Noname manuscript No.
(will be inserted by the editor)

SIRA: A Model for Propagation and Rumor Control with Epidemic Spreading and Immunization for Healthcare 5.0

Akshi Kumar · Nipun Aggarwal · Sanjay Kumar

Received: date / Accepted: date

Abstract Healthcare social networks have a significant role in providing connected and personalized healthcare environment with real-time capabilities. However, building resilient, robust and technology-driven healthcare 5.0 has its own barriers. Especially, with social media's high susceptibility to rumors and fake news, these networks can harm the society. Many researchers have been investigating the process of information diffusion, and it has been one of the most intriguing issues in network analysis. Modeling rumor propagation is one of the prominent researched topics in recent years. Traditional models assume that rumor propagation happens only in one direction, where only supporters are supposed to be active whereas, in a real-life situation, both supporters and deniers of the information operate simultaneously. In this paper, we introduce a model for the recovery of nodes in a setting where rumor propagation and rumor control happen simultaneously. We propose the Susceptible-Infected-Recovered-AntiSpreader ($SIRA$) model based on the notion of spreading of epidemics and also its applications to modeling the propagation of rumors and control of rumor. Our model assumes people have multiple forms of reactions to rumor, either posting it, deleting it, or announcing the rumor as fake. This paper also suggests how the model can act as a simulation method to compare two node centrality algorithms where spreaders chosen from one centrality al-

Akshi Kumar
Department of Computing and Mathematics, Faculty of Science and Engineering, Manchester Metropolitan University, Manchester, United Kingdom
E-mail: akshi.kumar@mmu.ac.uk

Nipun Aggarwal
Department of Computer Science and Engineering, Delhi Technological University, New Delhi-110042, India
E-mail: nipundtu@gmail.com

Sanjay Kumar
Department of Computer Science and Engineering, Delhi Technological University, New Delhi-110042, India
E-mail: sanjay.kumar@dtu.ac.in

gorithm try to spread the rumor, and the anti-spreaders chosen from other centrality try to dispel the rumor and vice-versa. We simulate the proposed algorithm on different weighted and unweighted real-world network data-sets and establish that the experimental results agrees with the proposed model.

Keywords Complex network · Healthcare social networks · Information diffusion · Node centrality · Rumor Control · Rumor propagation

1 Introduction

The 5.0 revolution in healthcare is eventually digitally transforming society, plummeting the overload on emergency and critical care, and letting people to reinforce their resilience and health with a faster response time. The use of AI, the Internet of Medical Things (IoMT) and social media is offering number of advantages over traditional analytic and clinical decision-making techniques. In this view, many Healthcare Social Networks (HSN) are evolving with the intent to augment patient care and awareness enticing the interest of clinical professionals and of the whole healthcare industry. Technological developments and digital platforms foster information diffusion possibilities to keep people safe, informed and connected. However, the same tools also expose possible perils of misinformation for patients. As the world combats with the outrageous and perilous coronavirus, the pandemic has demonstrated how the spread of misinformation, amplified on social media and other digital platforms, is proving to be as much a threat jeopardizing measures to control the pandemic.

Typically, rumors are information pieces that are unreliable, incorrect, and spread on the network due to limited knowledge of the recipient nodes or wrong intentions [3,4,5]. Rumors are the unverified spread of societal concerns, issues by different methods, and create misunderstandings between individuals. Rumors on the internet spread quicker and have a more extensive scope of an impact than word-of-mouth rumors. These rumors spread by word of mouth over a small area and in vast geographical regions like a forest fire on social networks [6], generally with malicious intent. Rumors alter and shape public opinion adversely affecting human behavior, causing multiple confusions [7] and instability. Many researchers have been investigating the process of information diffusion, and it has been one of the most intriguing issues in network analysis. Modeling rumor propagation is one of the prominent researched topics in recent years. The SIR [8] model is a traditional model that simulates information diffusion due to similarity in the field of epidemiology and social networks. It has been utilized by numerous researchers to model the process of information diffusion in social networks. A susceptible individual spreads data when influenced by a spreader and turns into a spreader. After some time, the spreader quits spreading the data and turns into a recovered node. The transformation parameters β and γ in the process are network characteristics-based probabilities, and the state transformation of all nodes is equivalent. While the SIR model serves as an excellent benchmark for information propagation, it doesn't generalize well to rumor propagation. A notable difference arises because individuals may also be immunizing other nodes regarding the rumor and preventing them from getting infected. This phenomenon is very different from actual epidemics modeling because, in an epidemiological setting, it is unlikely that a large portion of people who recover from the epidemic, make other people immune as well. However, in a social network, when a person propagated fake news intentionally or due to a lack of knowledge and after getting the truth, other persons can disseminate the fact about the misinformation. This makes their susceptible neighbors, directly recovered and immune to the rumor. We consider the above recovery aspects of information propagation in our paper and propose the Susceptible-Infected-Recovered-AntiSpreader ($SIRA$) model based on the notion of spreading of epidemics and its applications to modeling the propagation of rumors and control of rumor. Our model assumes people have multiple forms of reactions to rumor, either posting it, deleting it, or announcing the rumor as fake. The contribution of our papers are as follows:

- We establish mathematical model for the rumor propagation, rumor control and analyzing its equilibrium parameters and reproduction rate.
- The proposed model introduces the notion of anti-spreaders that immune its susceptible neighbors from the rumors to minimize the effect of rumors in the network.
- We consider the case where the information is not a rumor, and the proposed model simplifies to information propagation model for the influence maximization.
- We discuss the novel approach of directly comparing two different node centrality algorithms for the selection of spreaders and anti-spreaders. The spreaders, chosen from one centrality algorithm, try to spread the rumor, and the others try to dispel the rumor and vice-versa.

The paper organization is as follows. In Section 2, we present some traditional models used to study rumor propagation and clarify the enhancements that our model proposes. In Section 3, we describe the proposed SIRA model of rumor propagation and rumor control with mathematical modeling using the system of differential equations. This section also explores the analysis of the proposed model on homogeneous networks, calculation of reproduction number, and how our model can be simplified to a general information diffusion model as the SEI model followed by the simulation of the model on a toy network. In Section 4, we discuss the different weighted and unweighted real-world data-sets used in this work, and we prove the mathematical integrity of our model for information propagation, rumor propagation, rumor control, and node centrality algorithm performance comparison on various data-sets used. Finally, Section 5 concludes the paper.

2 Related Work

Social networks are complex networks with non-trivial topology and intricate interactions between users. The effective analysis of complex systems is essential to model and comprehend the method of rumor propagation. The spreading of rumors on online social networks is very similar to epidemic spreading. Researchers from various disciplines have extensively studied models of rumor propagation in complex networks. The majority of the conventional models for investigating information diffusion depend on elementary models of epidemics, for example, SI (Susceptible Infected), SIS (Susceptible Infected Susceptible), SIR (Susceptible Infected Recovered) [9,10].

Tripathy et al. [11] claimed that the effect of rumor can be combated by injecting messages called anti-rumors, just like the rumor process. They studied two metrics, namely the belief time, which is the duration for which a person believes the rumor to be true, and point of decline, which is a point after which the anti-rumor process dominates the rumor process. Huang and Jin [12] modified the SIR model to prevent the rumor propagation by devising immunization strategies, namely, the random immunization and the targeted immunization to the rumor model on small-world networks. Wan et al. [13] introduced the SIERsEs model, which is an extension of the SIR spreading model with the addition of the elimination of the rumor process. Zhu and Ma argue that certain people may not believe in the rumor and do not spread it, a different state called a hesitant state added to the traditional SIR model and proposed SHIR model [14]. This model also takes care of the varying degree of nodes, which commonly happens in online social networks. Indu and Thampi argue that the spread of rumors in social networks is similar to the spread of fire in a forest and identified various components to model the spread of rumors in online social networks [15]. Wang et al. [16] explored the propagation of rumor using an improved energy model applicable to multi-layer social networks. The model also introduces negative energy to simulate the mitigation process of rumors. Zhu et al. [17] recently proposed a SIR like an epidemic rumor propagation model by employing forced silence function, time delay, and network topology. Recently, Yang et al.[18] developed an ILSR rumor propagation model based on the notion of the degree of different nodes in networks. They introduced a new type of users known as lurkers users who have heard the rumors, but temporarily not spreading rumors in the system. Kaur et al. [19] proposed a fake news detection system using multi-level voting model by exploiting various feature extraction techniques and machine learning classifiers. Qui et al. [3] introduced the SIR rumor spreading model with an influence mechanism in social networks, named SIR-IM, by integrating the number of current spreaders into the spreading probability and time function.

3 Proposed Methodology

3.1 $SIRA$ Model description

We propose the Susceptible-Infected-Recovered-AntiSpreader ($SIRA$) model for the recovery of nodes in a setting where rumor propagation and rumor control happen simultaneously. The entire population is divided into four categories:- Susceptible nodes (S), that are unaware of the rumor, Infected nodes (I) that are actively spreading the rumor, Recovered nodes (R), that are not spreading the rumor, and Anti-spreader nodes (A) that are dispelling the rumor. We consider people have multiple forms of reactions to rumor, either posting it, deleting it, or announcing the rumor as fake. We introduce the notion of anti-spreaders nodes, such nodes that get to know about the irrelevance of the information, and the piece of the information is a rumor, tend to dispel the rumor by announcing the truth. Hence, the Anti-spreaders nodes are immunizing their susceptible neighbor's nodes from false information and play a crucial role in controlling the rumor. As evident from the Fig. 1, solid

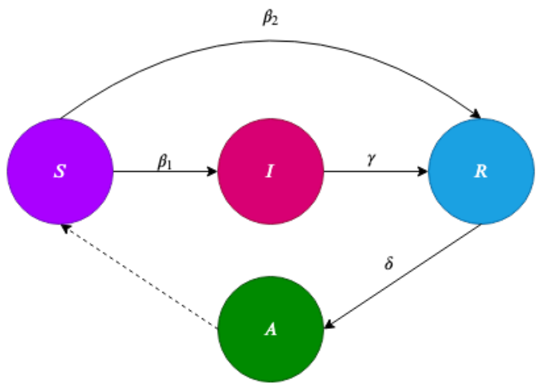


Fig. 1: $SIRA$ Model

lines represent direct conversion of a node from one state to another, for examples $S- > A$, $I- > R$. However, the dotted line corresponds to indirect conversion from one state to other. For example, the conversion from A to S signifies the indirect effect of anti-spreaders on their susceptible neighbors. This makes susceptible neighbors transition to the recovered state with the rate β_2 and becomes immune to the fake news. The steps of the algorithm are as follows:

- Node population starts from certain infected nodes and certain anti-spreader nodes that spread the rumor or dispel it, and all other nodes are susceptible
- Infected nodes affect their susceptible neighbors and convert them into infected members that propagate the rumor. This conversion happens with a probability β_1

- Probability that the infected members recover is at a rate of γ
- Under these conditions recovered nodes can now become anti-spreaders at the rate of δ
- Nodes that become Anti-spreaders affect β_2 fraction of their susceptible neighbors, converting them directly into recovered state and immune to the rumor. Due to the influence of Anti-spreaders nodes, some of their susceptible neighbors can directly go to a recovered state, and they will not be affected by the rumor.

Also, other than the rate of conversion β_1 from $S- > I$, the nodes in susceptible state can only transit to an infected state, if they are direct neighbors of an infected node. Similarly, other than the rate of conversion β_2 from $S- > R$, the susceptible nodes can directly transition into recovered nodes if they are direct neighbors of anti-spreader nodes.

At the end of simulation when there are no infected nodes i.e., $I(t) = 0$ and equilibrium is reached then some nodes can still remain in susceptible state other then anti-spreaders. It is not like that the final state is anti-spreader, and all the nodes should be in this state at the end; rather, it depends on the various conversion rates considered. Also, other than the rate of conversion β_1 from $S- > I$, the nodes in susceptible state can only transit to an infected state, if they are direct neighbors of an infected node. Similarly, other than the rate of conversion β_2 from $S- > R$, the susceptible nodes can directly transition into recovered nodes if they are direct neighbors of anti-spreader nodes.

3.2 System of differential equations

The population is divided into the 4 states Susceptible S, Anti-Spreaders A, Infected I and Recovered R and at any time t their sum is 1 as in Eq. No. 1.

$$S(t) + A(t) + I(t) + R(t) = 1 \tag{1}$$

As evident from Fig. 1, some recovered nodes may become anti-spreaders at the rate δ as per the Eq. No. 2. Anti-spreaders are those nodes who know the ingenuity of the information. The node in the anti-spreader(A) state continues to be in the same state throughout.

$$\frac{d}{dt}A(t) = \delta R(t) \tag{2}$$

Recovered nodes increase through multiple methods, either through the direct recovery of susceptible nodes due to influence of anti-spreaders (edge between S to R in the Fig. 1) or by direct conversion from infected to recovered (edge between I to R in the Fig. 1). Some recovered nodes may also become anti-spreader at the rate δ. The change in recovered nodes is presented in Eq. No. 3.

$$\frac{d}{dt}R(t) = -\delta R(t) + \gamma I(t) + \beta_2 * k * S(t) * A(t) \tag{3}$$

where k is the average degree of the nodes in the network. Change in infected nodes, Eq 4, arises from the conversion of susceptible nodes to infected nodes with the rate β_1 and infected nodes go to recovered state with the rate γ.

$$\frac{d}{dt}I(t) = -\gamma I(t) + \beta_1 * k * S(t) * I(t) \tag{4}$$

$$\frac{d}{dt}S(t) = -\beta_1 * S(t) * I(t) * k - \beta_2 * S(t) * A(t) * k \tag{5}$$

Initially, all the nodes are assumed susceptible except few seed nodes and slowly, they decay into infected state with rate β_1 and recovered state with rate β_2 as given in the Eq. No. 5. Now Let us solve these equations by substituting $R(t)$ with $1 - S(t) - A(t) - I(t)$ from Eq. No. 3 and hence getting the system given by 6, 7, 8.

$$\frac{d}{dt}A(t) = \delta(1 - S(t) - A(t) - I(t)) \tag{6}$$

$$\frac{d}{dt}I(t) = -\gamma I(t) + \beta_1 * k * S(t) * I(t) \tag{7}$$

$$\frac{d}{dt}S(t) = -\beta_1 * S(t) * I(t) * k - \beta_2 * S(t) * A(t) * k \tag{8}$$

3.3 Calculation of reproduction number

The basic reproduction number (R_0) is an important parameter in epidemics which measures the propagation potential of contagious diseases. The reproduction number (R_0) is an epidemiologic measure used to characterize the contagiousness or transmissibility of infectious agents. We have applied a similar concept for the rumor propagation. Since for successful reproduction of the rumor, we require that initially $\frac{d}{dt}I(t) > 0$, this produces Eq. No. 9

$$-\gamma I(t) + \beta_1 * k * S(t) * I(t) > 0 \tag{9}$$

Thus, giving us Eq. No. 10,

$$R_t = (\beta_1 * k * S(t)/\gamma) > 1 \tag{10}$$

$R_t > 1$ means that the rumor is active in the system. For the initial spread of rumor, $R_0 > 1$ and since $S(0) \cong 1$ we get Eq. No. 11

$$R_0 = (\beta_1 * k/\gamma) > 1 \tag{11}$$

Now, when $R_t < 1$, there is no rumor in the system. When $R_t < 1$ then $\frac{d}{dt}I(t) = 0$. This implies there is no infected node in the system and hence there is no rumor in the system.

3.4 Model Analysis in a homogeneous network

We analyze the proposed $SIRA$ model on a homogeneous network and prove the validity of the model at the stable equilibrium points. An equilibrium point of a dynamic mathematical framework produced by an independent system of ordinary differential equation (ODEs) is a solution point where the count of participating entities doesn't change with time and hence, their rate of change becomes 0 [31]. In the case of epidemics, the disease-free equilibrium is asymptotically stable, and the infection asymptotically vanishes. Similarly, in the case of information propagation, rumor-free stable equilibrium is stable, and information is no longer being spread, and the network has no active rumor propagating nodes. We prove the stability of our differential equation at the equilibrium condition given by Eq. No. 12

$$\frac{d}{dt}S(t) = 0, \frac{d}{dt}A(t) = 0, \frac{d}{dt}I(t) = 0, \frac{d}{dt}R(t) = 0 \tag{12}$$

The Jacobian Method [32] as given by Eq. No. 13. The Jacobian method is a popular iterative technique used for determining the solutions for the system of linear equations, which is diagonally dominant.

$$J_E = \begin{bmatrix} -\beta_1 * I * k - \beta_2 * A * k & -\beta_2 * S * k & -\beta_1 * S * k & 0 \\ 0 & 0 & 0 & \delta \\ \beta_1 * k * I & 0 & -\gamma & 0 \\ \beta_2 * k * A & \beta_2 * S * k & \gamma & -\delta \end{bmatrix} \tag{13}$$

The equilibrium points exist if I(t) = 0 and under this condition we have Eq. No. 14

$$E_i = (S, A, 0, R) \tag{14}$$

The Jacobian matrix is given by Eq. No. 15:

$$J_E = \begin{bmatrix} -\beta_2 * A * k & -\beta_2 * S * k & -\beta_1 * S * k & 0 \\ 0 & 0 & 0 & \delta \\ 0 & 0 & -\gamma & 0 \\ \beta_2 * k * A & \beta_2 * S * k & \gamma & -\delta \end{bmatrix} \tag{15}$$

We will calculate the eigenvalues of J_E as in Eq. No. 16:

$$J_E = \begin{vmatrix} -\beta_2 * A * k - \lambda & -\beta_2 * S * k & -\beta_1 * S * k & 0 \\ 0 & -\lambda & 0 & \delta \\ 0 & 0 & -\gamma - \lambda & 0 \\ \beta_2 * k * A & \beta_2 * S * k & \gamma & -\delta - \lambda \end{vmatrix} \tag{16}$$

Expanding on the third row, we arrive at the following characteristic equation in Eq. No. 17:

$$(\gamma + \lambda)(\lambda)(\lambda^2 + (\delta - \beta_2 k A)\lambda - \beta_2 k \delta (A + S)) \tag{17}$$

As evident, $(-\gamma, 0)$ are eigen values and also the solutions of $\lambda^2 + (\delta - \beta_2 kA)\lambda - \beta_2 k\delta(A+S)$ which we find and analyze as follows using the Quadratic Formula:

$$\begin{cases} (\lambda^2 + (\delta - \beta_2 kA)\lambda - \beta_2 k\delta(A+S)) \\ D = ((\delta - \beta_2 kA)^2 + 4 * \beta_2 k\delta(A+S)) \\ \therefore \delta^2 + \beta_2^2 k^2 A^2 - 2\delta\beta_2 kA + 4\beta_2 \delta kA + 4\beta_2 \delta kS \\ \cong (\delta + \beta_2 kA)^2 + 4 * \beta_2 k\delta S \end{cases} \tag{18}$$

.
Here, since, $(\delta + \beta_2 kA)^2 \geq 0 \because (x^2 > 0, \forall x \in R)$ and also $S, k, \beta_2, \delta \geq 0$, we have Eq No. 19:

$$D \geq 0 \tag{19}$$

Thus the Eigenvalues are given by Eq. No. 20:

$$(0, -\gamma, (-(\delta - \beta_2 kA) - \sqrt{D})/2, (-(\delta - \beta_2 kA) + \sqrt{D})/2) \tag{20}$$

Now, for a stable equilibrium, real eigenvalues must be ≤ 0 and since $\gamma, \delta, \beta_2 > 0$ and $S, I, R, A \geq 0$, we know that eigenvalues satisfy these properties in the assumptions of our system for equilibrium. The first eigenvalue is 0 and $-\gamma$ and $-(\delta - \beta_2 kA) - \sqrt{D})/2$ are ≤ 0 as given by Eq. 21.

$$(-(\delta - \beta_2 kA) + \sqrt{D})/2 \leq 0 \tag{21}$$

Thus rearranging and squaring both sides, we get to Eq. No. 22:

$$(\delta - \beta_2 kA)^2 \leq (\delta - \beta_2 kA)^2 + 4\beta_2 \delta kA + 4\beta_2 \delta kS \tag{22}$$

Hence, Eq. No. 22 holds true $\forall A, S \geq 0$, and at this point, the mathematical stability of the system is proved quantitatively by finding the eigenvector related with every eigenvalue. An equilibrium point is hyperbolic if none of the eigenvalues have zero real part. One of our eigenvalues is zero. Thus the equilibrium is non-hyperbolic. If all eigenvalues have a negative real part, the equilibrium is steady, as given by the Lyapunov Stability theorem [33]. As evident in Eq. No. 20 - 22 that the corresponding eigenvalues are ≤ 0, and the equilibrium for our model is stable.

3.5 The simplification to information propagation SEI model

When the information is non-rumor, the model can directly be simplified to an information propagation model that leads to influence maximization achieved through anti-spreaders.

The roles of the nodes however remain almost the same. Anti-spreaders are still the carriers of correct information and the recovered nodes, still are the dormant nodes that know about the information but are not spreading it. Similarly, susceptible nodes still remain the nodes that are unaware of the

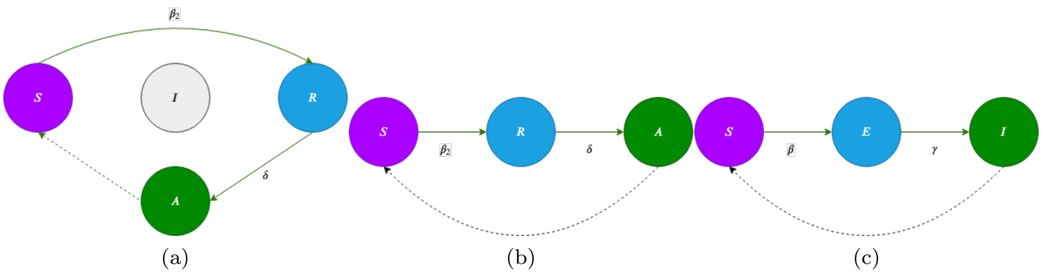


Fig. 2: Simplification of the model from $SIRA$ to SEI for non-rumor information

information and can likely be turned into recovered nodes by anti-spreaders. This effect can be seen when we plug in $I(t) = 0$, $\frac{d}{dt}I(t) = 0$ into the model (Eq. 1, 2, 3, and 5) and we get the following system of equations namely, Eq. No. 23 24, 25, and 26.

$$S(t) + A(t) + R(t) = 1 \tag{23}$$

$$\frac{d}{dt}A(t) = \delta R(t) \tag{24}$$

$$\frac{d}{dt}R(t) = -\delta R(t) + \beta_2 * k * S(t) * A(t) \tag{25}$$

$$\frac{d}{dt}S(t) = -\beta_2 * S(t) * A(t) * k \tag{26}$$

For the purpose of an analogy between these states we rename the states for direct correspondence with the SEI model as follows:

1. Anti-spreader nodes (A) become source spreaders or seed nodes (I)
2. Recovered nodes (R) are renamed as exposed nodes (E)
3. Susceptible nodes remain in the same naming convention (S)

These seed nodes (I) affect β_2 fraction of their neighbors, and they convert into exposed (E). Exposed users can become spreaders of the information with the rate δ.

Thus, the proposed $SIRA$ model for the rumor propagation and rumor control simplifies to information spreading SEI [34] model for influence maximization. This conversion is depicted in Fig. 2

3.6 Simulation of the proposed model on a toy network

A toy network, as in Fig. 3 is chosen as a simulation network. The most influential node in the network is chosen as the spreader and $\beta_1 = 0.5, \beta_2 = 1, \gamma = 1, \delta = 1$ are assumed without loss of generality. Initially, all the nodes, except the one chosen as the influential node based on Out-Degree, are susceptible, and the given node is in an Infected state. As time progresses, random neighbors of this node become infected, and the node recovers. After recovery, it is dormant, and other nodes continue spreading information. The recovered node decides to dispel the rumor and makes itself an anti-spreader. This effect makes all its susceptible neighbors instantly recover. The spreading continues till all the nodes are either against the rumor or have never heard of the rumor or its dispel. We describe the spread in Fig. 3 as a change-of-state in an event-driven simulation.

The change of state in an event-driven simulation are as follows:
t1: Node F is infected, chosen because of maximum out-degree, all other nodes are currently susceptible.
t2: Node F has recovered, and has infected its neighbors with a uniform probability of $\beta_1 = 0.5$, thus randomly infecting nodes E and J.
t3: Node F is now an anti-spreader, all the infected nodes recover at the rate of $\gamma = 1$, and their neighbors are infected at the rate β_1.
t4: Node F has recovered its susceptible neighbor H at rate β_2, all recovered nodes became anti-spreaders at the rate of $\delta = 1$.
t5: Anti-spreaders are beginning to take more portion of the network. However, there is still an infection going on in the network.
t6: No neighbors of the current infected nodes are susceptible.
t7: All reachable nodes in the network have seen the information. However, there is still a node G, which has never heard of the rumor or anti-rumor.
t8: All nodes that saw the rumor or the truth are now anti-spreaders. Other nodes like G are still susceptible. After the equilibrium point, all nodes will either end as susceptible, or anti-spreaders and no nodes are spreading the rumor if the parameters $\beta_1, \beta_2, \gamma, \delta$ are coherent to the reproduction number.

In above toy network simulation in Fig. 3, the value of various probabilities considered are $\beta_1 = 0.5, \beta_2 = 1, \gamma = 1, \delta = 1$, which are relatively very high values. The conversion rate from $S- > I$ is 0.5%, i.e., an infected node can randomly infect 50% of its susceptible neighbors. Similarly, each of the conversion rate from $I- > R$, $S- > R-$, and $R- > A$ is 1. It implies the probability of conversion from susceptible to recovered, infected to recovered and, recovered to anti-spreaders is 100%. At the end of the simulation, all the nodes are in the anti-spreader state except one node in the susceptible state. The node 'G' is non-reachable from all other nodes and hence, it is not influenced by any nodes during the simulation process and its color remains unchanged.

4 Result and analysis

In this section, we simulate our proposed methodology and analyze the different findings related to rumor propagation and rumor control using real-world various weighted and unweighted networks given in Subsection 4.1. Tab. 1 depicts the brief description used data-sets in this paper. We use the spreading characteristics defined by $F(t)$ giving the count of the susceptible, infected, anti-spreaders, and recovered nodes as a function of time. Infection scale, i.e.,

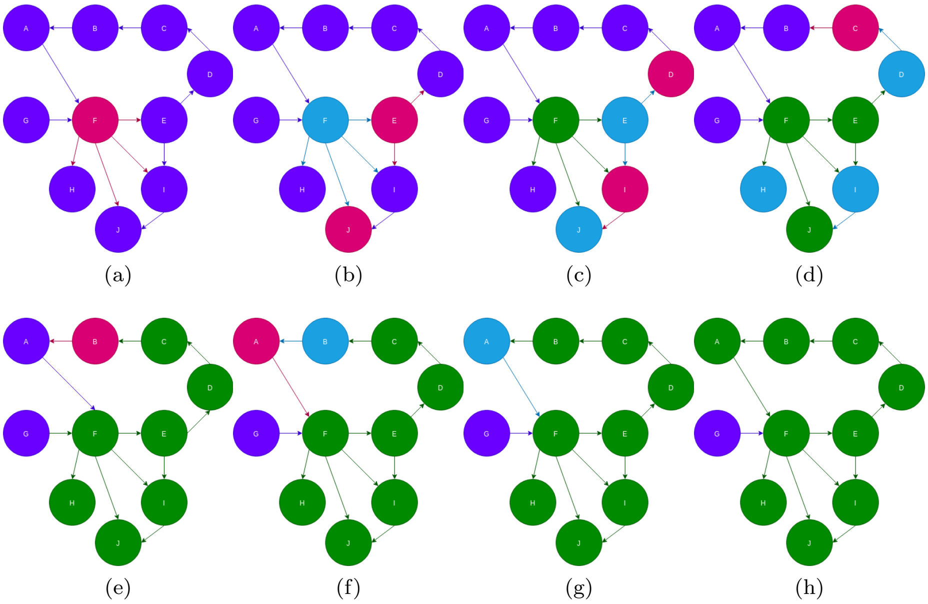


Fig. 3: All nodes are initially susceptible and an influential node is chosen as the spreader with $\beta_1 = 0.5, \beta_2 = 1, \gamma = 1, \delta = 1$. Here, purple nodes are susceptible, pink nodes are infected, blue nodes are recovered and green nodes are anti-spreaders.

$F(t)$, is an important evaluation parameter. When information propagates in a network then at any time t, $F(t)$ denotes the fraction of nodes in the network which have received the particular information. A higher value of $F(t)$ indicates that information spreads rapidly in the network. Section 4.2 discusses the propagation of non-rumors, similar to the traditional SEI information propagation model, using the proposed $SIRA$ model. Section 4.3 discovers the role of anti-spreaders in containing the harmful effects of the rumor. In section 4.4, a novel method for comparison of centralities is proposed. Here, we compare the recently introduced NCVoteRank [26] centrality algorithm for unweighted networks and WVoteRank [27] centrality on weighted networks for finding the initial spreaders. These centrality algorithms are simulated in comparison to traditional algorithms, DegreeRank [28] for unweighted networks, and its weighted version WDegreeRank for weighted networks.

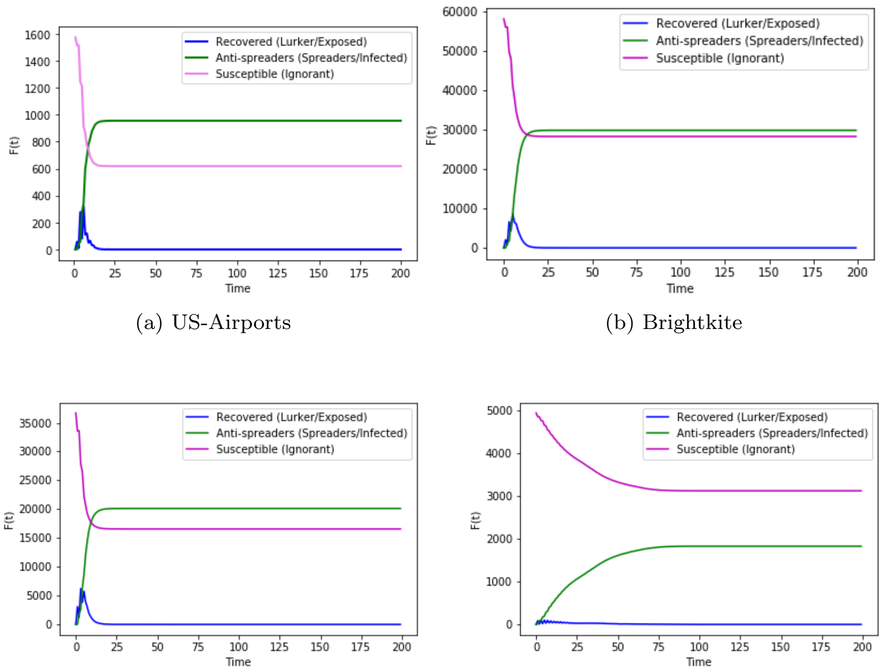
4.1 Datasets used

Dataset Name	Characteristics	Description
US Airports [20]	Weighted, Directed, 1,574 vertices, 28,236 edges	The network of the most crowded trading airports in the United States. An edge is there between airports, if there is a flight between them. The weights are the number of seats available on these flights.
Brightkite [21]	Unweighted, Undirected, 58228 nodes, 214078 edges	Brightkite was a location-based social network. Users can share their locations by check-ins.
Email-Enron [22]	Unweighted, Undirected, 36692 nodes, 183831 edges	Nodes of the network are email addresses. If address A sends one or more email to address B, the graph contains an undirected edge between these nodes.
US Powergrid [23]	Undirected, Weighted, 4941 nodes, 6,594 edges	This network is the high-voltage power grid in the USA. The nodes are substations, transformers, and generators, and the connections are voltage power lines.
CA-GrQc [24]	Unweighted, Undirected, 5242 nodes, 14496 edges	Arxiv GR-QC (General Relativity and Quantum Cosmology) research network is from the research papers and covers scientific referencing between papers in General Relativity and Quantum Cosmology category.
p2p-Gnutella08 [25]	Unweighted, Undirected, 6301 nodes, 20777 edges	This is a Gnutella peer-to-peer file-sharing network dataset collected from 08 August 2002 where a vertex represents host in the network, and the edge denotes the connection between the Gnutella hosts.

Table 1: Description of data-sets used

4.2 Information propagation on real-world networks

As highlighted in Section 3.5, the proposed algorithm reduces to a Susceptible-Exposed-Infected (SEI) model where susceptible nodes remain susceptible (S), recovered nodes act as exposed (E) nodes, and anti-spreader nodes are the carrier of the correct information and act as infected (I) nodes. The information diffuses based on two parameters, namely the rate of infection as β_2 and rate of recovery as δ, leading to a simplified model of information propagation. As evident, the model has similar characteristics to those of the SEI model with similar parameters and has expected behavior on the real-world weighted and unweighted networks, as shown in Fig. 4. We use NCVoteRank centrality [26] and to identify one top spreader as infected node on the unweighted networks like Brightkite and Email-Enron. Similarly, we utilize WVoteRank centrality [27] to identify one top spreader as infected node to start the information diffusion process on the weighted networks like US-Airports and Powergrid. Notice that all the graphs behave similarly to the SEI epidemic model. As evident from Fig. 4(a), the spreader on weighted network US-Airports is able to infect around 25% (750 nodes of the network receive the information at the



(a) US-Airports

(b) Brightkite

(c) Email-Enron

(d) Powergrid

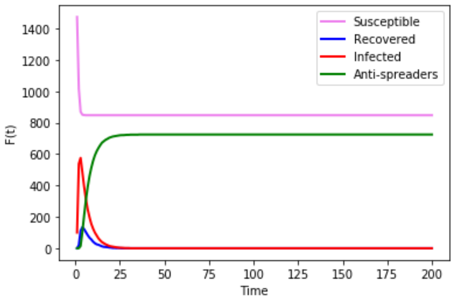
Fig. 4: Information propagation on real-world networks when there is no rumor, and anti-spreader nodes act as the information carrier. All nodes are initially susceptible. An influential node is chosen as the spreader using the NCVoteRank method in case of an unweighted network and WVoteRank method in weighted networks. The results are obtained over 100 independent simulation of SEI model with $\beta = 0.2$, and $\delta = 0.8$

end of the simulation, out of the total population of 3000) of the network. In Fig. 4(b), the spreader on Brightkite is able to infect around 50% of the network and ended up having a stable equilibrium. In Fig. 4(c), the spreader on Email-Enron is able to infect more than 50% of the network, notice that the peak reaches an early timestamp due to high connectivity. On the weighted PowerGrid network, notice that the change in the numbers is gradual, and very few nodes eventually become exposed as in Fig. 4(d).

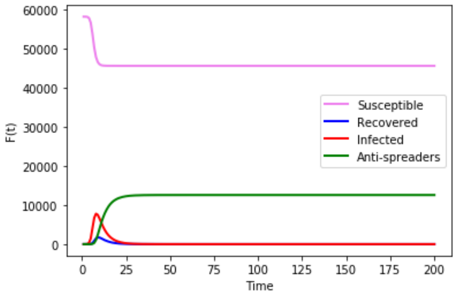
4.3 Numerical simulations for basic reproduction number and the equilibrium point

In this section, we discuss the equilibrium characteristics of the proposed method near the rumor-free equilibrium point, along with endemic equilibrium characteristics and the role of anti-spreaders nodes to minimize the effect of rumor on various network datasets. On reaching equilibrium on real-world networks, the rumor vanishes in the network only after the complete network has no more rumor-spreading nodes. The equilibrium is stable, only if the conditions are such that $R_t < 1$, as given by calculation in Eq. 10. We assume the values of the parameters β_1, β_2, and γ based on the reproduction number and network topology. The value of δ shall indicate how likely an informed (recovered) node is to dispel the rumor by becoming an anti-spreader. The value of δ depends on the preference of the given group and the severity of the information. As evident from Fig. 5(a), the spreaders on weighted network US-Airports can infect around 35% of the network in the beginning, whereas around 50% nodes of the network recovers and becomes anti-spreader during the simulation and minimizing the spread of the rumor. In Fig. 5(b), the spreader on Brightkite is able to infect around 15% of the network, and about 20% of the nodes became anti-spreader. Due to the influence of anti-spreaders, approximately 80% of the nodes remains susceptible means unaffected by the rumor. Hence, the roles of anti-spreaders nodes become vital to confine the rumor in the system. In Fig. 5(c), the spreader on Email-Enron is able to infect about 32% of the network initially, and 43% nodes are anti-spreaders. It is clear that due to the impact of ant-spreaders about 70% of nodes remains susceptible. On the weighted PowerGrid network, the spreader is able to infect about 15% of the network initially, and 22% nodes are anti-spreaders, which enables approximately 70% nodes are susceptible and are unaffected by the rumor as depicted in Fig. 5(d). Similarly, in the case of GrQc and p2p Guntella08 networks, as evident from Fig. 5(e) and 5f, respectively, the anti-spreader nodes are able to control the rumor, and most of the nodes remain susceptible in the system after reaching the equilibrium.

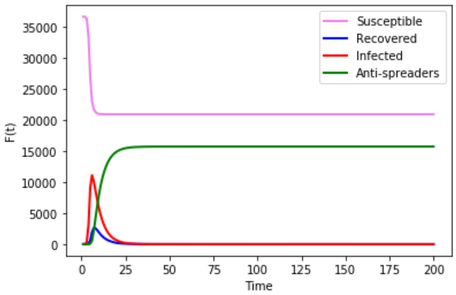
Here, in case of the results obtained on real-life network datasets in Fig. 5, results are obtained over 100 independent simulation of SIRA model with the value of various probabilities as $\beta_1 = 0.1, \beta_2 = 0.1, \gamma = 0.2, \delta = 0.9$. The probability of conversion from $S- > I$ is 0.1, i.e., an infected node can randomly infect 10% of its susceptible neighbors. The probability of conversion $S- > I$ is 0.1, i.e., an infected node can randomly infect 10% of its susceptible neighbors. The probability of conversion from $S- > R$ is 0.1, which implies due to the influence of anti- spreader nodes, their 10% susceptible neighbors go directly in recovered state. Similarly, the rate of conversion from infected state to recovered state is 20%, and the rate of conversion from recovered state to anti-spreader state is 90%. Hence, in this case, we have performed the simulation on relatively lower conversion rates as compared to the simulation performed in the toy network as in Fig. 3. This results in a significant number of nodes are in the susceptible state apart from in anti-spreader state when the



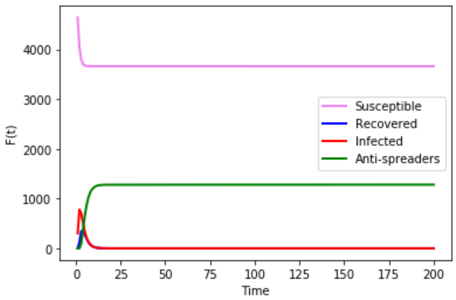
(a) US-Airports



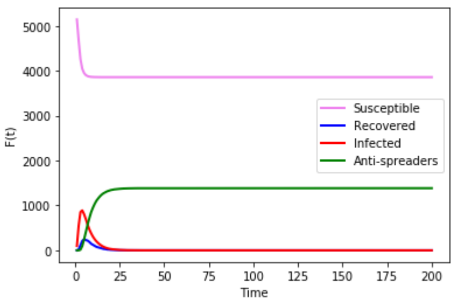
(b) Brightkite



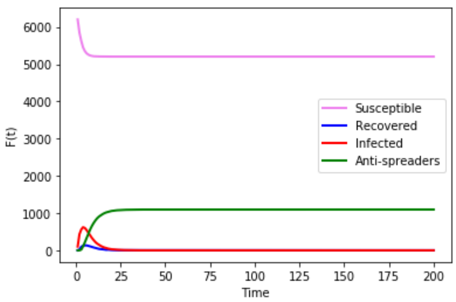
(c) Email-Enron



(d) US-Airports

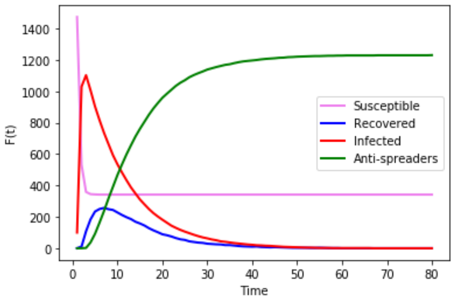


(e) Brightkite

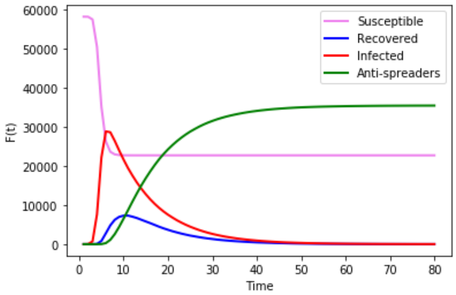


(f) Email-Enron

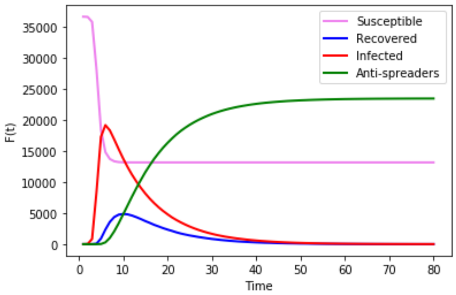
Fig. 5: The equilibrium characteristics of $SIRA$ model and impact of anti-spreaders to control the rumor as seen on various used datasets. The results are obtained over 100 independent simulation of the proposed model with $\beta_1 = 0.1, \beta_2 = 0.1, \gamma = 0.2,$ and $\delta = 0.9$



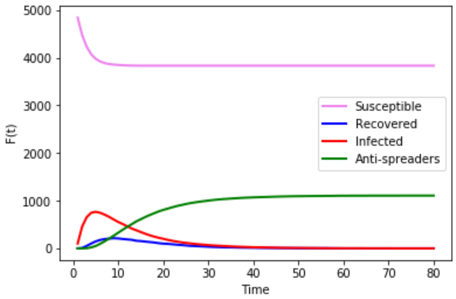
(a) US-Airports



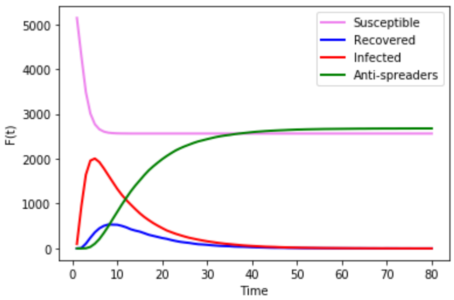
(b) Brightkite



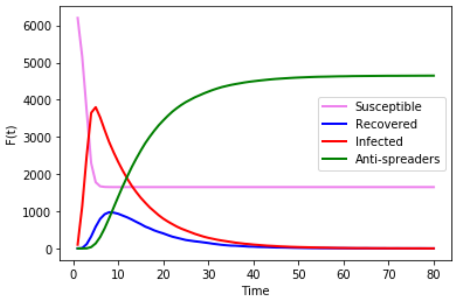
(c) Email-Enron



(d) US-Airports



(e) Brightkite



(f) Email-Enron

Fig. 6: The endemic equilibrium characteristics of $SIRA$ model observed on different used datasets. The results are obtained over 100 independent simulation of SIRA model with $\beta_1 = 0.4, \beta_2 = 0.01, \gamma = 0.1, \text{and } \delta = 0.3$

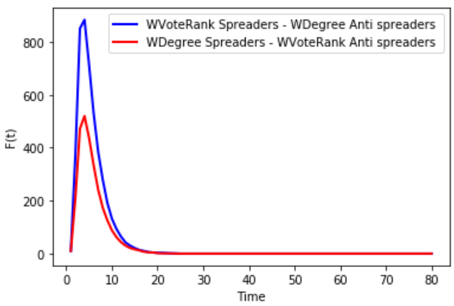
equilibrium is reached in the system. The increase in the value of the conversion rate, especially β_1 for $S->I$ and β_2 for $S->R$, results in less number of susceptible nodes at the end of the simulation when equilibrium is reached. Also, many real-life networks like US-Airports, Brightkite, and social networks are sparse networks where many nodes exist with relatively less degree. Here, degree means the number of direct neighbors or number of friends in social networks, and some of such nodes may not come in contact with either infected or anti-spreader nodes in the system and remains in the suseptible state or unware of the information.

The endemic equilibrium state is the state where the disease or rumor cannot be totally eradicated but remains in the population [35]. For the disease or rumor to persist in the population, the anti-spreader class, the susceptible class, the infectious class, and the recovered class must not be zero at the equilibrium state, i.e., $(S, I, R, A) \neq (0, 0, 0, 0)$. In this case the equilibrium is stable, only if the conditions are such that $R_t > 1$, as given by calculation in Eq. No. 10. The results obtained in Fig. 6 on the various used datasets confirms that there exits an endemic equilibrium point and its stability in the system.

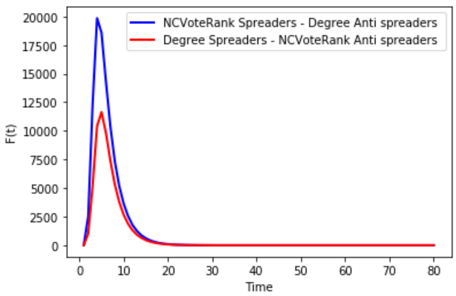
4.4 Spreader and Anti-Spreader seeding: Comparing Centrality algorithms

The proposed $SIRA$ model can compare two different node-centrality algorithms for evaluating the spread and control of information to model the real-life situation. In the framework, we can change the comparing centrality algorithms. Node centrality methods intend to assign some score to each node of the network based on its significance and different parameters [29, 30] In this case, spreaders are chosen from one centrality algorithm spread the rumor, and the anti spreaders are chosen from the second algorithm dispel the rumor and vice-versa. We apply $F(t)$, the infected scale, as a function of time, as the performance measure for this comparison. We employ the top 10 selected spreaders and top 10 selected anti spreaders using the specified centrality algorithm for performing the comparison study. We intend to develop a generic framework where any two node centrality can be taken to compare the performance. In the case of unweighted networks, we consider NCVoteRank and Degree centrality as the choice of node centrality algorithms. However, in the case of the weighted networks, WVoteRank and WDegree centrality is viewed as the choice of node centrality algorithms. During the simulation, due to a non zero value of δ, the count of anti-spreaders increases, and also newly became anti-spreaders can immune their neighbors as per the dynamics of the model.

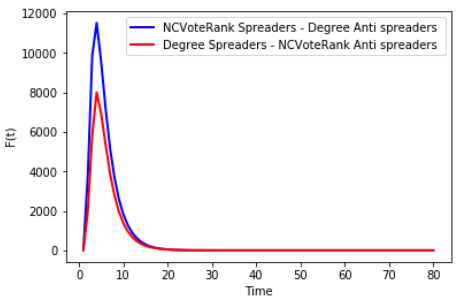
The obtained results of comparing centrality is depicted in Fig. 7 for various considered datasets by using the parameters $\beta_1 = 0.4, \beta_2 = 0.3, \gamma = 0.3, \delta = 0.5$. From the results, it can be inferred that initially, the count of infected nodes increases due to the influence of selected seed nodes. However,



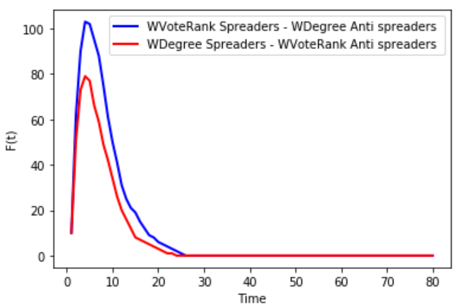
(a) US-Airports



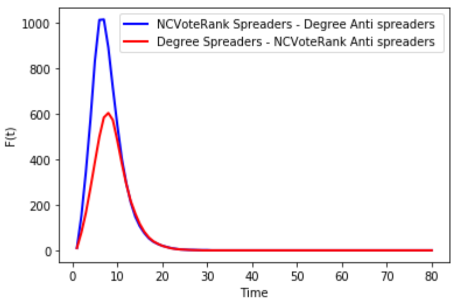
(b) Brightkite



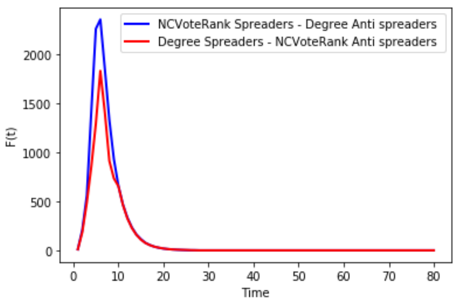
(c) Email-Enron



(d) US-Airports



(e) Brightkite



(f) Email-Enron

Fig. 7: The spreading of the information using comparing centrality algorithms on the various network datasets with NCVoteRank and DegreeRank used on unweighted networks, and WVoteRank and WDegreeRank on weighted networks with parameters values as $\beta_1 = 0.4, \beta_2 = 0.3, \gamma = 0.3, \delta = 0.5$

shortly the initial seed and the converted Anti-Spreader nodes tend to slow down the conversion of susceptible nodes to the infected node and rapidly decrease the count of spreaders by converting them to recovered nodes based on the parameters considered. PowerGrid network is a weighted network, and we use WVoteRank centrality and WDegreeRank centrality for the selection of spreaders and anti-spreaders respectively and vice-versa. WVoteRank as the choice for spreaders and WDegreeRank as the choice for anti-spreaders can get 350% more infected users than WDegreeRank as the method for selecting spreaders and WVoteRank for the anti-spreaders, as evident in Fig. 7(d). CA-GrQc is an unweighted network, and hence, NCVoterank and Degreerank are applicable for comparison. NCVoteRank as the choice for spreaders and DegreeRank as the choice for anti-spreaders is able to get 80% more infected users than DegreeRank as the method for the selection of spreaders and NCVoteRank for the anti-spreaders, as evident in Fig. 7(e). Hence, we can conclude that NCVoteRank, on unweighted networks, and WVoteRank, on weighted networks, act as better node selection algorithms at both the spreading, and the control of the rumor and fake news as compared to degree on the unweighted and weighted degree on the weighted networks.

5 Conclusion

The healthcare social networks present potential risks for patients due to the possible distribution of poor-quality or wrong information. In this paper, we proposed the $SIRA$ rumor propagation model to investigate the process of rumor propagation and control in healthcare social networks. We added a new state to the model called the anti-spreader, that spread the truth, to prevent the harmful effects of the rumor. Considering the variety in people's reactions to rumors, we divided nodes into four categories, namely susceptible, anti-spreaders, recovered, and infected. We modeled the system of differential equations mathematically and determined the reproduction number to find the equilibrium point and also confirmed the consistency of the proposed model on real-world data-sets. We generalized our model into an information propagation model similar to the traditional models for modeling information diffusion to achieve influence maximization. In both weighted and unweighted networks, we compared two different centrality algorithms, one for choosing the spreaders and the other for choosing the anti-spreaders and vice-versa and the amount of information spread. The studies showed that our model is coherent with the spread of rumors in social network topology, and our proposal gives a great reference to simulating and controlling the spread of rumors and fake news on the network. The proposed model is not applicable for heterogeneous networks. In the near future, we can extend our model to heterogeneous networks.

Declrations

Conflict of interest: The authors declare that they have no conflict of interest.
Funding: No funding has been received.
Conflicts of interest/Competing interests: The authors certify that there is no conflict of interest in the subject matter discussed in the manuscript.
Ethics approval: The work conducted is not plagiarized. No one has been harmed in this work.
Consent to participate: All the authors have given consent to submit the manuscript.
Consent for publication: Authors provide their consent for the publication.
Author's contributions: All the authors have contributed equally in the research, compiling results and manuscript preparation.

References

1. O'Keeffe GS, Clarke-Pearson K (2011) The Impact of Social Media on Children, Adolescents, and Families. PEDIATRICS, 127(4) 800–804.
2. Kijkuit B, Ende J (2007) The Organizational Life of an Idea: Integrating Social Network, Creativity, and Decision-Making Perspectives. Journal of Management Studies, 44(6) 863–882.
3. Qiu L, Jia W, Niu W, Zhang M, Liu S (2021) SIR-IM: SIR rumor spreading model with influence mechanism in social networks. Soft Computing, 25(22), pp.13949-13958.
4. D. Spohr, Fake news and ideological polarization. Business Information Review, 34(3) (2017) 150–160.
5. Sivasankari S, Vadivu G (2021) Tracing the fake news propagation path using social network analysis. Soft Computing, pp.1-9.
6. Pfeffer J, Zorbach T, Carley KM (2013) Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. Journal of Marketing Communications, 20(1-2).
7. Lai K, Xiong X, Jiang X, Sun M, He L (2020) Who falls for rumor? Influence of personality traits on false rumor belief. Personality and Individual Differences, 152, p.109520.
8. Harko T, Lobo FSN, Mak MK (2014) Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and the SIR model with equal death and birth rates. Applied Mathematics and Computation, 236 (2014) 184–194.
9. Britton T, Stochastic epidemic models: A survey. Mathematical Biosciences, 225(1) (2010) 24–35.
10. S. Kumar, M. Saini, M Goel, N. Aggarwal, Modeling Information Diffusion In Online Social Networks Using SEI Epidemic Model. Procedia Computer Science, 171 (2020) 672-678.
11. Tripathy RM, Bagchi A, Mehta S (2010) A study of rumor control strategies on social networks, in Proceedings of the 19th ACM International Conference on Information and Knowledge Management - CIKM '10. Presented at the 19th ACM international conference, 1817-1820.
12. Huang J., Jin X (2011) Preventing rumor spreading on small-world networks. Journal of Systems Science and Complexity, 24(3) pp.449-456.
13. Wan YP, Zhang DG, Ren QH (2015) Propagation and inhibition of online rumor with considering the rumor elimination process. Acta Physica Sinica, 64(24) p.240501.
14. Zhu H, Ma J (2018) Analysis of SHIR rumor propagation in random heterogeneous networks with dynamic friendships. Physica A: Statistical Mechanics and Its Applications, 513, 257-271.
15. Indu V, Thampi S M (2019) The nature-inspired approach based on the Forest Fire model for modeling rumor propagation in social networks. Journal of Network and Computer Applications, 125, pp. 28-41.

16. Wang C, Wang G, Luo X, Li H (2019) Modeling rumor propagation and mitigation across multiple social networks. Physica A: Statistical Mechanics and its Applications, 535, 122240.
17. Zhu L, Liu W, Zhang Z (2020) Delay differential equations modeling of rumor propagation in both homogeneous and heterogeneous networks with a forced silence function. Applied Mathematics and Computation, 370, p.124925.
18. Yang A, Huang X, Cai X, Zhu X, Lu L (2019) ILSR rumor spreading model with degree in a complex network. Physica A: Statistical Mechanics and its Applications, 531, 121807.
19. Kaur S, Kumar P, Kumaraguru P (2020) Automating fake news detection system using multi-level voting model. Soft Computing, 24(12), pp.9049-9069.
20. Colizza V, Pastor-Satorras R, Vespignani A (2007) Reaction-diffusion processes and metapopulation models in heterogeneous networks. Nature Physics 3, 276-282.
21. Cho E, Myers SA, Leskovec J (2011) Friendship and Mobility: Friendship and Mobility: User Movement in Location-Based Social Networks ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)
22. Leskovec J, Lang K, Dasgupta A, Mahoney M(2009) Community Structure in Large Networks: Natural Cluster Sizes and the Absence of Large Well-Defined Clusters. Internet Mathematics 6(1) 29–123,
23. Watts DJ, Strogatz S H (1998) Collective dynamics of "small-world" networks. Nature, 393, 440-442.
24. Leskovec J, Kleinberg J, Faloutsos C (2007) Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data (ACM TKDD), 1(1)
25. Ripeanu M, Foster I, Iamnitchi A (2002) Mapping the Gnutella Network: Properties of Large-Scale Peer-to-Peer Systems and Implications for System Design. arXiv preprint cs/0209028
26. Kumar S, Panda BS (2020) Identifying influential nodes in Social Networks: Neighborhood Coreness based voting approach. Physica A: Statistical Mechanics and Its Applications, 124215.
27. Kumar S, Panda A, (2022) Identifying influential nodes in weighted complex networks using an improved WVoteRank approach. Applied Intelligence, 52, 1838–1852
28. Freeman L C, Centrality in social networks conceptual clarification. Social Networks, 1(3)(1978) 215–239.
29. Kumar S, Gupta A, Khatri I (2022) CSR: A community based spreaders ranking algorithm for influence maximization in social networks. World Wide Web, pp.1-20.
30. Kumar S, Mallik A, Khetarpal A, Panda B S (2022) Influence maximization in social networks using graph embedding and graph neural network. Information Sciences.
31. Roussel MR (2005) Stability analysis for ODEs. Nonlinear Dynamics, lecture notes, University Hall, Canada
32. Izhikevich EM (2007) Equilibrium. Scholarpedia, 2(10):2014.
33. Lyapunov AM (1992) The general problem of the stability of motion. International journal of control, 55(3) pp.531-534.
34. L. Guihua, J. Zhen Global stability of an SEI epidemic model. Chaos, Solitons & Fractals, 21(4) (2004) pp.925-931.
35. Ugwa KA, Agwu IA, Akuagwu NA (2013) Mathematical analysis of the endemic equilibrium of the transmission dynamics of tuberculosis. International Journal of Technology Enhancements and Emerging Engineering Research, 2(12) pp.263-9.