Analysing the spreading of a meme on social media

Lasya Mundrathi, Sneha Gadade, Disha Danej, Pablo Fatas

***Abstract*—With the increasing prevalence of social media, there are now greater opportunities to analyze human social behavior by exploring the networks and data streams that it generates. Social Network analysis involves examining social structures and patterns using network and graph theory. Networks are described in terms of nodes, edges, or links to connect them. This paper provides an overview of networks, their visualization, and analysis. It presents a straightforward example of conducting network analysis using data from the "Reddit Hyperlink Network", which focuses on the Susceptible-Infected (SI) model, where nodes are either infected or susceptible. The project's outcome provides insights into meme spreading dynamics on social media networks and how network design can influence these patterns.**

I. Introduction

In recent years, social media platforms have seen a huge increase in the volume of information exchanged there. Although there is a potential that spam, rumors, slander, and other types of incorrect material will be disseminated through these uncontrolled platforms. Memes are infectious cultural knowledge patterns that travel from person to person and have a direct impact on a social group's primary behaviors and attitudes. Network analysis has the capacity to estimate complex association patterns, and the network structure can be looked at to determine the network's key properties.

Nodes signify the individual users within a social network, and analyzing them includes understanding their behavior, interest, and interactions within the network. Whereas Graph structure represents social networks on social media platforms. They identify the users and connection between users and examine the pattern of interaction between them.
Overall, evaluating social media data using communities' graphs, and nodes can offer appreciated visions into the behavior and relations of users on social media platforms [13]. It can help us identify patterns, trends, and influential users.

A. Research problem

Understanding the range of memes on social media is crucial for grasping their effects on civic thoughts and behavior [3]. This includes using epidemic spreading models to discover how memes spread across the network or whether they are concentrated in specific clusters or more broadly.

B. Motivations

Examine the linkage between the nodes of the Reddit HyperLink network by recognizing features that influence the spread of memes which contribute to the expansion of the models and tools for analyzing social networks and predicting the spread of information.

C. Challenges

The network's enormity poses challenges in visualizing and analyzing it, while its dynamic and complex nature makes detecting patterns and communities challenging. Biases in the data from Reddit user's demographic and behavior may also impact the generalizability of findings.

II. Dataset and network presentation

A. Dataset

The Reddit online community platform served as the source of the dataset for this study. Users can join communities on the website Reddit called subreddits where they can post things including photographs, videos, and links to articles. The Reddit dataset utilized in this study was compiled from 40 months of data of publicly available Reddit posts and comments, that is between January 2014 and April 2017.

The dataset consists of Post and comment data from Reddit communities, commonly referred to as subreddits, make up the dataset.Cross-links across subreddits make up the dataset used in this study. A cross-link is a hyperlink that points to a post in another subreddit in a post from one subreddit. The dataset is a directed network in which the target node refers to the subreddit that the hyperlink points to and the source node to the subreddit where it was posted.The dataset utilized in this study contains 137821 cross-links made between 35766 subreddits after overlapping cross-links were removed. The data is organized as a directed network, with each node standing for a different subreddit and each edge representing a link (hyperlink) from one subreddit to another. Subreddit embeddings, which are representations of the subreddits in a lesser-dimensional space, are also included in the dataset.

B. Network

A directed network is the Reddit Hyperlink (RH) Network. Subreddits, or interest-based communities on the Reddit platform, are represented by the nodes in the network. The network's edges stand in for the hyperlinks that link content from one subreddit (the source) to content from another subreddit (the target). The edges' orientation matches the hyperlink's direction, which is from the source subreddit to the target subreddit.The direction of these hyperlinks determines how the network's nodes and edges are arranged.

The way the nodes and edges are set up in the RH network allows for hyperlink connections between each subreddit node and one or more additional subreddit nodes. A directed system like the RH network allows some nodes to have both in-degree and out-degree while only allowing others to have one of the two. Nodes with zero in-degree and out-degree total 15170 and 7913, respectively. 7,913 subreddits and 15,170 subreddits have no links from any other subreddit in the network. There are also 15,170 subreddits that have no links from any other subreddit in the network. Regarding hyperlink exchanges, these subreddits can be seen as being apart from the rest of the network. The number of nodes and edges within the network can be used to gauge its size. There are 35766 nodes in the network, which is exactly the same as the number of subreddits that are included in the dataset. The total number of links between subreddits in the dataset is represented by the network's 137821 edges.

The identification of node 59, which ranks highly in betweenness centrality, is the main discovery. This node is crucial in tying together various components of the network and facilitating information transfer between them. Using this knowledge, specific tactics for disseminating memes or other types of information throughout the network can be developed. There are many nodes that have a clustering coefficient of either 1 or 0. In the network, nodes that are highly coupled and have a clustering coefficient of 1 often form close-knit clusters that help information propagate more quickly. The propagation of information can be slowed down by nodes having a clustering coefficient of 0, which operate as barriers to information flow. The important implications for understanding the underlying mechanisms driving information diffusion in the RH network. This is a large and complex network with a relatively high density.

III. Methodology

A. Community detection algorithm on a complex network

1) Community detection algorithm with 5 different "Resolution" values: To extract communities from the network, we will use Gephi's community detection technique. Using louvian modularity optimisation, the technique divides the network into communities. The modularity of community structures is a quality metric, with higher values indicating a better structure. The approach seeks to maximise modularity by repeatedly relocating nodes across communities until no more improvement is possible [14].

We will run the community detection algorithm with five different values of the "Resolution" parameter to show how each value affects the resultant community structure. The resolution option adjusts the granularity of the division to manage the size of the communities.Communities turn into smaller patterns graphs and more densely linked when a higher resolution value is used, whereas bigger and more loosely connected communities are created when a lower resolution value is used.

By estimating the number of communities, the average size of each community, and the standard deviation of community size, we will examine the community structure produced by each resolution constraint value. This analysis will aid us associate the quality of the community structure across different resolution settings and choose the most optimal community.

2) Comparison of detected communities using basic network metrics: Now, we'll look at one of the five partitions and see how similar or distinct the detected communities are in that partition. To aid our discussion, we will use basic network metrics such as centrality measurements, clustering, and assortativity.

Degree centrality, betweenness centrality, and eigenvector centrality are centrality measurements that provide information about the relevance or influence of nodes in a network. Clustering measurements, such as the clustering coefficient or transitivity, characterize how closely nodes cluster together.

We may learn more about the similarities and differences between the discovered groups by analyzing these measures inside them. As an illustration, if two communities have comparable clustering coefficients but differing assortativity values, this indicates that they have a similar level of internal cohesion but differ in the way nodes with similar features are connected. Similarly, if two communities have different degree centrality distributions, it is likely that they have varying node importance or influence levels. By comparing communities we can gain a better understanding of the structure and traits of the discovered communities within a particular partition.

B. SI models and centrality measures

1) SI model: To measure the impact of each of these centrality measures we will be using an SI model. SI models have one or more starting infected nodes. In our case this is the subreddit where the meme was first created or at least posted to the platform. The rest of the nodes are classified as susceptible meaning that they can be infected. When looking at an SI model the choice of starting node is important. Choosing a poorly connected node or even a disconnected node from the biggest connect component will lead to the information to spread slowly or even not at all throughout the network. The next parameter we need to chose is the infection rate which determines the probability that a susceptible node next to an infected node gets infected on any given time step. As we are simply comparing spreads between simulations and not interested in modelling or predicting any particular one; we can chose an arbitrary alpha bigger than 0 as long as it is the same for each simulation. In this section we will look at different centrality measures and observe how picking our starting node based of those measures affects the spread of memes across subreddits. We will be looking at 2 centrality measures; eigenvector centrality, betweeness centrality. We will discuss how they differ from standard degree centrality and there relevancy in an SI mode [12].

2) Eigenvector Centrality: Eigenvector centrality evaluates a node's significance based off of its neighbours centrality more than its own. For example, a person who has 5 friends with no other friends would be considered less connected than a person who has 1 friend how has 1000 friends. In the context of an SI model this is significant because it is much more important to the spread of the information that that 1 friend with 1000 friends is "infected" than those other 5 friends with few friends.

3) Betweeness centrality: Betweenness centrality [1] is a way of detecting the amount of influence a node has over the flow of information in a graph. It is often used to find nodes that serve as a bridge from one part of a graph to another. The way it is computed is by looking at all the shortest paths between each pair of nodes and counting how many times a node is part of those path. This highlights less how connected a certain node is but emphasizes how crucial a node is to the connectivity between large components in the graph. In the context of an SI model, this is an important metric if we want the "infection" to spread across the entire network and not get stuck in one single large community because it has not yet infected the connecting nodes to other communities.

4) Implementation: To implement this we use the built in methods in the networkx package to get our most and least central nodes based off of our centrality measures. We then run our SI model for both measure's most and least central nodes (10 times each). Using these simulations we will plot a graph where data point is a node's time it was infected and distance from the starting node. This will give us a good idea of how the infection is spreading relative to the starting node.

5) Analyzing the Relationship between Infection Time and Network Distance in a Simulated Epidemic Spread: We would first replicate the SI epidemic spreading model for both the most and least central nodes in order to get a scatter plot of Ti vs Di for each node in the network. Di is the distance from the initially infected node, and Ti is the amount of time it takes for each node to get infected. To represent the centrality measure (betweenness or eigenvector), the plot would employ colour or form. We may be able to comprehend the role of centrality in the transmission of infection and develop focused therapies by analysing this plot.

C. Analyzing Spreading within and across the 5 Largest Communities

1) Analyzing Number of infected nodes over time in each communities: Community identification and network simulation methods are utilised for the specific job of running a SI model where the first infected node is a randomly picked node from the community and visualising Nc vs t. The Louvain approach for community detection is used to identify the top five biggest communities in the network, and a random node is chosen as the initial infected node for the SI model for each community. To simulate the propagation of the memes inside the network, the EoN (Epidemics on Networks) package is used, and the number of infected nodes for each community is recorded at each time step.

We may investigate how the size and structure of a community impact the rate and extent of meme transmission by randomly picking an initial infected node within it. Furthermore, by analysing the number of affected nodes in each community over time, we may identify communities that are more susceptible to the spread and may require focused measures to avoid further spread.

D. Simulating meme spreading using SI epidemic spreading model with two different centrality measures

1) Analyzing the effect of node removal on meme spreading within the largest community: The top five largest communities in the network are determined using the Louvain method for community detection, and a random node is selected for each community to serve as the initial infected node for the SI model. The spread of the memes inside the network is simulated using the EoN (Epidemics on Networks) programme, and the number of infected nodes for each community are noted at each time step. Randomly selecting an initial infected node within a community helps to investigate how a community's size and structure affect the rate and extent of meme transmission [10] .

A random network with a predetermined number of nodes and a probability of p that any two nodes are connected by an edge is generated by the Erdos-Renyi model. On the other hand, the GNM random graph model creates a random graph with a predetermined number of nodes and edges. In order to produce random graphs with the same amount of nodes and edges as the original network, both models can be used [6].

The graphs derived from the original network and the random graph differ in that the nodes in the random graph are distributed more uniformly among communities. This is due to the possibility that the original network's design reflected some underlying social or geographic organization, which may have an impact on how memes spread among different communities. A random graph, on the other hand, lacks this structure, and the memes will spread more evenly among communities.

2) Analyzing the effect of node removal of highest eigen vecter centrality on meme spreading within the largest community: Running a SI model with the first infected node randomly selected from the largest community after randomly eliminating a portion of the network's nodes. The spread of memes throughout communities slows as the fraction of removed nodes rises. This is to be expected as the network becomes less connected as nodes are removed, making it more difficult for memes to spread [8]. Although not all communities experience the same pace of decline, some are more resistant to node removal than others.

Eigenvector centrality is a measure that considers a node's neighbors' centrality and gives a higher score to nodes that are connected to other highly central nodes. finding the nodes with the highest eigenvector centrality using the NetworkX library's nx.eigenvector centrality function, which sorts the nodes based on their centrality score and removes the top X % of nodes. We may run the SI model simulation again after eliminating the nodes.

By eliminating nodes with high eigenvector centrality (task (b)), we can see that removing highly central nodes has a greater impact on meme transmission than random node removal. Specifically, removing the top 5% of nodes with high eigenvector centrality results in a faster and significant decrease in the number of infected nodes when compared to removing the same proportion of random nodes.

IV. Results and Discussion

A. Community detection algorithm with 5 different "Resolution" values

The strength of a network's separation into communities or clusters is evaluated using modularity in the context of network analysis. High modularity networks have sparse connections between communities but dense connections within communities. A network with low modularity, on the other hand, has fewer clearly defined groups and more links between them. Higher resolution will produce smaller and more communities, whilst lower resolution would produce larger and fewer communities.Our goal is to develop communities with few connections between them and many connections inside them.

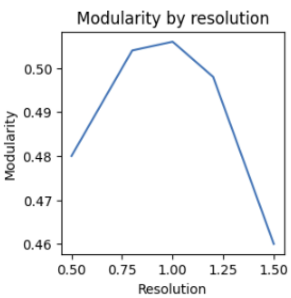


Fig. 1: Modularity vs resolution

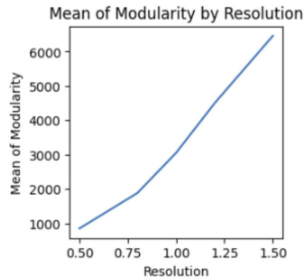


Fig. 2: Mean of Modularity vs Resolution

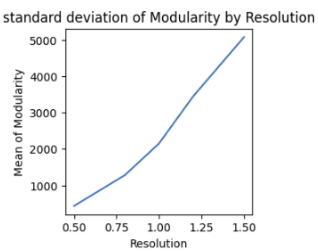


Fig. 3: Standard Deviation of Modularity vs Resolution

The mean size of communities in a network with low modularity is often larger, indicating fewer, larger, and less coherent communities.The mean size of communities in a network with high modularity tends to be lower, indicating a greater number of tightly connected groups.A low standard deviation suggests that the community sizes in the network are relatively comparable, whereas a high standard deviation shows that the community sizes vary greatly [4].

This reveals more specific patterns or structures are identified when the resolution is 1.0, and there are 541 numbers of communities which provide insights into the underlying processes or mechanisms that control the network's behavior.

B. SI models and centrality measures

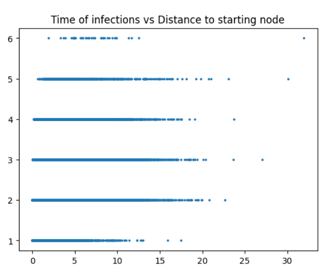


Fig. 4: High eigenvector centrality

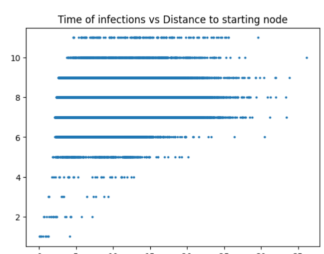


Fig. 5: Low eigenvector centrality

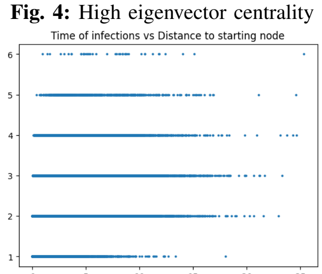


Fig. 6: High betweeness centrality

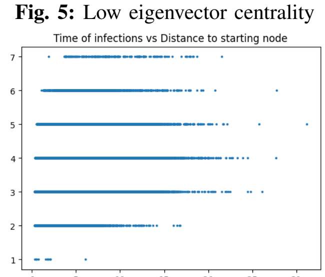


Fig. 7: Low betweeness centrality

In these figures we see each node represented as a point in the scatter plot as a data point with there time of infection and distance from the starting node. First of all we should note that the starting node for highest eigenvector centrality and highest inbetweeness centrality are actually the same which explains how similar they look. Another important note is that the starting node for both of the lowest centralities are not actually the lowest scoring central nodes as the least central node is actually one of two nodes in a disconnected pair with no other connections meaning the meme could not spread to the rest of the network which is not an interesting comparison.

Now if we actually compare the figures we can see that the when choosing a central node that the meme spreads quicker and that the average distance between each node and the starting node is shorter which is to be expected. More interestingly when we compare the two different centrality measures it seems having a lower inbetweeness centrality was not as detrimental to the spread of the meme. The difference between the spread of using a starting node with high or low betweeness centrality is actually quite small. This leads us to think that the eigenvector centrality of the starting node is a better indicator of the spread of a meme across a network.

1) Plotting Distance for nodes in the community over infection time: The locations of the initial infected node and time steps were plotted on a scatter plot, with the x-axis representing distances. Every data point corresponded to a particular node in the network, and the colour denoted the centrality metric applied during the simulation. We utilised the graphic to look for any patterns or trends in the data, such as whether nodes with higher betweenness or eigenvector centrality tended to have lower time values or shorter distances.

C. Spreading Analysis within and across 5 Largest Communities with SI Model and Random Graph Simulation

1) Analysis of Spreading Patterns in a Reddit Network Within and Across Communities: The first infected node is a randomly chosen node from the community

Community	Size	Clustering	Average centrality
0	6824	0.2	0.0013
1	5786	0.19	0.0008
2	2752	0.14	0.0014
3	2222	0.14	0.0014
4	1725	0.22	0.0032

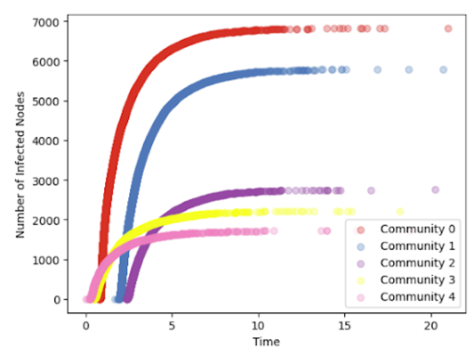
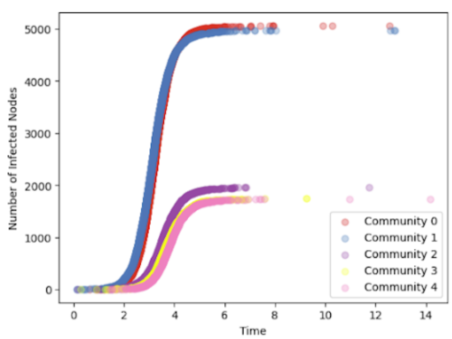


Fig. 8: Number of infected nodes over time in each largest community

The SI model was run on five largest communities with a randomly chosen infected node, and the number of nodes infected (Nc) was plotted against time (t) for each community. The analysis showed that the rate of spreading varied across communities, with Community 0 having the fastest spread and Community 4 having the slowest [15]. Additionally, the communities with higher clustering coefficients and average centrality tended to have faster spread rates. The results suggest that the structure of the network and the location of the infected node can significantly affect the spread of the virus within and across communities.

Community	size	clustering	Average centrality
0	5060	0.00015	0.0007
1	4976	0.0005	0.0007
2	1958	0.00	0.0014
3	1742	0.00016	0.0015
4	1731	0.0017	0.0015

2) Impact of Community Properties on Meme Spreading: A Comparison between Real and Random Networks:



The results of making a random graph for each of the five communities that had the same amount of edges and nodes as the original network are displayed in the output above. The random graphs exhibit significantly lower clustering coefficients than the original network, which suggests that nodes are less likely to be connected to their neighbors [5]. In the SI model simulations, the random graphs show a slower spread of the meme compared to the original network for all communities except for Community 4, which showed similar spreading dynamics.The attributes of each community, including their size, clustering coefficient, and average centrality, can be used to explain the variations in spreading. The dynamics of spreading within larger communities with higher clustering coefficients and average centrality tend to be faster than those within smaller communities with lower clustering coefficients and average centrality. These findings collectively imply that a community's characteristics can have a significant impact on the transmission of a meme both inside and between communities.

The analysis of the SI model simulations for each of the five largest communities in the partitioned network, along with their characteristics of size, clustering, and average centrality, demonstrates that larger and more clustered communities typically have a slower spread of the meme, whereas smaller and less clustered communities typically have a faster spread. This shows that interactions among communities are key to the meme's ability to propagate [11]. The findings of the simulations using random graphs also show that the meme spreads more slowly in the original communities than in the simulations using random graphs, indicating that the original communities' characteristics have an effect on the meme's spread. In general, the characteristics of each community can significantly influence the transmission of the meme, both inside and beyond communities, and an understanding of these characteristics can aid in the development of efficient meme propagation tactics.

D. SI Model Analysis on Random and Highest Eigenvector Centrality Nodes Removal in the Largest Community

1) Impact of Node Removal on Spreading within the Largest Community: Removing the network's node.

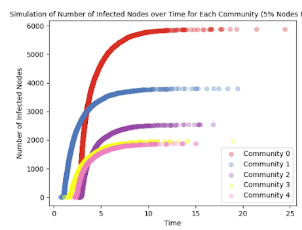


Fig. 9: Removal of 5% of Node Simulation Infected vs Time

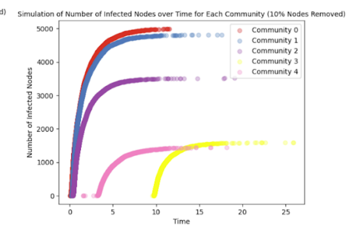


Fig. 10: Removal of 10% of Node Simulation Infected vs Time

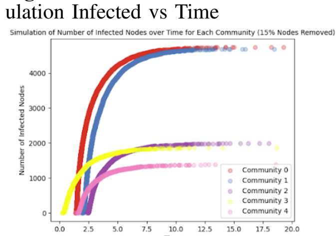


Fig. 11: Removal of 15% of Node Simulation Infected vs Time

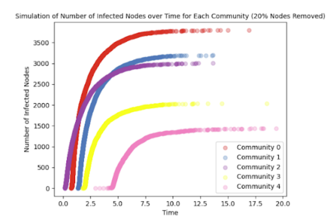


Fig. 12: Removal of 20% of Node Simulation Infected vs Time

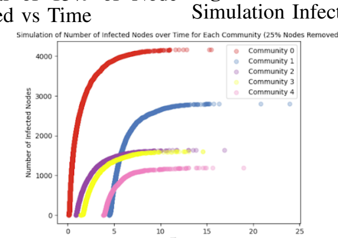


Fig. 13: Removal of 25% of Node Simulation Infected vs Time

As more nodes are removed randomly from the network, the size of the communities decreases, along with a decrease in clustering coefficient and average centrality. Specifically, the largest community size reduces from 6904 to 3983, and the average centrality drops from 0.0014 to 0.0012, when the number of removed nodes increases from 5% to 25%. Interestingly, the clustering coefficient initially decreases but then slightly increases in the case of 20% and 25% node removal, which could be attributed to the formation of tightly connected small clusters among the remaining nodes, These findings demonstrate the network's resilience to the random node removal and provide insights into its performance under different stress conditions [2].

2) Node Removal Based o n Eigenvector Centrality: Removing the Nth highest eigenvector centrality.

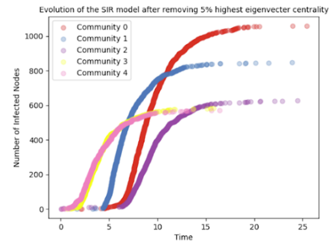


Fig. 14: SIR Model Removal of highest eigenvector centrality 5%

Fig. 15: SIR Model Removal of highest eigenvector centrality 10%

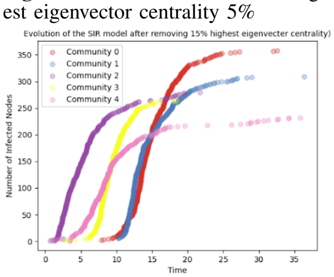


Fig. 16: SIR Model Removal of highest eigenvector centrality 15%

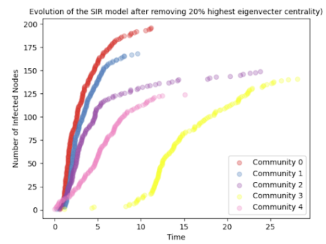


Fig. 17: SIR Model Removal of highest eigenvector centrality 20%

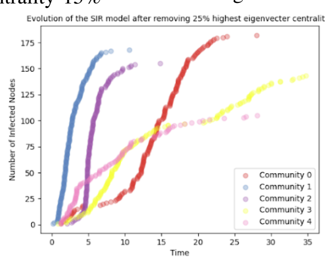


Fig. 18: SIR Model Removal of highest eigenvector centrality 25%

The results suggested that removing nodes with high eigenvector centrality significantly impacts community structure, as community size decreases and the average centrality of nodes increases with the removal of more nodes. However, the clustering coefficient shows a more complex pattern, with some communities exhibiting high clustering and others showing low clustering. This indicates that removing high centrality nodes may not always lead to a more uniform distribution of connections within communities[.

To identify key nodes that facilitate the flow of memes between communities, we can use the betweeness centrality metric, which quantifies the number of shortest paths passing through a node and reveals the node's significance in linking different parts of the network [7]. Nodes with high betweeness centrality are likely to be critical nodes that enable the flow of memes between communities.

Other network metrics that could be useful in this context include degree centrality, which measures the number of connections a node has, and closeness centrality which measures how close a node is to all other nodes in the network [9]. Additionally, community detection algorithms such as modularity can identify communities within the network and help analyze the flow of memes within and between communities.

V. Conclusion

In conclusion,The study provides insight into the dynamics of meme propagation in a network and how different network characteristics influence its spread on social networks. Overall, the analysis indicated how each community's characteristics (size, clustering, average centrality) may impact meme dissemination inside, away from, or towards the community. Metrics of centrality, such as eigenvector and betweenness centrality, are used to identify the critical nodes that allow memes to move between communities.

When random nodes are removed the change in the density of the network and the centrality measurements is minimal. Removing prominent nodes from a network can have a major influence on the network topology. Betweenness, or eigenvector centrality, refers to nodes that are crucial to the network's connectedness.Nodes with high eigen vector centrality are sometimes referred to as "hubs" because they play an important role in linking various portions of the network.When such nodes are eliminated, the network may become less linked or fragmented, disrupting information flow and weakening overall network connection.Understanding the impact of node removal on network topology is critical for forecasting network resilience and stability, as well as creating effective techniques to limit or prevent information propagation in the network.

Node centrality properties compute the logical inference of a node's significance in a network. Central nodes have a substantial impact on the flow of information. Removing nodes can affect shattering and loss of connectivity, leading to diminished proficiency in terms of resource allocation and data transmission. It's important to understand the consequences of removing nodes when examining networks.

A network's structure and performance can be significantly impacted by the removal of a large portion of its nodes. While removing nodes based on eigenvector centrality can result in a more intricate pattern of community structure changes, random node removal can result in a decrease in community size, clustering coefficient, and average centrality. Designing and enhancing communication techniques in social networks might benefit from identifying key nodes that support information flow and understanding the network's resilience under stress circumstances.

Our findings imply that social network analysis might be an excellent technique for understanding the dynamics of information transmission and developing successful information dissemination methods. The findings of this study can benefit policymakers, marketers, and other stakeholders in their efforts to use social networks for a variety of reasons, including public health initiatives, commercial campaigns, and political campaigns.

References

[1] *2016 IEEE Conference on Intelligence and Security Informatics (ISI)*. IEEE Press, 2021.

[2] Lada A Adamic and Eytan Adar. Friends and neighbors on the web. *Social Networks*, 25(3):211–230, 2003.

[3] Aseel Addawood, Rezvaneh Rezapour, Omid Abdar, and Jana Diesner. Telling apart tweets associated with controversial versus non-controversial topics. In *Proceedings of the Second Workshop on NLP and Computational Social Science*, pages 32–41, Vancouver, Canada, August 2017. Association for Computational Linguistics.

[4] Amy Binns. Don't feed the trolls! *Journalism Practice*, 6(4):547–562, 2012.

[5] Ian Cooper, Argha Mondal, and Chris G. Antonopoulos. A sir model assumption for the spread of covid-19 in different communities. *Chaos, Solitons Fractals*, 139:110057, 2020.

[6] Srayan Datta, Chanda Phelan, and Eytan Adar. Identifying misaligned inter-group links and communities. *Proc. ACM Hum.-Comput. Interact.*, 1(CSCW), dec 2017.

[7] William L. Hamilton, Rex Ying, and Jure Leskovec. Representation learning on graphs: Methods and applications, 2018.

[8] Saike He, Hongtao Yang, Xiaolong Zheng, Bo Wang, Yujun Zhou, Yanjun Xiong, and Daniel Zeng. Massive meme identification and popularity analysis in geopolitics. In *2019 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pages 116–121, 2019.

[9] Kristina Lerman, Rumi Ghosh, and Tawan Surachawala. Social contagion: An empirical study of information spread on digg and twitter follower graphs, 2013.

[10] Ka Yin Leung, Pieter Trapman, and Tom Britton. Who is the infector? epidemic models with symptomatic and asymptomatic cases. *Mathematical Biosciences*, 301:190–198, 2018.

[11] Adam Lonnberg, Pengcheng Xiao, and Kathryn Wolfinger. The growth, spread, and mutation of internet phenomena: A study of memes. *Results in Applied Mathematics*, 6:100092, 2020.

[12] Elaine O. Nsoesie, Richard J. Beckman, Sara Shashaani, Kalyani S. Nagaraj, and Madhav V. Marathe. A simulation optimization approach to epidemic forecasting. *PLOS ONE*, 8(6):1–10, 06 2013.

[13] Joseph D O'Brien, Ioannis K Dassios, and James P Gleeson. Spreading of memes on multiplex networks. *New Journal of Physics*, 21(2):025001, feb 2019.

[14] Songsri Tangsripairoj, Pasathorn Luengpattana, Teerapat Sriboonsong, and Thanapat Piyakulpinyo. Meme scraper: A web application for internet meme spread pattern analysis. In *2022 6th International Conference on Information Technology (InCIT)*, pages 35–38, 2022.

[15] Lilian Weng, Filippo Menczer, and Yong-Yeol Ahn. Virality prediction and community structure in social networks. *Scientific Reports*, 3(1), aug 2013.