Dynamic Analysis of Information Diffusion in Social Networks: Evidence from Facebook

Yiwei Qi, Hongzhou Tao

Graduate School of Arts and Sciences, Georgetown University, USA

August 3, 2024

Instructor: James Hickman

Abstract

This paper explores the factors which influence the way information is shared in social networks, putting emphasis on the role of various network structures. It is by accessing diverse network types that we want to look for the way they forward the spread and mode of information. Our investigation involves the use of Facebook, the juxtaposition of social network graphs, the computation of network structure metrics, and detection of supernodes. We also developed a model for the diffusion process of information besides comparing the information patterns of different ones such as news, rumors, and ad placements. Our study shows that the structure of the network is the most significant factor in the information spreading process. Data deputation through a high level of interconnection brings about the fast delivery of information, while the pronounced clustering can cause and effect for it to happen. We exposed that the network structure is an essential element in devising the strategy of the information dissemination process.

Keywords: Information Diffusion, Social Networks, Network Analysis, Information Propagation

1. INTRODUCTION

Nowadays, in today's digital age, the transference of knowledge to people through intersections like social media has caught the attention of many researchers. The fast growth of the digital world and social media is making it easier for different information to be spread around quickly and that has implications far and wide, a big audience can be reached in a short span of time. The knowledge of spreading an idea in the network is not only relevant to optimizing communication strategies but also to predicting the coverage and the influence of the different types of data. This understanding is very important in the development of the performance and accuracy of information exchange, thus, that the necessary information will get to the right person at the right time.

Information spreading in social networks is a very complex and diverse phenomenon. Different network structures like random networks, small-world networks, and scale-free networks are a significant factor in the way information diffusion occurs. Specific characteristics of the structures are at play determining the mode of information propagation and the speed of diffusion respectively for each of the structures. One common pattern is that random networks may display different diffusion patterns than small-world networks which are characterized by high clustering and small average path lengths. Furthermore, Scale-free networks which consist of few highly connected nodes (hubs) show a different set of transmission dynamics. Thus, the analysis of information dissemination in social networks extends one's ability to disclose the environmental state and the influencing factors, which yield a profound knowledge of the data flows over these intricate systems.

The reason that we take the study of the dissemination of information as an important aspect lies in the manifold applications that are related to it. The research in this field can be useful for building marketing strategies that will ensure the successful promotion of disinformation, offer solutions to disinformation problems, and allow for a quick response to public health.

In the case of natural disasters, for instance, information flows can help to make sure that the initial messages that citizens receive are comprehensive, clear, and arrive quickly, which will be very beneficial to the crisis response community. Also, with this knowledge, organizations, and groups will be able to come up with clear and effective communication strategies and thus have a higher influx of information and its influence.

Companies are frequently developing the ability to harness the power of information diffusion to develop more attractive and influential ways to quickly market products or services and to mold new rival products and services. These will be some of the potential benefits that will be available as a result of the better-targeted marketing campaigns, the increase in the customer conversion rate, and the higher customer interactivity.

Mastering information dissemination patterns also helps the government and the agencies provide accurate and timely information, thereby preventing the emergence of rumors and misinformation during such outbreaks as the coronavirus, and in the meantime, assuring the public safety. The proper use of communication strategies can be through the rapid reception of health guidelines and treatment plans by the population, which is a sure way of stifling the uncertainty generated by the diseases.

The knowledge of the processes involved in the dissemination of information in politics and the resulting activation of public opinion can become of great assistance in the forming of campaign strategies that can attract more voters. By scrutinizing the characteristics of political messages that are distributed on social media networks, the campaigns will get to tinker with their strategies to reach the highest number of people and, at the same time, be able to impact them with the message.

Another field that examines the spread of false news and misinformation is the science of information dissemination. In a time when fake news could spread through social media at an incredible speed, it is necessary to find out how to prevent and neutralize its dissemination. This study may be the basis for the creation of algorithms and strategies that help to recognize the spreading of misinformation and the consequent ensuring of the accuracy and reliability of the information.

The start goal of the survey is to analyze the dynamic characteristics of information diffusion in social networks and to determine how different network architectures can influence the speed and pattern of information diffusion. Through the use of the latest modeling tools, we have a projected idea of the information breakdown and thus coming up with elements that are prerequisites for the dissemination of the outcome. We will in a special manner carry out a simulation along with the data analysis with the aim of studying the features of different network environments, which enable knowledge sharing. Through the completion of these modules, we will be able to comprehend the roles played by network topologies and user interactions in the propagation or withholding of information.

This paper is organized in such a way that the various constituents of information diffusion in social networks are thoroughly examined. Chapter 2, Literature Review, is a synopsis of existing studies and a basis for the analysis. Chapter 3, Experimental Design, covers data collection, preprocessing, model construction, and experimental methods and parameters. Chapter 4, Experimental Results and Discussion, presents social network structure analysis, information diffusion modeling, user influence analysis, the study of different information diffusion patterns, and methods for optimizing information dissemination paths. Chapter 5, Conclusion, recaps the results, shows the dynamic mechanisms of information diffusion, explains key influencing factors, and hints at new research outlets for the future.

Then we list 10 question we want to explore:

- What are the degree distribution characteristics of the social network?
- What are the clustering coefficient and average path length of the social network?
- Are there super nodes (super spreaders) in the social network?
- How can the node and edge data in the social network be used to model the information diffusion process?
- How do different types of information (such as news, rumors, advertisements) spread differently in social networks?
- How can a user's influence in the social network be measured?
- What role do high-influence users play in the information diffusion process?
- How do the diffusion range and speed of different types of information vary in social networks?

- How can the information propagation path in social networks be optimized to minimize propagation time or maximize propagation range?
- What are the differences in information propagation efficiency before and after optimization?

2. LITERATURE REVIEW

In his research work "Structure and Function of Complex Networks" Newman (2003) gave the attention to the impact of the network topology in the process of information dissemination and mainly focused on the small-world effect and scale-free characteristics in promoting rapid dissemination as the main targets the paper touches upon. By looking at the nodes and edges forming the network structure, he indicated that high clustering and short path lengths in small-world networks are the two most significant factors responsible for the speedy transmission of information within the network's own little clusters, while the role of a few highly connected nodes in scale-free networks in the widespread, but uneven diffusion of information was also discussed. Here, beyond the matter of degree distribution, clustering coefficient and the average path length must also be mentioned, as these are three factors which also play a major role in the propagation of the information. Xu and Shen (2015) take into account the revelation that opportunistic links and node mobility made significant changes in the mobile social network communication they discovered in the simulation experiments by brainstorming on how they change the path and speed of the information dissemination that would then lead to the corresponding effect on communication. These investigations offer newfound directions for grasping the net's tenor and the tautness of spationetworks along.

The process of creating and simulating information diffusion models is the center of studying the dissemination of information. Morales et al. (2014) had researched Twitter data to know how it goes about influencers and information dissemination that leads to the partaking of the Twitter population in a social contagion (KINER & ROZENFELD, 2011). This study has reached quite a few remarkable results-few and far between as well. It's a well-conducted survey of the Twitter population. For the first time, this study was carried out on a population of a real social platform. They have pointed out that the users with many followers and who post information frequently are the big players in information dissemination and, thus, they can change the way in which others get to find out information which may involve some changes to marketing strategies. Liu et al. (2019), in a case of the hell raising that comes out of online emergencies, talked about the nonlinear dynamics that are evident when information is propelled through an emergency by introducing a nonlinear dynamic model that is used to investigate the mode of information propagation under emergencies and that can be used to identify the crucial drivers that orchestrate transmission in the system. These researches provide a foundation that could be used for the construction and simulation of the diffusion and dissemination of information.

The various forms of data can show significant differences in the ways that they spread out in social networks. Zhang and Zhang (2016), for example, largely concentrated on the subject of illnesses, network security as well as public opinion observation, and informed that diffusion models are not only useful in practicality but are also significant in theory. Information dissemination patterns are going to be in a clearer and predictive way be utilized if they are applied to almost all the existing fields and the new one, which in turn will enable response strategies to be more efficient. These studies lay a foundation for capturing the essential features of information's variation and coming up with methods to deliver it.

Analysis of user influence is another very important thing when the study of information dissemination is presented. Jiang and Liu (2014) conducted an evolutionary game theory-based model to demonstrate the dynamic communications strategies in social networks. They contrived a game theory model of sightseeing the developing procedure of different communication strategies in the network. The users have been revealed to be the primary contributors in information communication through their choices in strategy and their interactions with the people they communicate with. This is highly informative and relevant to the highly challenging and changing process of information dissemination in social networks. Chatman (1986) propounded that information dissemination is based on these three factors - network structure, information content and the cognitive state of the receiver. He preached that network structure actually controls the precise route and speed of information dissemination, also information content practically alters dissemination effectiveness, and the cognitive state of the receiver truly decides, respectively, the acceptance and readiness of dissemination, in this way, providing the theoretical support for the understanding of multiple factors in information dissemination. These studies present ways and indicators for monitoring user influence in social networks.

Overview of existing research indicates that the network structure as well as user behavior can be considered as highly important factors in the dissemination of information, but there are still many issues that should be researched. By building upon the previous research that we have done, we will try to explore the significance of the various types of network structures and the user behaviors and their relationship to the information dissemination in the interest of revealing the dynamic mechanisms of information dissemination and providing theoretical and practical advice for fine-tuning the dissemination strategies.

3. EXPERIMENTAL DESIGN

This study utilizes three primary datasets: Social Circles: Facebook, Graph Embedding with Self Clustering, and Facebook Large Page-Page Network from SNAP (https://snap.stanford.edu/index.html). These datasets offer comprehensive node and edge data, allowing for in-depth analysis of information diffusion characteristics and optimization strategies within social networks. Data statistics are presented in Table 1 below.

Table 1: Data Statistics

Social circles: Facebook		Graph Embedding with Self Clustering: Facebook			Facebook Large Page-Page Network	
		Category	Nodes	Edges		
Nodes in largest WCC	4039 (1.000)	Government	7,057	89,455	Directed	No.
Edges in largest WCC	88234 (1.000)	New Sites	27,917	206,259	Node features	Yes.
Nodes in largest SCC	4039 (1.000)	Athletes	13,866	86,858	Edge features	No.
Edges in largest SCC	88234 (1.000)	Public Figures	11,565	67,114	Node labels	Yes. Binary-labeled.

Average clustering coefficient	0.6055	TV Shows	3,892	17,262	Temporal	No.
Nodes	4039	Politician	5,908	41,729	Nodes	22,470
Edges	88234	Artist	50,515	819,306	Edges	171,002
Number of triangles	1612010	Company	14,113	52,310	Density	0.001
Fraction of closed triangles	0.2647				Transitvity	0.232
Diameter (longest shortest path)	8					
90-percentile effective diameter	4.7					

Experiment first verifies the over all characteristics of the network creating social network graph, computing the Network structure parameters related to the Degree Distribution, Clustering coefficient, Average Path Length.

page to find that page or to add new information on that page. The specific method includes the use of. Each of them is an edge file to build up the social network graph, we then compute the degree distribution, clustering coefficient, and average path length of the network. While recognizing super nodes, information concerning the degree distribution and centrality measures (the degree centrality, betweenness centrality, and closeness centrality measures are used). To simulate the information diffusion process, we use the SIR (susceptible-infected-recovered) model, which includes three states: consists of susceptible (S), infected (I) and recovered (R) individuals and infection rate (β), and recovery rate (y) as its parameters. By randomly selecting the initial infected node and setting appropriate parameters, we simulated the information propagation process in the social network, and simulated three types of information: news, rumours and advertisements, the ability of news ' and rumours compared to that of advertisements to spread their messages around in a short time span.

In the context of the user influence analysis, a volume of users' activity in the social network can be estimated by using the specific indexes, including the degree of vertex centrality, the betweenness of the vertex centrality, and the closeness of the vertex centrality. Degree centrality gives the number of connections of a node, betweenness centrality gives the importance of a node in the shortest path, and closeness centrality gives the average shortest path distance from a node to the other nodes. Thus, using the prescriptions of the methods, we were able to define the most central individuals in the network and study their contribution to the information spread. High degree centrality means a high number of edges connecting a particular user and they are primarily used in spreading out information; High betweenness centrality entails one is a link in the transmission path of information between two widely separated parts of the network; Users with high closeness centrality can easily access most other users within the network and are critical in the rapid transfer of information.

In terms of optimizing the information dissemination path, we used the Dijkstra algorithm and the Edmonds-Karp algorithm to optimize the dissemination path. The Dijkstra algorithm is used to calculate the shortest path, while the Edmonds-Karp algorithm is used to calculate the maximum flow. By comparing the effects of different optimization strategies, we evaluated the information dissemination efficiency before and after optimization, and used visualization methods to display the optimized dissemination path. The construction of these methods and models enables us to deeply explore the characteristics of information diffusion in social networks, compare the diffusion patterns of different types of information, and optimize the information dissemination path, ultimately providing theoretical support and practical guidance for the formulation of effective dissemination strategies.

4. EXPERIMENT RESULTS

4.1 Network Structural Characteristics

First, using the various network parameters, the structural properties of the social network were determined using the degree distribution, clustering coefficient, and average path length. These measures give a basic knowledge about the other shape and connectivity of the whole network and the existence of super nodes or super spreaders.

In our analyzed network, the degree distribution revealed that majority of nodes possess a low degree and few of the nodes possess high degree. This makes one suspect the existence of super nodes which establish many connections hence are very vital in determining the connectivity of the network. The degree distribution gives a scale-free network where few of the nodes have most of the connections. The degree distribution of the specific type identified in this study is depicted below (Figure 1).

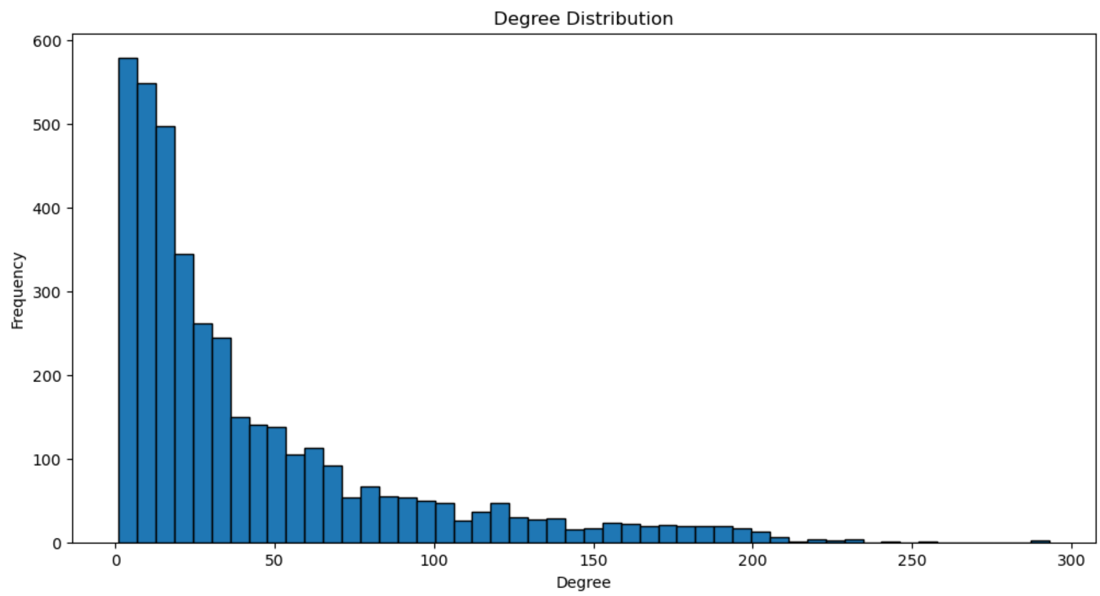


Figure 1

The over-all clustering coefficient of the network is 0.5437, which probably means that in the network nodes prefer to stay together and are closely connected. This has a high clustering coefficient which means that the local clustering is evident, friends tend to associate with friends, thus giving many numbers of close-knit groups of friends. This characteristic is quite expectable for social networks as people tend to form groups.

Thus, the analysis of the specified network showed that the network is not complete; in fact, it contains 13 connected components. This disconnected state does not allow calculating a unified value of the average path length of the entire network. However, for each connected component one can calculate the average path length separately, which will give more information about the structure of the network. If there are two or more connected components, then this means that although some part of the network is densely connected, there exist other small dense sub-graphs. Super nodes were identified using degree centrality. The top 10 nodes with the highest degree centrality are as follows:

Table 2: The Top 10 Nodes with the Highest Degree Centrality

Node	Centrality
2543	0.0740
2347	0.0733
1888	0.0639
1800	0.0616
1663	0.0591
1352	0.0589
2266	0.0589
483	0.0584
1730	0.0568

1985	0.0563

These super nodes are also very influential in the network, as they connect many other nodes in the network and ensure timely sharing of information in the network. The splendor illustrates the critical role of several kinds of nodes in providing the overall organization of the network and the stable conditions to facilitate message exchange.

The totality of the examined structural characteristics underlines that the social network reveals scale-free characteristics with the high local clustering coefficient and multiple connected parts. The identification of super nodes also shows that there are some nodes to keep connected the network and guarantee information diffusion. These discoveries yield practical information and a topological map and the routes of access to them by the nodes in the network.

4.2 Information Diffusion Model Analysis

The proposed diffusion model with a SIR structure was used to simulate the diffusion process of the various types of information circulating in a social network. The largest component was chosen for the experiment. Through selection of the first infected nodes and parameters that define the spread of infection (β / γ and γ) the process of information dissemination in the network was properly simulated. Amount and frequency of varied sorts of information in different types covering discrete time intervals were measured regarding their diffusion rates. This analysis compared the transmission details of various forms of information, news, rumors, and advertisements in the network by assessing identical features in the range and speed of the operating diffusion wave. The findings were also presented in graphic form in order to explain how different types of information diffuse.

Simulating the spread of information in a network can be done easily with help of the SIR model.

Table 3: SIR Model Description

Symbol	Full name	Description
S	Susceptible	End-users can be defined as persons who have not come across the information before。
I	Infected	Persons, who have met and are disseminating the information themselves.
R	Recovered	Persons who have interacted with the information and Hub, but have ceased to popularize it.
β	Beta	The transmission rate, which is the daily probability of a person who gets infected transmitting the information to a person who has not been infected.
γ	Gamma	Recovery rate, by which we mean the probability of an infected person moving from the infected state to the recovered state.

According to the range of the initial infecting nodes and the choice of β and γ values, the propagation of information in the social network was modeled.

The simulations were conducted for different types of information: gossips, rumors and special bulletins News. The results are detailed in table below:

Table 4: Diffusion Characteristics of News, Rumors, and Advertisements

Type of Spread	Infection Rate	Recovery Rate	Diffusion Characteristics

News Spread	0.03	0.01	The spread results show that news initially spreads rapidly, reaching a peak in the 20th iteration, then gradually decreases as a large number of users recover.
Rumor Spread	0.05	0.01	The spread results show that rumors spread faster, with a higher peak number of infected individuals, reaching the peak in a shorter time, then gradually decreasing as the number of recovered individuals increases.
Advertisement Spread	0.02	0.01	The spread results show that advertisements spread more slowly, with a lower peak number of infected individuals, taking a longer time to spread, but eventually, a larger number of individuals recover.

Advertisements increase the rate of infection at a slower pace than news and rumor with a comparatively low number of infected users at the peak. The spread duration is longer but after a certain period, a large population of users are cured.

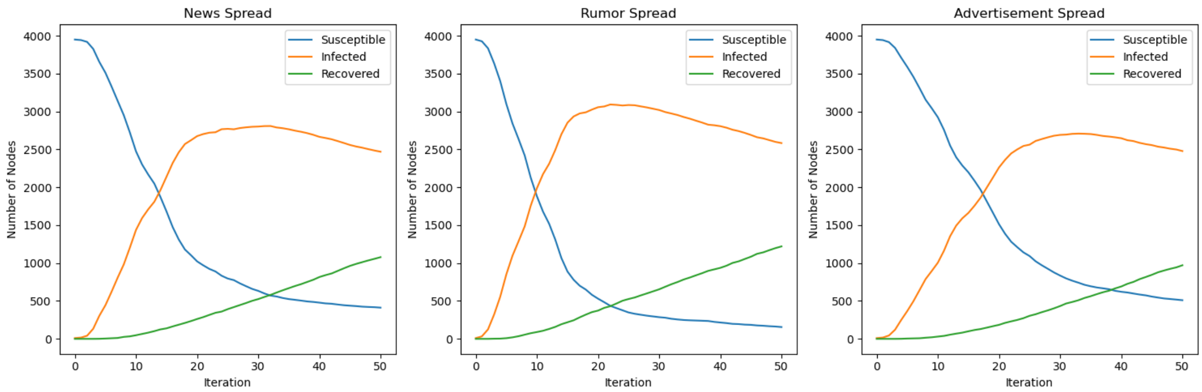


Figure 2

The patterns of diffusion regarding news, rumours, and advertisements are different with respect to the velocity of the spread and the infection rate peak. News infects its audience at a slow rate but it’s possible to have very high rates when it hits the peak, it gains public attention quickly and forms a permanent part of society. Gossip, on the other side circulated in the shortest time possible and covered a bigger crowd, resulting to the highest number of infections within the shortest time possible and the overall infected persons at the time of the epidemic. Nevertheless, they have a short life cycle as they are condemned soon. Advertisements were being slow to spread, especially in the countries that had the least number of peak rates of infection, but in return, they occupied a smaller part of the total population while influencing them for a longer period of time.

From the SIR model simulations, it is observed that the diffusion of various kinds of information differs from each other in the context of the social network. News has a fast diffusion rate and a high level of forgetting, rumors, on the other hand, diffuse faster and penetrates further into the community; however, advertisements take the longest time to diffuse, but the message circulates for a long time. It is imperative to comprehend these phenomena in order to construct methods for organizing and utilising the circulate of information in social networks.

4.3 Measuring User Influence in a Social Network

By analyzing centrality metrics within the social network, we can measure user influence and understand the role of highly influential users in the information diffusion process. But how to Measure User Influence in a Social Network? There are three centrality to depend.

While degree centrality is a measure of the node's number of connections, the value obtained does indicate the node's predominant status within the local network. Finally, we find the degree centrality of each user and determined the highest degree centrality users equal to 10.

Degree centrality shows the number of primary connections a user has and these are highly connected users who are significant in fast and efficient dissemination of information.

Table 5: Top 10 Users by Degree Centrality

User	Degree Centrality	Betweenness Centrality	Closeness Centrality
2543	0.074027	0.044374	0.239661
2347	0.073269	0.014075	0.220938
1888	0.063921	0.001429	0.247286
1800	0.061647	0.001456	0.259547
1663	0.059121	0.004362	0.26254
1352	0.058868	0.004888	0.263037
2266	0.058868	0.007084	0.211806
483	0.058363	0.039944	0.246784
1730	0.056847	0.001365	0.24936
1985	0.056342	0.001051	0.199012

From all the measures, Betweenness centrality estimates the centrality importance of the node based on the shortest paths of the network signifying the intermediary nature. We approximated the betweenness centrality for every user and then selected the first ten users in betweenness centrality order. High betweenness centralities of users make them act as shuttles in the network and help to establish multiple relations between different parts of the network.

Table 6: Top 10 Users by Betweenness Centrality

User	Degree Centrality	Betweenness Centrality	Closeness Centrality
1085	0.01617	0.24917	0.256843
1718	0.038656	0.168706	0.27036
698	0.006569	0.090464	0.177853
1577	0.042446	0.090379	0.255747
862	0.005811	0.088428	0.212117
1405	0.012633	0.085796	0.250951
1165	0.006316	0.080415	0.269424
136	0.033098	0.072715	0.230307
107	0.006316	0.069074	0.25771
171	0.005558	0.066085	0.244737

Closeness centrality might be defined as the average distance of a node to all other nodes and demonstrates how easily a node can be reached from all the other nodes. Next, we calculated the closeness centrality for each user and short listed 10 users with maximum closeness centrality. Closeness centrality represents the ability of certain users to access other people in the network; consequently, users with high values of this indicator can contribute greatly to the spread of the information.

Table 7: Top 10 Users by Closeness Centrality

User	Degree Centrality	Betweenness Centrality	Closeness Centrality
1534	0.008338	0.061843	0.270943
1835	0.04573	0.008152	0.270566
1718	0.038656	0.168706	0.27036
1165	0.006316	0.080415	0.269424
1173	0.028802	0.011828	0.269108
1376	0.04573	0.006351	0.269108
1509	0.032845	0.005378	0.267831
1312	0.01996	0.003516	0.26591
1334	0.025013	0.023087	0.264358
1590	0.021981	0.003741	0.264197

Based on the analysis of these centrality metrics, we identified highly influential users in the network and understood their role in the information diffusion process: Based on the analysis of these centrality metrics, we identified highly influential users in the network and understood their role in the information diffusion process:

High connections, these are the users with many connections to other users and mainly responsible for spreading information. This is due to their ability in disseminating information to many users within a short span thus achieving the wide diffusion of information.

High Betweeness Centralization Users, these are the users that exist within the information transfer paths along bridges, they are the intermediaries. One can link the different areas of the network thus aiding in the transaction of information from one area of the network to another and thus goods and services don’t remain stagnated in certain sections of the network.

Evaluating the centrality parameters suggest that high closeness centralities, these are user who in quires limited 'hops’ to get in touch with other users in the network and thus are important for distributing information. They can also quickly reduce the time it takes for information to pass through the network so in essence, information can easily reach every node on the network.

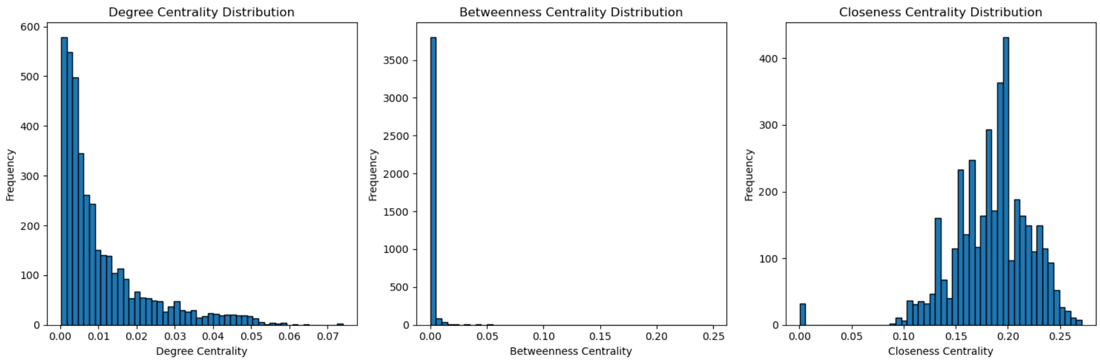


Figure 3

Applying the degree centrality, betweenness centrality and closeness centrality one will be in a position to quantify the influence of a user in a given social network. The opinion leaders are very important in the diffusion process as they act as the connectors disseminating information to the other users in the network with great speed and range.

4.4 Diffusion pattern analysis of different types of users

The diffusion process of different types of information was simulated in the network. By setting initial infected nodes and propagation parameters (such as infection rate β and recovery rate γ), the dissemination process of information in the network was simulated. The diffusion range and speed of different types of information over certain time steps were recorded. The spread patterns of government, news sites, athletes, public figures, TV shows, politicians, artists, and companies information in the network were compared, identifying their similarities and differences in propagation range and speed. The experimental results were visually displayed to show the diffusion paths of different types of information.

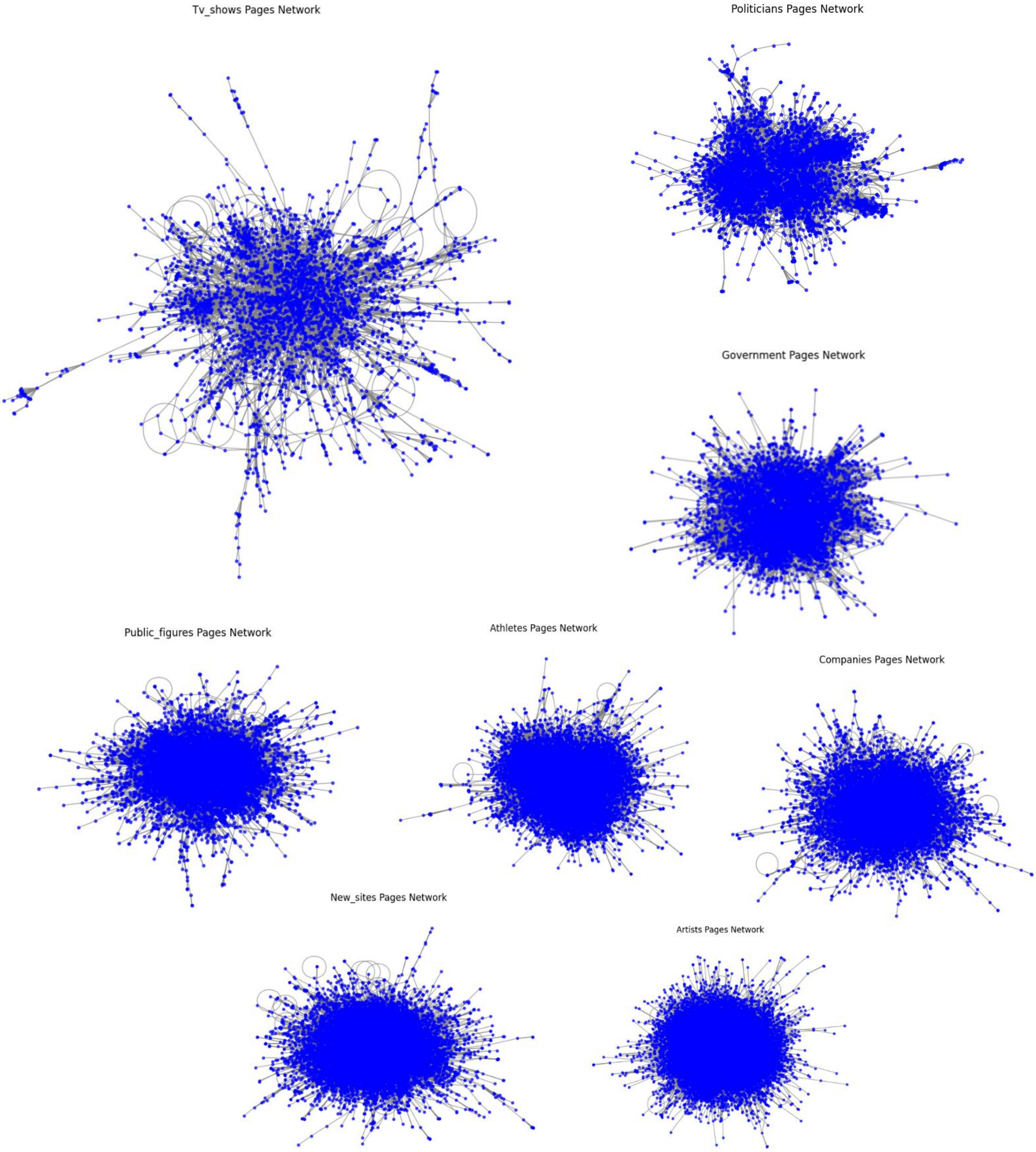
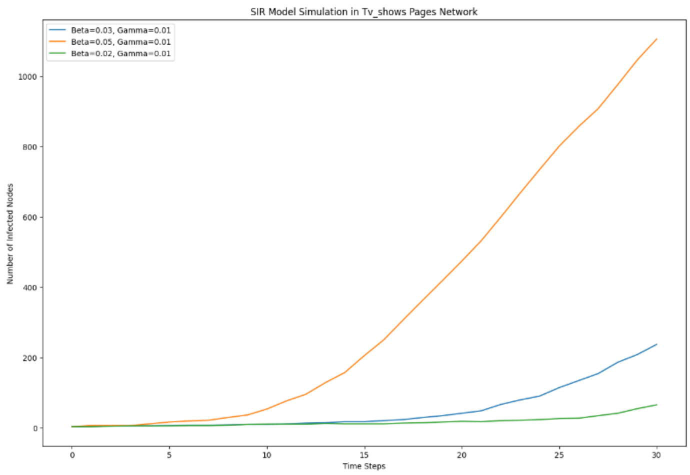
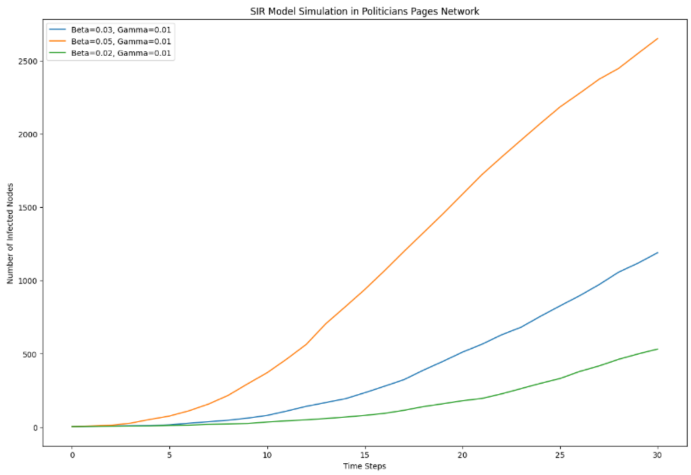
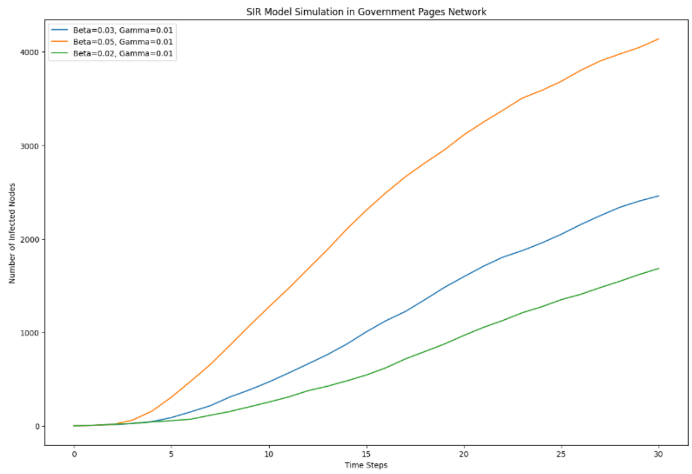


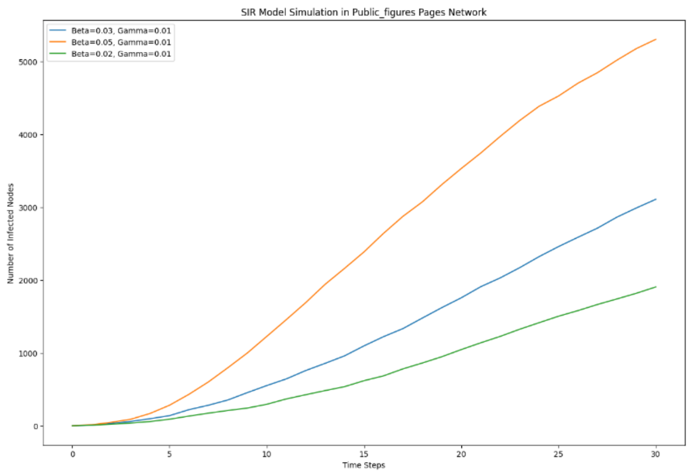
Figure 4

Government information spreads at a moderate speed with a wide diffusion range, covering most network nodes within a short period. News sites information spreads quickly with a wide diffusion range, rapidly reaching various parts of the network. Athletes information spreads at a moderate speed with a concentrated diffusion range, primarily within specific communities. Public figures information spreads at a moderate speed with a wide diffusion range, showing high dissemination effects. TV shows information spreads at a moderate speed with a smaller diffusion range but effectively among specific nodes. Politicians information spreads quickly with a wide diffusion range, rapidly covering a large number of network nodes. Artists information spreads at a moderate speed with a wide diffusion range, maintaining dissemination over a longer period. Companies information spreads slowly with a relatively smaller diffusion range but has a longer duration of propagation.









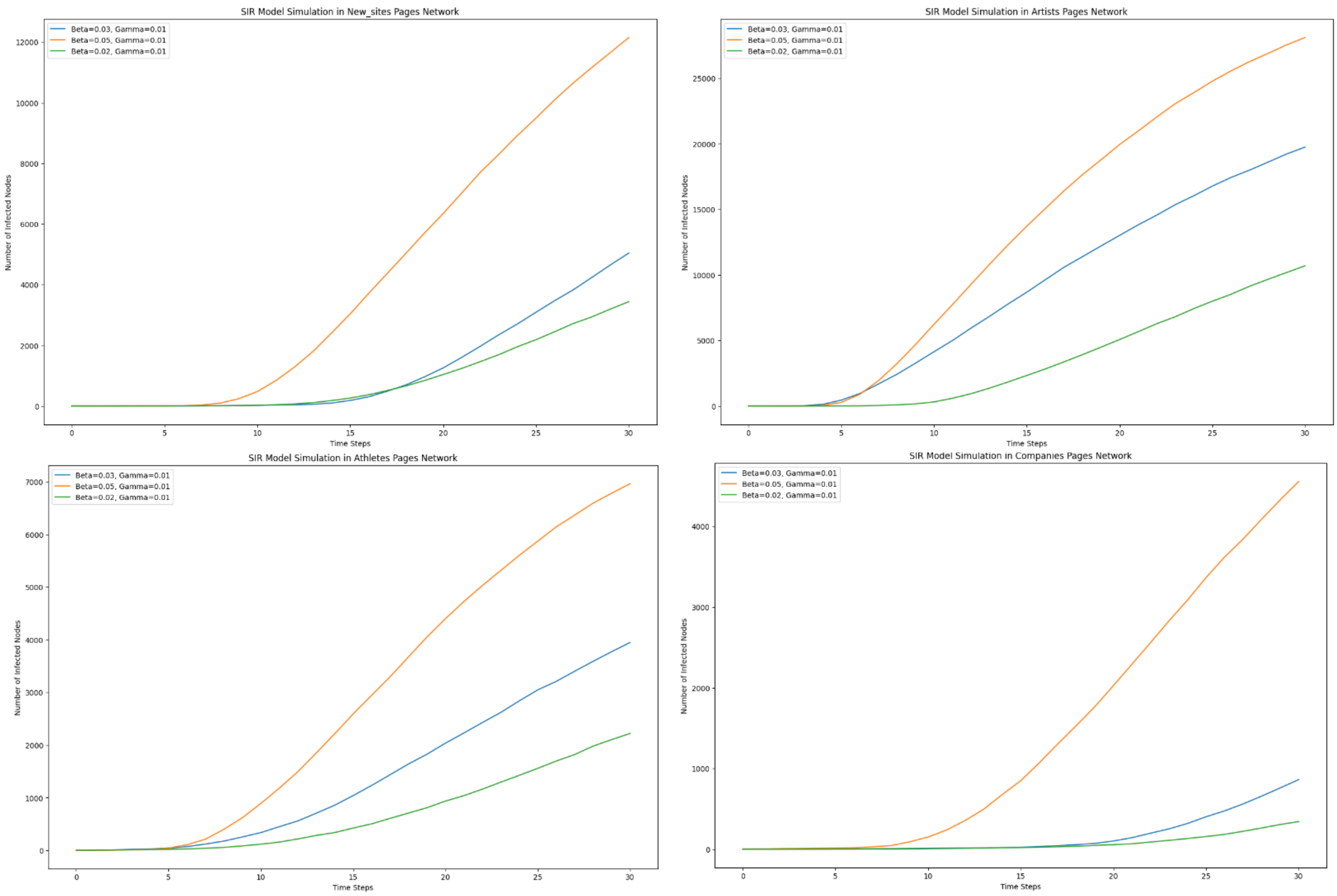


Figure 5

The experimental results indicate significant differences in the spread patterns of different types of information in social networks. Government information spreads at a moderate speed initially, covering a wide range and then stabilizing gradually. News sites information, due to its timeliness and broad appeal, spreads quickly and covers the entire network rapidly. Athletes information primarily spreads within specific communities, with a concentrated diffusion range. Public figures information shows high dissemination effects and spreads widely. TV shows information spreads effectively among specific nodes but has a smaller overall diffusion range. Politicians information spreads quickly, covering a large number of network nodes rapidly. Artists information spreads at a moderate speed and maintains dissemination over a longer period. Companies information spreads slowly, with a smaller diffusion range but a longer duration of propagation.

In each type of user network, we simulated the information propagation process. The initial infected nodes and propagation parameters, such as infection rate (β) and recovery rate (γ), were set to simulate the diffusion of information in the network. The diffusion range and speed of different types of information in a specific time step were recorded, and the diffusion paths and characteristics of different types of information were displayed through visualization methods.

For the government user network, the simulation results show that the information diffusion of news, rumors, and advertisements has its own characteristics. The highest infection rate (β =0.05) leads to rapid information spread, covering most of the network nodes in a short time. Lower infection rates (β =0.03 and β =0.02) lead to slower propagation and a smaller number of infected nodes. News information spreads faster in the government network and quickly reaches a high number of infected nodes, while rumors spread the fastest and have the highest number of infected nodes. Advertising spreads relatively slowly, but can also cover a wider network for a longer time. In the news website user network, news information spreads the fastest, showing a wide range of spread and a high number of infected nodes. The information diffusion speed of rumors follows closely, quickly covering most of the nodes in the network. Advertising information spreads relatively slowly, but the number of infected nodes increases gradually over time.

For the athlete user network, news and rumors spread at a moderate speed. The highest infection rate causes information to spread rapidly within a specific community, while lower infection rates show a more concentrated spread pattern. Advertising information spreads slowly in this network, but eventually affects a considerable number of nodes. The public figure user network has a high information diffusion efficiency. News and rumors spread rapidly in this network, especially under the highest infection rate condition, the number of infected nodes increases rapidly. Advertising information spreads slowly in this network, but its spread range gradually expands over time.

In the TV program user network, information spreads slower than other types. Although information spreads relatively quickly under the highest infection rate, in general, the spread of news, rumors, and advertising is concentrated in specific nodes, showing the specificity and concentration of their audiences. For the politician user network, news and rumors spread quickly, especially under the highest infection rate condition, quickly covering most of the nodes of the network. Advertising information spreads slowly, but eventually affects a considerable number of nodes.

In the artist user network, the information diffusion of news and rumors shows high propagation efficiency. The information diffusion range under the highest infection rate is wide, and the number of infected nodes increases rapidly. Advertising information spreads slowly in this network, but eventually covers a large network range. Information propagation speed in the company user network is slow. Even under the highest infection rate conditions, the speed of information diffusion is slow, but over time, the information spread range of news, rumors, and advertisements gradually expands.

4.5 Optimization in Social Networks

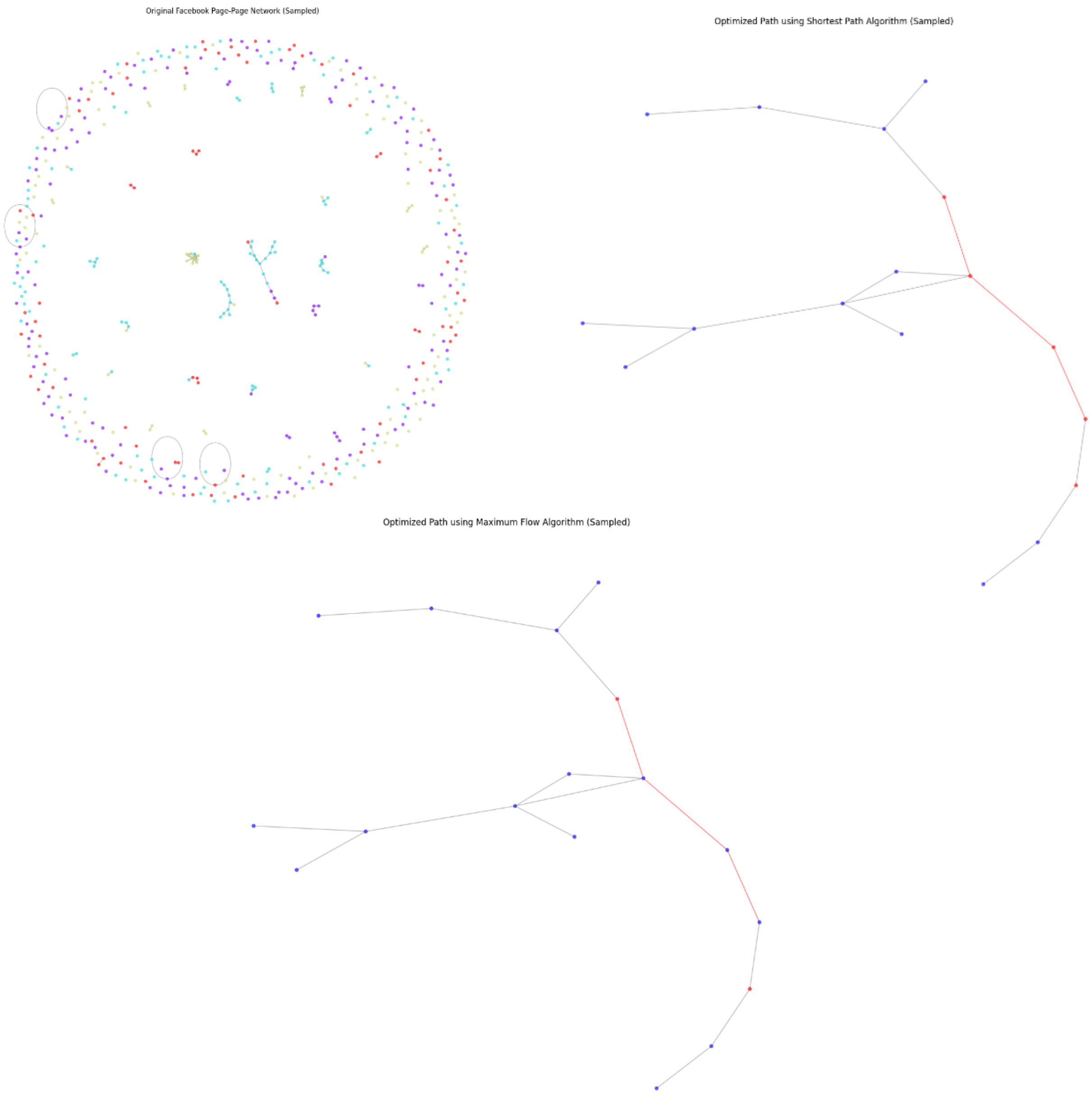


Figure 6

Through the shortest path algorithm, we find the optimal path for information dissemination. We can see the result in Table 7. The results show that the optimized shortest path length is 5. This means that information propagated in the network can reach the target node in the shortest number of steps through this optimal path, which significantly reduces the time for information propagation. This optimization result shows significant improvement compared to the original network diameter 9. The shortest path algorithm improves the efficiency of information dissemination by reducing the propagation time. This result has important application value for scenarios where key information is expected to be disseminated in the shortest possible time, such as breaking news dissemination or emergency notification issuance.

On the other hand, the application results of the maximum flow algorithm show that the maximum flow value from the source node to the destination node is 1. This result shows that, given the network structure and constraints, the maximum flow of information or resources from source to destination is 1. Although the maximum flow value is low, by optimizing the path, the effective delivery of information in the network can be ensured. The maximum flow algorithm improves the stability and continuity of information transmission by optimizing the path of information flow in the network. This is of great significance for scenarios where stable transmission of information or resources in the network needs to be ensured, such as data synchronization or resource allocation.

Table 8: Simulation Results

Shortest Path Length	5
Maximum Flow Value	1
Original Network Diameter (Sampled)	9
Shortest Path Length after Optimization	5

5. CONCLUSION

In conclusion, its principal value lies in deepening our knowledge about dynamic processes by which information spreads through social networks. It gives us a good handle on what structures within networks and who influences those structures cause information diffusion to take different paths. Consequently, the present results are also relevant with respect to the development of effective patterns for the transmission of information in fields of marketing, public health and crisis communication.

Future work should examine how some more factors (e.g., user behavior, content characteristics, and temporal dynamics) impact information diffusion. Meanwhile, integrating real-time data and leveraging advanced modeling technologies may further improve the performance of the current methods. We will continue our efforts in this line to develop more elaborate mechanisms for managing and maximizing information dissemination in social networks with increasing complexity.

6. REFRENCE

[1] Chatman, E. A. (1986). Diffusion theory: A review and test of a conceptual model in information diffusion. Journal of the American Society for Information Science, 37(6), 377-386.

[2] Jiang, Y., & Liu, M. (2014). Evolutionary dynamics of information diffusion over social networks. PLoS ONE, 9(10), e110486. https://doi.org/10.1371/journal.pone.0110486

[3] Li, X., & Zhang, Y. (2017). A survey on information diffusion in online social networks: Models and methods. Information Systems, 71, 111-128. https://doi.org/10.1016/j.is.2017.08.001

[4] Liu, Y., Wang, Z., & Zhang, Q. (2019). Information diffusion nonlinear dynamics modeling and evolution analysis in online social network based on emergency events. Physica A: Statistical Mechanics and its Applications, 523, 1280-1291. https://doi.org/10.1016/j.physa.2019.04.150

[5] Morales, A. J., & Benito, R. M. (2014). Efficiency of human activity on information spreading on Twitter. Scientific Reports, 4, 4473. https://doi.org/10.1038/srep04473

[6] Newman, M. E. J. (2003). The structure and function of complex networks. SIAM Review, 45(2), 167-256. https://doi.org/10.1137/S003614450342480

[7] Xu, K., & Shen, H. (2015). Epidemic information dissemination in mobile social networks with opportunistic links. IEEE Transactions on Emerging Topics in Computing, 3(3), 399-409. https://doi.org/10.1109/TETC.2015.2428107

[8] Zhang, Z., & Zhang, X. (2016). Dynamics of information diffusion and its applications on complex networks. Applied Mathematics and Computation, 291, 44-53. https://doi.org/10.1016/j.amc.2016.05.024

[9] Song, X., Lin, C. Y., Tseng, B. L., & Sun, M. T. (2005, August). Modeling and predicting personal

information dissemination behavior. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (pp. 479-488).