically, the positioning of a node within the social network exerts a significant influence on the manner in which it serves as the originating point for the dissemination of information to other nodes. In cases where certain nodes are intentionally designated to direct the flow of information within the social network, it becomes necessary to devise strategic approaches aimed at exerting influence over these pivotal nodes. By doing so, the objective is to expedite the dissemination of information throughout the social network, thereby achieving a more expeditious and cost-effective information propagation process.

3.3.2. Definition of Node Transmissibility in Social Networks

Based on the aforementioned propagation model, this section outlines the methodology employed in utilizing the model to simulate propagation and generate multiple propagation processes for the purpose of elucidating the transmissibility of nodes within a social network. In this paper, the concept of node transmissibility is delineated from two distinct perspectives: (1) the transmissibility of a node when serving as the initial information source and (2) the transmissibility of a node during the process of information dissemination.

The assessment of node transmissibility as initial information sources involves the selection of nodes situated at various locations within the network to act as the starting points for information dissemination. Propagation simulation experiments are then conducted using the chosen models. These experiments serve to contrast the differing propagation extents of information originating from distinct starting nodes within a predefined time frame. Additionally, they facilitate an examination of the varying time intervals required for information to propagate to specified ranges from diverse starting nodes, each associated with a predetermined propagation range size.

In this experimental setup, multiple nodes, characterized by distinct penetration centrality values within the social network, are designated as the initial information sources.

The chosen information propagation model is employed to conduct these propagation experiments. The experiments are executed repeatedly, and the resultant average propagation time, as influenced by the changing attributes of the starting nodes, serves as a critical evaluation criterion.

This approach enables a comprehensive analysis of how different nodes, based on their penetration centrality, impact the dissemination of information throughout the entire network. It offers insights into the dynamics of information propagation and the role played by various nodes in facilitating or hindering the spread of information within the social network.

Transmissibility of Nodes in the Information Dissemination Process: In the context of the social network information dissemination model, this parameter characterizes the transmissibility of individual nodes. This research quantifies the node's ability to propagate information to other nodes during the information dissemination process and denotes this transmissibility as S_a. It serves as a fundamental metric in evaluating the node's influence and effectiveness in the information diffusion within the social network. The calculation method is shown in Equation (9):

$$S_a = \frac{\text{count}_{out}}{\text{count}_{in}} \tag{9}$$

In which *count_in* indicates the number of times a node propagates information in the process of information dissemination. Click count_out refers to the number of times the information is spread. To measure the transmissibility of nodes in the information dissemination process, a small number of nodes with different degrees of centrality in the social network are used as the starting information nodes. The information dissemination model is used for the dissemination experiment. The information dissemination cascade obtained from the experiment is used as the data. The statistics in Equation (9) are carried out for all nodes in the dissemination cascade to obtain the transmissibility of each node in the social network in the information dissemination process, that is, the dissemination influence. According to the method of Equation (9), a value greater than or equal to 0 can be calculated for each node in the social network to indicate the number of times that a node can cause other nodes to spread information by spreading a piece of information.

4. Result

4.1. Analysis of Simulation Results of Information Diffusion in Social Networks

The concluding section of this research harnesses the information propagation model, as previously defined, to emulate the dissemination of information within social networks. It commences by utilizing the identical initial node as in the real-world information propagation scenario, simulating the propagation of factual information. Moreover, it employs nodes situated at various positions within the network as the origin points for information dissemination, simulating the cascading effect of multiple information propagation instances. This section undertakes a comprehensive analysis of these simulated data while concurrently assessing the model's efficacy. Particular attention is devoted to investigating the influence of the initial node's spatial location on the transmission process.

In the process of evaluating the performance of our propagation model, it becomes imperative during the creation of cumulative distribution curves to define the vertical axis as the percentage of information diffusion across nodes. This consideration arises from the discrete nature of nodes within the social networking framework discussed earlier. It is crucial to incorporate this granularity to ensure that it does not impede the evaluation of the model's performance. Additionally, a comparative analysis with other established models is undertaken to gauge the fidelity of the model simulation presented in this

paper. Specifically, the communication model proposed in this study is juxtaposed with the SI model, the SEI model, and the communication dynamics of real-world information propagation. The comparative findings are visually represented in Figure 6.

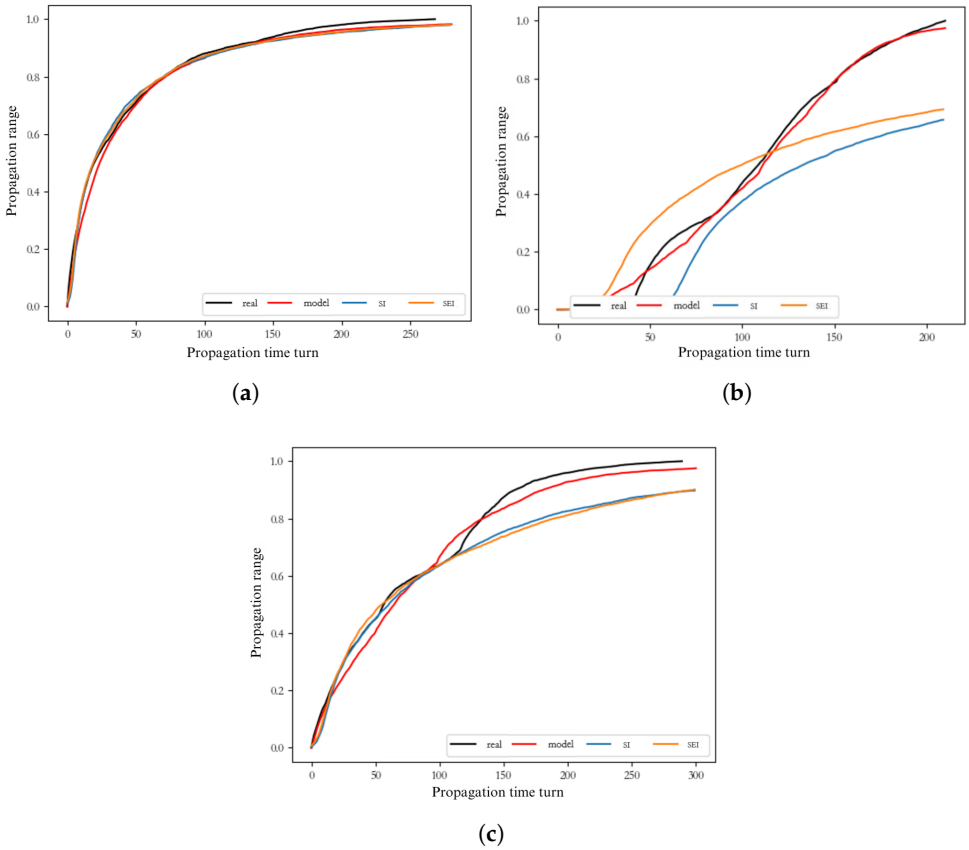


Figure 6. (**a**–**c**) are the comparison results with the actual information dissemination.

By conducting a comparative analysis between the empirical data on information dissemination and the cumulative distribution curves derived from three models, the Susceptible-Infectious (SI) model, the Susceptible-Exposed-Infectious (SEI) model, and the novel model proposed in this study, several key conclusions can be drawn.

In scenarios where the communication process unfolds without distinct stages of promotion or hindrance, the cumulative distribution curves obtained from our proposed model, the SI model, and the SEI model exhibit a remarkable degree of consistency. This convergence suggests that in cases where information transmission remains unhindered, the designated social network adheres to the principle of 'fast first and slow second,' implying an initial rapid dissemination followed by a gradual slowdown.

Conversely, in situations characterized by pronounced promotional phases, information propagation experiences an abrupt acceleration at a specific point in time. Notably, neither the SI model nor the SEI model adequately captures this phenomenon. This limitation arises from the fixed propagation probabilities inherent in the SI model and the reliance on probability-based state transitions in the SEI model. In contrast, our proposed model successfully simulates the observed scenario where information transmission accelerates from a slow to a rapid pace at a specific temporal juncture.

Finally, when considering information transmission processes marked by obstructive phases, the initial stages exhibit minimal growth, and information can only be relayed to other nodes once it reaches a particular node at a specific time. In such instances, the SI and SEI models fail to represent this intricate process accurately. Our proposed model, however,

accounts for this nuanced phenomenon, reflecting the delayed propagation observed when specific conditions are met.

In conclusion, our study underscores the importance of tailoring modeling approaches to the distinct characteristics of information dissemination processes, acknowledging that traditional models such as SI and SEI models may not capture the full spectrum of dynamics observed in real-world social networks. Our novel model offers a more comprehensive framework for understanding information transmission behaviors, encompassing scenarios of unhindered flow, accelerated propagation, and obstructed pathways, thereby contributing to a deeper understanding of complex communication dynamics within social networks.

By conducting a comparative analysis of the cumulative distribution curves, we can draw the following conclusions. In contrast to the SI model, the model proposed in this study demonstrates a remarkable capacity to closely replicate the cumulative distribution curve of real-world information transmission. This finding underscores the model's enhanced fidelity to the inherent propagation dynamics of actual information dissemination. Furthermore, our experimentation results corroborate that by adjusting the probability of information transmission in accordance with the relevant characteristics, the information transmission model achieves a higher degree of accuracy in simulating the dynamics of real-world information transmission. Consequently, the simulated information generated by substituting the initial information node also exhibits heightened authenticity.

4.2. Transmissibility of Nodes as Starting Points for Information Diffusion

When evaluating the ability of nodes to spread information, this paper selects different degrees of centrality as the basis for selecting the starting node of information propagation, groups the nodes in the network according to the degrees of centrality value, selects a number of nodes with different degrees of centrality values as the starting nodes of propagation, uses the propagation model to conduct propagation experiments, and analyzes the propagation results obtained by using different nodes as the starting nodes and finds the relationship between the degree of centrality of penetration and the time required for the node to act as the propagation starting node and the specified propagation range that the information can reach. The experimental results are shown in Figure 7.

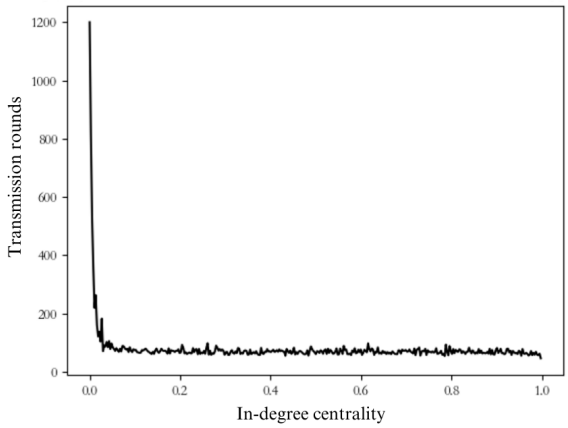


Figure 7. Transmission rounds required for entire network using nodes with different degree centrality.

When assessing the efficacy of nodes in disseminating information, this study adopts varying degrees of centrality as the foundational criteria for selecting the initial node for information propagation. The network nodes are categorized based on their centrality values, and a subset of nodes with diverse centrality values is chosen as the initiators of the propagation process. Subsequently, a propagation model is employed to execute

propagation experiments, and the results obtained from employing distinct nodes as initiators are scrutinized. These analyses aim to elucidate the interplay between centrality degree, the time required for a node to assume the role of an initiator, and the designated range within which the information can propagate effectively. The empirical findings are visually depicted in Figure 7.

Furthermore, we compare the mean absolute and relative errors of different propagation models (including SI and SEI) for four different tweets in Table 1. The improved SI and SEI models show a significant error reduction compared to other models, demonstrating the improved prediction accuracy of the improved both models in capturing the dynamics of information propagation.

Table 1. Group mean absolute error and relative error of prediction results of different propagation models.

Tweet ID	Our Model	SI Model	SEI Model
Tweet 1	131/0.44	3426/14.96	562/1.47
Tweet 2	162/0.49	3334/13.07	422/0.92
Tweet 3	91/0.41	3456/18.27	350/0.84
Tweet 4	130/0.34	3355/10.40	1478/3.23

4.3. Transmissibility of Nodes in the Process of Information Diffusion

In this section, we juxtapose the method of calculating node propagation influence with existing techniques for evaluating network node influence. Our objective is to elucidate the constraints imposed on the information propagation capabilities of influential nodes within the social network. We employ different numbers and calculation methodologies to ascertain the influential nodes. Subsequently, we employ the model to disseminate information, enabling us to scrutinize the impact of nodes with substantial influence as determined by the proposed communication influence calculation method, in contrast to nodes with significant influence determined by alternative methodologies.

Within the domain of social network analysis, numerous techniques are employed to assess the influence of nodes. Two predominant approaches exist, one rooted in intrinsic node attributes, specifically the number of followers, and the other derived from the node's network position. For instance, the UserIndex value and PageRank value are indicative of a node's significance within the network structure to varying degrees. Notably, UserIndex encapsulates the representation of a user's influence in the social network by assigning distinct weight combinations to each attribute in accordance with the node's attributes within the social network. The attributes employed in this method encompass the node's follower count and tweet activity. The calculation method is presented in Equation (10):

$$\text{UserIndex} = w_{\text{flerCnt}} \times \text{flerCnt} + w_{\text{tweetCnt}} \times \text{tweetCnt} \tag{10}$$

This refined version maintains a scholarly tone and adheres to the style typically expected in the conclusion section of an academic article.

The two existing methodologies exhibit inherent limitations in their respective applications. These approaches primarily assess the impact of nodes at the attribute level and their significance in preserving network structure, but they may not fully capture their importance in the communication process. As a remedy, this study introduces an evaluation framework that assesses various influence calculation methods within the context of a communication model.

The comparative experimental protocol is outlined as follows: it entails the utilization of UserIndex values, PageRank scores, and the propagation influence values derived in this study to rank nodes within the network in descending order of importance. Subse-

quently, the top k nodes are targeted for propagation suppression. The ensuing propagation experiments commence from distinct starting nodes, allowing for an exploration of the repercussions of suppressing varying quantities of nodes based on distinct calculation methodologies. The experimental findings are visually presented in Figure 8.

In both Figures 8 and 9, propagation initiates from two nodes characterized by divergent penetration centrality levels, with the objective of disseminating information across the social network. The extent of information diffusion reaches its maximum when the top 0 to 1 nodes, as determined by different influence calculation methods, are suppressed within the network.

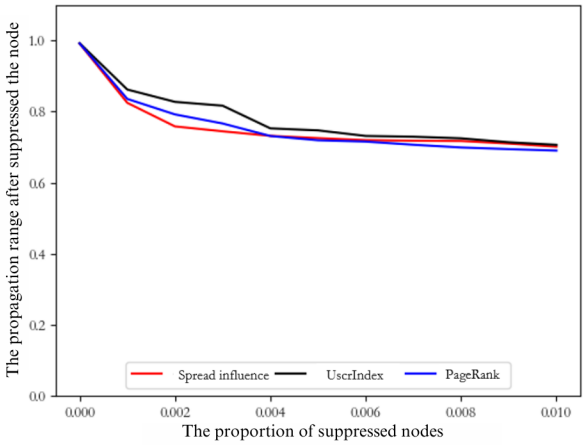


Figure 8. The value of ingress centrality of the propagation start node: 0.033.

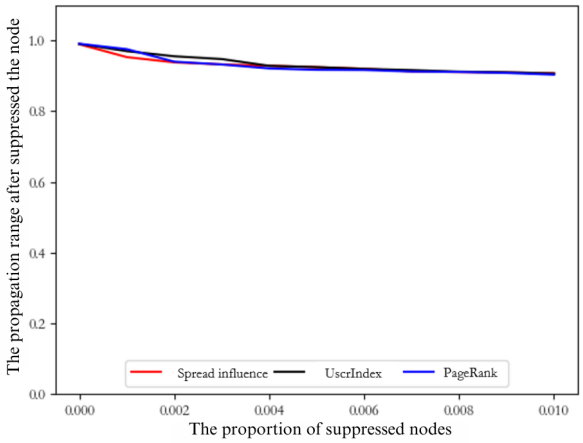


Figure 9. The value of ingress centrality of the propagation start node: 0.591.

Different methods for calculating influence are employed to curtail the spread of nodes, resulting in varying degrees of reduction in information propagation. In accordance with the influence propagation calculation method proposed in this study, the suppression of information propagation in the top k nodes with the highest influence yields superior control over the propagation range, outperforming alternative methodologies. As depicted in the accompanying figure, it is evident that nodes with the highest propagation influence, ranking within the top 0.2 percentile of the inhibition network, exhibit greater efficacy in curtailing the spread of information compared to nodes with the highest UserIndex or PageRank values within the same percentile of the inhibition network. In comparison to alternative node influence assessment techniques, the method introduced in this paper demonstrates a more effective capacity to influence the dynamics of information dissemination within the network.

In the context of the social networks illustrated in the figure, the choice of nodes with distinct centrality in information transmission has a discernible impact on the effectiveness of communication inhibition. As elucidated in Section 3.3.1, for scenarios where the initiating node directly exerts influence over the majority of network nodes, inhibiting nodes other than itself does not yield optimal results. This is due to the fact that even in the absence of external node promotion, the initiating node can disseminate information to a significant portion of the social network. Conversely, when the initiating node lacks direct influence over other network nodes, inhibiting these nodes substantially diminishes the range of information propagation. This empirical experiment validates the pivotal role played by nodes with substantial propagation influence in the following multiple rounds of propagation through the established model for influence assessment of network nodes. Consequently, the methodology for calculating node propagation influence, as delineated in this paper, is deemed to be both rational and well-founded.

5. Conclusions

In conclusion, this paper presents a comprehensive investigation of information dissemination within social networks. The initial sections introduce and analyze common information diffusion models, thoroughly examining node parameters within real-world social network data. By exploring the complex relationships between various parameter characteristics and the processes of information diffusion, we enhanced the Independent Cascade (IC) model. This enhancement involved redefining the calculation method for information dissemination probability, ultimately leading to the development of a practical and realistic information diffusion model.

Following the establishment of this diffusion model, we further examined the role of edge nodes as initiators and their transmissibility in the process of information dissemination. The analysis of the two generated diffusion models highlights that nodes with different centrality levels yield different dissemination outcomes. Notably, when nodes with high centrality assume the role of information spread initiators, they can accelerate the dissemination of information throughout the entire network.

In summary, the findings of this study provide valuable insights into the dynamics of information dissemination in social networks and propose an innovative model to capture the complexity of this process. The definition of diffusion influence makes a notable contribution to the field, with potential applications in optimizing information dissemination strategies within social networks. This research lays a foundation for future studies on information dissemination and influence within online communities.

Author Contributions: Methodology, J.D., Z.L. and H.H.; Software, R.L.; Formal analysis, H.H.; Data curation, X.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by The Fundamental Research Funds for the Centra Universities (No. ZYGX2022T001).

Data Availability Statement: The datasets generated during this study are subject to access restrictions as they constitute sensitive research materials from an ongoing social network analysis research. Qualified researchers may request access by contacting the corresponding author.

Conflicts of Interest: Author Jianwei Ding was employed by the company 30th Research Institute of China Electronics Technology Group Corporation. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Sheng, C.; Zhou, W.; Han, Q.L.; Ma, W.; Zhu, X.; Wen, S.; Xiang, Y. Network Traffic Fingerprinting for IIoT Device Identification: A Survey. *IEEE Trans. Ind. Inform.* **2025**, 1–14. [CrossRef]
2. Celi, F.; Allibhoy, A.; Pasqualetti, F.; Cortés, J. Linear-Threshold Dynamics for the Study of Epileptic Events. *IEEE Control Syst. Lett.* **2020**, *5*, 1405–1410. [CrossRef]
3. Zaman, T.; Fox, E.B.; Bradlow, E.T. A bayesian approach for predicting the popularity of tweets. *Ann. Appl. Stat.* **2014**, *8*, 1583–1611. [CrossRef]
4. Ruchjana, B.; Chaerani, D. Information diffusion model using continuous time Markov chain on social media. *J. Phys. Conf. Ser.* **2021**, *1722*, 012091.
5. Lobel, I.; Sadler, E. Information diffusion in networks through social learning. *Theor. Econ.* **2015**, *10*, 807–851. [CrossRef]
6. Shen, H.; Wang, D.; Song, C.; Barabási, A.L. Modeling and predicting popularity dynamics via reinforced poisson processes. In Proceedings of the AAAI Conference on Artificial Intelligence, Québec City, QC, Canada, 27–31 July 2014; Volume 28. [CrossRef]
7. Hunter, D.; Smyth, P.; Vu, D.Q.; Asuncion, A.U. Dynamic egocentric models for citation networks. In Proceedings of the 28th International Conference on Machine Learning (ICML-11), Bellevue, WA, USA, 28 June–2 July 2011; pp. 857–864.
8. Zhao, Q.; Erdogdu, M.A.; He, H.Y.; Rajaraman, A.; Leskovec, J. Seismic: A self-exciting point process model for predicting tweet popularity. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, 10–13 August 2015; pp. 1513–1522.
9. Rogers, E.M.; Singhal, A.; Quinlan, M.M. Diffusion of innovations. In *An Integrated Approach to Communication Theory and Research*; Routledge: London, UK, 2014; pp. 432–448.
10. Hernandez, A.; Ng, K.; Iamnitchi, A. Using deep learning for temporal forecasting of user activity on social media: Challenges and limitations. In Proceedings of the Companion Proceedings of the Web Conference 2020, Taipei, Taiwan, 20–24 April 2020; pp. 331–336.
11. Dey, A.K.; Gel, Y.R.; Poor, H.V. What network motifs tell us about resilience and reliability of complex networks. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 19368–19373. [CrossRef] [PubMed]
12. Namtirtha, A.; Dutta, A.; Dutta, B.; Sundararajan, A.; Simmhan, Y. Best influential spreaders identification using network global structural properties. *Sci. Rep.* **2021**, *11*, 2254. [CrossRef] [PubMed]
13. Tafti, A.; Zotti, R.; Jank, W. Real-time diffusion of information on Twitter and the financial markets. *PLoS ONE* **2016**, *11*, e0159226. [CrossRef] [PubMed]
14. Zhou, T.; Liu, J.G.; Bai, W.J.; Chen, G.; Wang, B.H. Behaviors of susceptible-infected epidemics on scale-free networks with identical infectivity. *Phys. Rev. E* **2006**, *74*, 056109. [CrossRef]
15. Newman, M.E. Spread of epidemic disease on networks. *Phys. Rev. E* **2002**, *66*, 016128. [CrossRef] [PubMed]
16. Nazemian, A.; Taghiyareh, F. A Mutual Independent Cascade Model for Customer Behavior Propagation. *Int. J. Inf. Commun. Technol. Res.* **2013**, *5*, 35–44.
17. Liu, L.; Qu, B.; Chen, B.; Hanjalic, A.; Wang, H. Modelling of information diffusion on social networks with applications to WeChat. *Phys. A Stat. Mech. Its Appl.* **2018**, *496*, 318–329. [CrossRef]
18. Kumar, S.; Saini, M.; Goel, M.; Aggarwal, N. Modeling information diffusion in online social networks using SEI epidemic model. *Procedia Comput. Sci.* **2020**, *171*, 672–678. [CrossRef]
19. Ver Steeg, G.; Galstyan, A. Information transfer in social media. In Proceedings of the 21st International Conference on World Wide Web, Lyon, France, 16–20 April 2012; pp. 509–518.

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.