A Comparative Evaluation of the SIR and SEIZ Epidemiological Models to Describe the Diffusion Characteristics of COVID-19 Polarizing Viewpoints on Online Social Networks

Maryam Maleki
Assistant Professor, Systems Engineering Program,
California State University, Dominguez Hills, USA
mmaleki@csudh.edu

Nitin Agarwal
COSMOS Research Center, UA - Little Rock, Arkansas, USA
International Computer Science Institute,
University of California, Berkeley, USA
nxagarwal@ualr.edu

Abstract

To understand and characterize the diffusion trends of opposing viewpoints on Twitter (X), we applied two epidemiological models to six datasets related to COVID-19. We compared the results of the SIR (Susceptible, Infected, Recovered) and the SEIZ (Susceptible, Exposed, Infected, Skeptics) epidemiological models. We collected six datasets indicative of polarizing viewpoints related to contentious subjects surrounding the COVID-19 pandemic. Three of the datasets fall into an anti-subject hashtag group, and three of the datasets fall into a pro-subject hashtag group. The timeframe of each dataset is from January 1, 2020, to the end of 2021. Our findings demonstrate that while both the SIR and the SEIZ models are able to evaluate the propagation trends of the polarizing viewpoints, the SEIZ model is more accurate with relatively less error compared to the SIR model. This work sets the stage for ultimately leading to the ability to develop methods to prevent the propagation of ideas that lack scientific evidence while promoting the spread of scientifically backed ideas.

Keywords: Mathematical modeling, Epidemiological model, SEIZ model, SIR model, COVID-19, Twitter (X), Social network analysis, Social contagion.

1. Introduction

Social media platforms such as Twitter (X) have become one of the most popular sources of information for people around the world. Twitter (X) facilitates the propagation of polarizing viewpoints.

Hashtags on Twitter (X) are utilized to categorize messages, amplify ideas, and promote specific topics and individuals (Cunha et al., 2011)

While some of these viewpoints are backed by scientific evidence, other viewpoints may not be as much. Non-scientifically backed viewpoints could tend to feed misinformation. Consequently, due to its popularity along with its accessibility, ease of use, and focus on user-generated posting and interaction, Twitter (X) is faced with the risk of the propagation of misinformation (Jain, Sharma, & Kaushal, 2016). Twitter (X) has also become a platform wherein users argue opposing viewpoints.

Although being exposed to public discourse that contains varying viewpoints can be healthy in terms of gaining reciprocal understanding and situational awareness, these arguments often contain misinformation. Spreading misinformation and confusion within public arguments about healthcare subjects can be dangerous and poses a serious threat to people's health (Van Bavel et al., 2020). Unfortunately, the social media information regarding the novel coronavirus of 2019 (SARS-CoV-2) and its resultant disease (COVID-19) that caused a global pandemic is a profound example.

Numerous items of false information, hoaxes, and scams about COVID-19 topics have been propagated on multiple online social media platforms, such as its causes, consequences, spreading, prevention, medicine, treatment, and vaccination (Tasnim, Hossain, & Mazumder, 2020).

The prolific nature and preponderance of this COVID-19 misinformation is such that it is increasingly referred to as an "infodemic" ("Managing the COVID-19 infodemic: Promoting healthy... - Google Scholar," n.d.). This infodemic has given rise to a vast array of public discussions of polarizing viewpoints on COVID-19-related subjects such as the effectiveness of lockdowns, masks, and vaccines.

Tasnim et al. (2020) have shown that misinformation can spread faster and further than legitimate information on social media, which harms the balance of the information ecosystem. We expect

URI: https://hdl.handle.net/10125/109142
978-0-9981331-8-8

that the propagation of polarizing viewpoints follows a similar trend, especially when they contain items of misinformation.

Li et al. (2020) showed that over 25% of the most-viewed COVID-19-related videos on YouTube included misinformation, which represented a total of 62 million views worldwide (Li, Bailey, Huynh, & Chan, 2020).

British communications services regulatory agency, Ofcom, conducted a study that revealed that in just one week, almost half of online adults in the United Kingdom were exposed to COVID-19 misinformation. Two-thirds of those exposed reported that they see misinformation on a daily basis. This is a significant concern because research in experimental psychology has shown that repeated exposure to false information leads to an increase in its perceived accuracy (Lazer et al., 2018).

This study is motivated by the influence of the propagation of polarizing viewpoints on social media in people's minds, specifically relative to public health issues.

We attempted to apply a mathematical model to explain how different viewpoints related to COVID-19 are propagated across Twitter (X).

We studied the trend of spreading different anti-hashtags and pro hashtags for three different categories including lockdown, face mask, and vaccine. We collected six different datasets containing anti and pro hashtags from January 1, 2020, until to the end of 2021, which was the peak time of this pandemic and in the spread of relative misinformation.

The primary objectives in this work include:

- Can we find epidemiological mathematical models that are able to evaluate the propagation trends of polarizing viewpoints related to COVID-19 on Twitter (X)?
- Are there any similarities or differences between the propagation characteristics of anti- and pro-viewpoints relative to a given subject disseminated via hashtags?
- Which model is more capable of modeling the diffusion trends of opposing viewpoints on an online social network?

To answer these questions, we were motivated to apply two different epidemiological models to determine, compare, and contrast the diffusion trends of opposing viewpoints related to COVID-19 on Twitter (X).

In this study, we have applied the SIR (Susceptible, Infected, Recovered) and the SEIZ (Susceptible, Exposed, Infected, Skeptic) models. Theoretically, the SEIZ model is a stronger model for the propagation of information compared to the other epidemiological models (Bettencourt, Cintrón-Arias, Kaiser, & Castillo-Chávez, 2006). In this study, we aim to see if this model can model different ideas more accurately than the SIR model.

The primary contributions of this work include:

- We demonstrated that epidemiological models such as SIR and SEIZ can be used to explain the diffusion characteristics of polarizing topics.
- While both pro and anti-viewpoints can be modeled by SIR and SEIZ, both models did better in explaining the anti-viewpoint diffusion characteristics.
- Between SIR and SEIZ, SEIZ was the clear winner, meaning SEIZ outperformed SIR for all the topics that were considered and for both the viewpoints (anti and pro).

The remainder of this paper is presented as follows.

Section 2 presents the related work that is germane to the application of epidemiological models to the spread of rumor, information, and misinformation. In section 3, the methodology used in this work is described, including our data collection process, data processing, and the models applied. This discussion includes a comparison between the SIR and SEIZ epidemiological models. Section 4 discusses our analysis and results and highlights the breakout findings from our research. Finally, section 5 discusses our overall conclusions and the impact of our research. We also discussed our ideas for future work.

2. Related work

In this section, we discuss the related works that have applied epidemiological models to the spread of various types of information on Online Social Networks (OSNs). We first give a brief explanation of the common epidemiological models. We then discuss the application of these models to the spread of information, rumor, and misinformation. We conclude the section with a brief discussion of some of the adaptations to the fundamental epidemiological models that have been proposed over time.

According to Bettencourt et al. (Bettencourt et al., 2006), quantitative measures derived from the modeling of epidemics and the spread of infections "generalize naturally to the spread of ideas and provide a simple means of quantifying sociological and behavioral patterns."

When applying epidemiological (epidemic) models to information on OSNs, the spread of the information is analogous to the spread of an infection or disease within a community or population.

Epidemiological models all begin with a basic framework, which involves dividing the population into different compartments.

The most basic epidemic model is the SI (Susceptible-Infected) model, which partitions the total population into the compartments of Susceptible and Infected. In this model, the Infected compartment involves individuals who are already carrying the infection (i.e., spreading the information), while the Susceptible compartment consists of people who are at risk of contracting the infection (i.e., being exposed to the information) (Kabir, Kuga, & Tanimoto, 2019).

The SIS model (Susceptible-Infected-Susceptible), extends the SI model to include the realization that people who are infected may be transferred to the Susceptible compartment again at some point (i.e., no longer sharing the information but returning to being exposed to it) (reports & 2016, n.d.).

The SIR model (Susceptible-Infected-Recovered) accounts for instances wherein users recover from infection and develop immunity (i.e., no longer share nor are susceptible to the information) (Colizza & Vespignani, 2007) (Abdullah & Wu, 2011).

The present work focuses on a novel approach to characterize the propagation of conspiracy theories through social networks by applying epidemiological models to Twitter data. A Twitter dataset was searched for tweets containing hashtags indicating belief in the "5GCoronavirus" conspiracy theory, which states that the COVID-19 pandemic is a result of, or enhanced by, the enrollment of the 5G mobile network.

Despite the absence of any scientific evidence, the "5GCoronavirus" conspiracy theory propagated rapidly through Twitter, beginning at the end of January, followed by a peak at the beginning of April, and ceasing/disappearing approximately at the end of June 2020. An epidemic SIR (Susceptible-Infected-Removed) model was fitted to this time series with an acceptable model fit, indicating parallels between the propagation of conspiracy theories in social networks and infectious diseases. Extended SIR models were used to simulate the effects that two specific countermeasures, fact-checking, and tweet-deletion, could have had on the propagation of the conspiracy theory.

The SEIZ model (Susceptible-Exposed-Infected-Skeptic) includes a Skeptic (Z) compartment consisting of users who know about the content but have made a decision not to continue to spread it, and an Exposed (E) compartment consisting of users who know about the content but have not yet decided whether or not to share it (Jin, Dougherty, Saraf, Cao, & Ramakrishnan, 2013).

The SI, SIR, and SEIZ models have been applied successfully to study the spread of information on web forums (Woo & Chen, 2016) as well as the spread of news and concepts on Twitter (Abdullah & Wu, 2011), (Shu, Wang, Lee, & Liu, 2020), (Works-Sprogvidenskabeligt & 2020, 2020) providing evidence to support the generalizability claim of Bettencourt et al. (i.e., the similarity between the propagation of disease and the spread of information on Twitter).

The comparison of the spread of rumors to that of epidemics goes back as far as 1964 with the works of Daley and Kendall (Daley, Kendall, DJ, & DG, 1964) (DALEY & KENDALL, 1965).

Both the SIR and SEIZ models have been applied to study the spread of rumors on online platforms such as LiveJournal (Zhao et al., 2011) and Twitter (Shu et al., 2020). Isea et al. (2017) used a version of the SEIZ model to evaluate the spread of a rumor for a specific case in Venezuela. In their model, the rumor propagates between two various scenarios Z_1 and Z_2 that do not share information with each other (Isea & Lonngren, 2017).

Epidemiological models have also been successfully applied to the study of misinformation on OSNs.

Cinelli et al. (Cinelli et al., 2020) applied the SIR model to data from Twitter, Instagram, YouTube, Reddit, and Gab to examine the diffusion of a mix of COVID-19 "reliable and questionable information sources". The authors were able to characterize the basic reproduction number R0 for each platform, stating that an R0 > 1 "indicates the emergence of an infodemic". Further, although they found evidence that each platform reacts differently to information produced by reliable and questionable news outlets and that Gab was more susceptible to the spread of misinformation, their results did not show any significant differences in the spreading behaviors between reliable and questionable information sources.

Kauk et al. (2021) used different extended SIR models to evaluate the propagation of tweets containing hashtags related to the '5GCoronavirus' conspiracy theory (Kauk, Kreysa, & Schweinberger, 2021).

Maleki et al. (Maleki, Mead, Arani, & Agarwal, 2021) applied the SEIZ model to study the spread of misinformation on Twitter. Analyzing the spread of an item of misinformation that falsely claimed the existence of a riot in Washington, D.C. during the Black Lives Matter movement in March 2020, the researchers were able to provide evidence for the ability of the SEIZ model to describe the spread of a specific misinformation item on an OSN.

In another study, Maleki et al. (Maleki, Mead, et al., 2021) applied the SEIZ model to evaluate the trend of the propagation of misinformation and legitimate hashtags from three different Twitter COVID-19 campaigns: lockdown, face mask, and vaccine. Their findings provided evidence for the ability of the SEIZ model to model the propagation of both misinformation and legitimate information campaigns on an OSN.

Over time, several adaptations of the fundamental epidemiological models have been proposed for application to the spread of various types of information on OSNs.

Jin et al. (2013) also used an epidemiological model to evaluate the propagation of news and rumors on Twitter. In this more expansive effort, the authors applied the SEIZ model to evaluate the diffusion trend of four news items and four rumors on Twitter. The SEIZ model adds a Skeptic (Z) compartment, which consists of users who know about the story but decide not to spread it, and an Exposed (E) compartment, which are users who know about the news but need some time to decide to spread it (Jin et al., 2013)

Rodrigues & Fonseca (2016) used the SIR (Susceptible, Infected, Recovered) epidemiological model to analyze the effects of the marketing communication strategy of viral marketing. Through simulations using the infectivity rate, the recovery rate, and the cost to send the message, the authors found that, as expected, when the number of initially infected people (size of the seed group to which the message was sent) was high, message virality peaked more rapidly. Additionally, however, the cost of increasing the seed group from 10% to 20% of the target population did not produce any significant benefits in terms of time and population reached. (Rodrigues & Fonseca, 2016)

Kereselidze (2019) considered the spread of false information as a form of "information warfare" and examined it using an expanded RAI (Risk, Adept, Immunity) model. The author states that the RAI groups correspond to the SIR groups of Susceptible, Infected, and Recovered. The authors emphasize the importance of inserting what they call "objective information" (N0) into the at-risk group (RG), which acts like a vaccine to reduce the spread of false information. (Kereselidze, 2019).

Rystrom investigated the use of the classical SI model and the sociology-inspired SEIZ model to understand the spread of concepts on Twitter. The authors use the term "concept" to refer to a word or a small set of words that describe the same phenomena. Concepts could be feelings such as "happiness" that can be described by a set of words such as happiness, joy, and well-being. The paper analyzes discourse around Rasmus Paludan, a far-right politician (Works-Sprogvidenskabeligt & 2020).

Castiello et al. (2023) employed the ISR model to analyze how information propagates on Twitter through hashtags related to various events, including 'DeathofQueenElizabethII', and 'TaylorSwiftsMidnights,' among others (Castiello, Conte, & Iscaro, 2023)

Escalante and Odehnal (2020) used a SIRS (Susceptible, Infected, Recovered, Susceptible) model in an attempt to understand the propagation of a rumor and how the insertion of a second deliberately counter rumor into a network will affect that propagation. In this work, the authors consider the first rumor as the infection, and the second rumor as an attempt at a vaccine (Escalante & Odehnal, 2020).

A SCIR (Susceptible, Contacted, Infected, and Refractory) has been presented to consider users who have attained immunity by way of a simple loss of interest in the information (Xiong, Liu, Zhang, Zhu, & Zhang, 2012).

Building upon their previous work, Zhao et al. (Zhao et al., 2012) proposed an SIHR (Susceptible-Infected-Hibernator-Removed) for rumor spreading to account for a new type of user compartment called Hibernators, which are spreaders who first experience the forgetting mechanism for some time, and then experience the remembering mechanism, which leads them to spread the rumor again. And, although there are numerous other examples, a final one is a theoretical STRS (Susceptible-Toxic-Recovered-Susceptible) proposed to study the spread of toxic content on YouTube (Obadimu, Mead, Maleki, & Agarwal, 2020).

Maleki et al. (2023) investigated how bots influence the effectiveness of the SIR and SEIZ epidemiological models in modeling the propagation of COVID-19-related ideas on Twitter (Maleki, Arani, Mead, & Agarwal, 2023).

To the best of our knowledge, there has been no prior study which has applied the SIR and the SEIZ epidemiological models to the propagation of opposing viewpoints related to COVID-19. Applying epidemiological models to the propagation of opposing viewpoints can enable researchers to evaluate the spread trend of non-scientifically backed viewpoints and find the best ways to control them.

3. Methodology

Methods of data collection are initially described, followed by an explanation of the epidemiological models applied in this work.

To provide the proper background, we compare the structure and components of two epidemiological models containing the SIR and the SEIZ models.

3.1. Data Collection and Processing

We used the Twitter academic API to collect tweets related to COVID-19 from January 1st, 2020, to the end of 2021. We collected data for different sets of hashtags that could best cover a broad range of topics related to COVID-19. For every category, we collected a broad range of hashtags to be able to include as much data as possible for each category. These hashtags are categorized in two categories: anti-hashtags and pro-hashtags and every category includes three subjects including mask, lockdown, and vaccine. A sample of hashtags collected for each Anti and Pro category subject as well as the number of tweets is shown in Table 1.

Table 1. Summary information for the datasets used in this work.

Dataset	Number of tweets	Sample of hashtags
Anti-Mask	751,302	#MasksareDangerous, #MasksDontWork, #Maskskill,
Pro-Mask	7,497,027	#WearAmaskSaveALife #masksaveslives, #WearADamnMask,
Anti-Lockdown	864,425	#nolockdowns, #saynotolockdown, #lockdownskill,
Pro-Lockdown	9,926,832	#respectlockdown, #lockdownsaveslives, #stayhomekeepsafe,
Anti-Vaccine	774,323	#sayNotoCovidVaccine, #covidvaccinekill, #NocovidVaccineForMe
Pro-Vaccine	816,290	#GetVaccinatedNow, #GetThecovidShot, #vaccinesaveslives,

3.2 Epidemiological Models

To evaluate the spread of different sets of opposing viewpoints related to COVID-19 on Twitter (X) based on indicative hashtags, we utilized the results of the two most common epidemiological models for evaluating the propagation of information, the SIR and the SEIZ.

3.2.1 SIR Model. The SIR model is a common fundamental epidemiological model. This model divides members of a population into three groups: Susceptible (S), Infected (I), and Recovered (R) (Figure 1).

Infected people consist of people who have the disease and are capable of spreading it to others. Susceptible people are individuals who are at risk of getting infected and the Recovered people are those who are immune from the infection or have died from the infection; consequently, they cannot cause another individual to be infected (Abdullah & Wu, 2011).

By allocating new definitions to these terms, the SIR model can be adjusted to consider the spread of information on Twitter (X), treating "information" as the "infection", and Twitter (X) users as the population.

In this revised SIR, to adjust this model to the spread of information on Twitter (X), we allocated new definitions to these terms (Maleki, Arani, Buchholz, Mead, & Agarwal, 2021).

A user is Susceptible if they follow someone who has posted a specific hashtag but has not yet posted a specific hashtag themselves. A user is Infected if they post a tweet using the specific hashtag and they are Recovered if they have not subsequently posted tweets with the specific hashtag within a certain time frame.

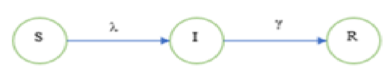

Figure 1. SIR model.

The following system of Ordinary Differential Equations (ODE) represents the SIR model (Abdullah & Wu, 2011).

$$\frac{dS}{dt} = -\lambda S$$

$$\frac{dI}{dt} = \lambda S - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

3.2.2 SEIZ Model. One important limitation of the SIR mode is that when a Susceptible individual is contacted by an Infected one, there is just one possible action, which is that the individual can only become Infected. However, this assumption does not apply properly to the propagation of information on Twitter (X).

Social media users may have different mindsets when they are exposed to information, especially information that is about polarizing subjects such as COVID-19.

Now a global issue, users have opposing viewpoints about different aspects related to COVID-19, such as agreement or disagreement with the use of lockdowns, masks, and vaccines.

Some users may need some time to decide if they believe the information and to decide if they want to spread the hashtags related to them. In addition, some users can be skeptical of the correctness of the information to which they are exposed and never show any reaction to tweets with these specific hashtags.

These possibilities can happen in reality but are not covered by the basic epidemiological models, such **as** SIR. In the context of analyzing the propagation of ideas on Twitter (X), the different compartments of the SEIZ model (Figure 2) are outlined below (Jin et al., 2013).

- Infected (I) relates to users who have used the specific hashtag in their tweets.
- Susceptible (S) represents users who follow the Infected individuals.
- Exposed (E) represents the users who have been exposed to the tweets containing a specific hashtag and had a delay of time before posting a tweet using the specific hashtag.
- Skeptic (Z) refers to individuals who have encountered the information via a tweet but decide not to tweet and use the hashtag (Shu et al., 2020).

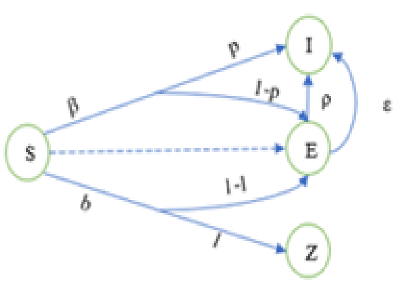

Figure 2. SEIZ model.

The following system of Ordinary Differential Equations (ODE) represents the SEIZ model (Jin et al., 2013).

$$\frac{dS}{dt} = -\beta S\frac{I}{N} - bS\frac{Z}{N}$$

$$\frac{dE}{dt} = (1-p)\,\beta S\frac{I}{N} + (1-l)bS\frac{Z}{N} - \rho E\frac{I}{N} - \varepsilon E$$

$$\frac{dI}{dt} = p\beta S\frac{I}{N} + \rho E\frac{I}{N} + \varepsilon E$$

$$\frac{dZ}{dt} = lbS\frac{Z}{N}$$

For the above-mentioned ODEs, the parameters are defined in Table 2. (Jin et al., 2013)

Table 2. Parameters of the SEIZ model

Parameter	Definition
β	Contact rate between S and I.
b	Contact rate between S and Z.
ρ	Contact rate between E and I.
p	Probability of S to I given contact with I.
1-p	Probability of S to E given contact with I.
ε	Transition rate of E to I (Incubation rate).
l	Probability of S to Z given contact with Z.
1-l	Probability of S to E given contact with Z.

When a Susceptible (S) - The follower of an Infected (I) user - contact with the Infected person (I) with β rate, they can immediately decide to share the hashtag with p probability, or that user may need some time to think about it and move to the Exposed (E) state with (1-p) probability.

In addition, a Susceptible may be in contact with a Skeptic (Z) - user who saw the hashtag but decided not to tweet about it - with the rate b. This contact can lead to two different scenarios.

The first possibility is that it can lead turning the user into another skeptic with the probability of l. It means the user chose not to tweet about it or not use the hashtag maybe they do not believe in the hashtag or believe it but decide not to do any reaction. The second one is that it may result in the unintentional outcome of leading the user into the Exposed (E) compartment with the probability (1-l).

Transferring users from the Exposed state to the Infected state can happen from two different scenarios. The first possibility is that the Exposed (E)- user who has heard about the hashtag but needs some time before tweeting about it and sharing the hashtag- may have more contact with Infected users who are users who shared the hashtag with a contact rate ρ and because of this further contact they will become Infected.

The second possibility is that the Exposed (E) user can move to the Infected compartment not because of contacting Infected users, but because of self-adoption with rate ε.

4. Analysis and Results

This section contains the discussion of the results of our applications of the SIR and SEIZ epidemiological models to each of the datasets used in this study. We fit the number of Infected people (those users who used the hashtags of interest in each experiment) in each 24-hour time interval as the Infected (I) compartment in the SIR and the SEIZ model by using the Scipy package in the Python programming environment.

The nonlinear least square solver MATLAB function called Lsqnonlin was applied to each of the datasets used in this study with the following parameters (method='Trust Region Reflective algorithm', Tolerance for termination by the change of the cost function=1e-08, Tolerance for termination by the change of the independent variables=1e-08, Tolerance for termination by the norm of the gradient=1e-08, and Characteristic scale of each variable=1.0).

To solve the ODEs, we used Odeint[1] from the same package to integrate a system of ordinary differential equations.

For every dataset there are a set of optimal parameters which can minimize the error between the real number of tweets (e.g., users of hashtags) and the estimated number of users in the Infected compartment, |I(t) - tweets(t)|. Parameter tables for all datasets are available upon request.

$$\frac{||I(t) - tweets\ (t)||_2}{||tweets(t)||_2}$$

[1] scipy.optimize.least_squares — SciPy v1.7.1 Manual, (n.d.). https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.least_squares.html (accessed August 31, 2021).

A sample of the SIR and the SEIZ model fit results for anti-lockdown were graphed in Figures 3 and 4 respectively. In addition, a sample of the SIR and the SEIZ model fit results for pro-mask hashtags were graphed in Figures 5 and 6 respectively, and will be discussed in the subsequent section.

The blue dots on the SEIZ model fit figures 4 and 6, and the green dots in the SIR model fit figures 3 and 5 are the actual tweets while the red line is the Infected (I) compartment of the models.

The “Error SEIZ”, and “Error SIR” columns in Table 3 display the difference between the actual number of tweets for every hashtag set and the Infected compartment predicted by the SEIZ and SIR model respectively, reflecting the relative error in 2-norm (Jin et al., 2013).

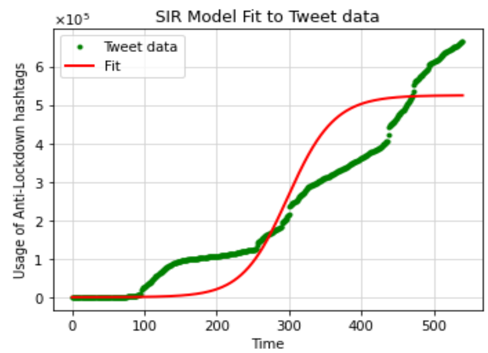

Figure 3. The SIR model fit for Anti-lockdown hashtags

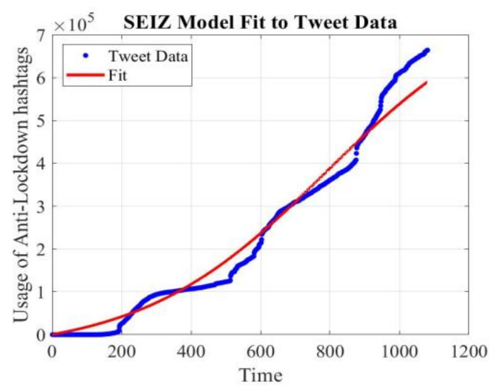

Figure 4. The SEIZ model fit for Anti-lockdown hashtags

When comparing the error of the SIR model for different hashtags, the lowest error is for the hashtag set related to Anti-mask (22.0%) and the highest error is for the hashtag related to the pro-vaccine hashtag set (30.8%). When comparing the error of the SEIZ model for different hashtags, the lowest error is for the hashtag set related to anti-lockdown (6.9%) and the highest error is for the hashtag related to the pro-vaccine hashtag set (10.5%). (Table 3).

The average error for the SEIZ model for all the datasets is 8.6% and the average error for the SIR model is 26.25%. In all the datasets, the error for the SEIZ model is less than the error for the SIR model (Table 3).

Based on the low error obtained in the results, we conclude that the SEIZ model is better able to model the spread of polarizing viewpoints represented by different hashtags related to COVID-19.

In addition, our findings show that both the SIR and the SEIZ models can model the anti-narrative of a subject more precisely than they can the pro-narrative of a subject.

The average error for the anti-viewpoint hashtags was 7.5% and 25.8% for the SEIZ and the SIR models, respectively. The average error for pro-viewpoint hashtags was 9.8% and 26.7% for the SEIZ and the SIR models, respectively. These results show that both models are more accurate with less error for anti-viewpoint hashtags.

Table 3. List of hashtags and corresponding errors when applying the SEIZ and SIR models

Category	Datasets	Error SEIZ	Error SIR
Anti-hashtags	Anti-lockdown	6.9%	26.4%
	Anti-vaccine	7.3%	29.1%
	Anti-mask	8.2%	22.0%
	Average-Anti	**7.5%**	**25.8%**
Pro-hashtags	Pro-lockdown	9.2%	23.4%
	Pro-vaccine	10.5%	30.8%
	Pro-mask	9.7%	25.8%
	Average-Pro	**9.8%**	**26.7**

In general, when we look at Table 3, both the SIR and the SEIZ models are able to model the trend of diffusion for opposing viewpoints. However, the SEIZ model is superior to the SIR model in terms of accuracy of prediction (reported errors).

Figures 3 and 4 provide an additional visual representation of the comparative performance of the SIR and SEIZ models for the Anti-lockdown narratives with regard to error. As can be seen from the figures, SIR models have difficulty modeling the anti-lockdown viewpoints compared to the SEIZ model. Recall the error metrics from Table 3. While the SEIZ model exhibits an excellent fit by 6.9% of error (Figure 4), the SIR model is relatively flawed by 26.4% of error (Figure 3).

Our next comparison is between the SEIZ and SIR performance for the anti-mask hashtag set (Figures 5 and 6). S-curve fitting was attempted in both models, but the SEIZ model had an error of 9.7%, while the SIR model had an error of 25.8%. The latter half of the trend was unable to be incorporated into the SIR model and the SIR model flattened rapidly (Figure 5).

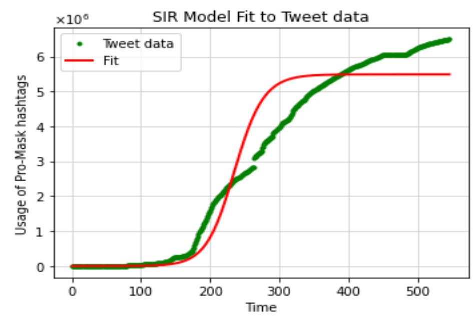

Figure 5. The SIR model fit for Pro-mask hashtags

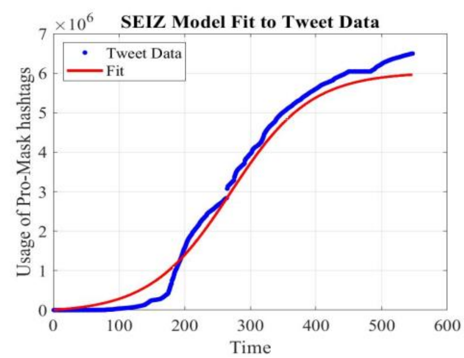

Figure 6. The SEIZ model fit for Pro-mask hashtags

To visually understand the performance of the SEIZ epidemiological model on anti and pro-subject hashtags, an example is provided in Figure 7. This figure illustrates the model's performance on anti and pro-lockdown hashtags. As shown, the propagation trends of the opposing viewpoints differ significantly, leading to distinct fits of the SEIZ model for each dataset. A complete set of visualizations comparing the performance of the SIR and SEIZ models across all pro- and anti-subjects is available upon request.

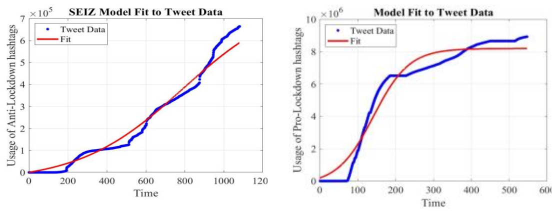

Figure 7. Comparison of the SEIZ model fit for anti and pro-lockdown hashtags

In the next and final section of this work, we present our overall conclusions. We present the highlights in terms of our research objectives. We also present our ideas for future work.

5. Conclusions and Future Work

In this study, we represented how the propagation of polarizing ideas regarding COVID-19 on Twitter (X) can be modeled by applying the SIR and SEIZ epidemiological models. We applied the I compartment of the SIR and the SEIZ model to six different datasets containing a comprehensive set of hashtags for anti-subject and pro-subject hashtags.

Our findings demonstrated that the SEIZ model is more accurate with a relatively low error rate than the SIR model in all datasets. The lowest error for the SEIZ model is 6.9% which is for the anti-lockdown dataset and the lowest error for the SIR model is 22.0% which is for the Anti-mask dataset.

Except for the pro-vaccine dataset where the error was relatively high, 10.5% in the SEIZ model, the rest of the datasets had an error less than 10%, presenting the SEIZ model's strength over the SEIZ model. Overall, the SEIZ model performs better when evaluating anti-subject ideas than pro-subject ideas.

Using the mathematical model to evaluate the spread of different opposing ideas on social media, especially Twitter (X), can provide the opportunity to study and predict trends of propagation. This evaluation can help researchers and policymakers to develop proper strategies for controlling and preventing the propagation of improper ideas and insights on social media.

In the future, we will apply additional epidemiological models to datasets from various domains, such as politics, healthcare, and religion. We will compare the results to other published findings to determine if any model outperforms the SEIZ model.

In addition, we plan to model diffusion characteristics of other compartments of the SEIZ model, such as the Skeptic (Z), Exposed (E), and Susceptible (S) compartments. As a result, we could find a way to bring more individuals from the Susceptible and Exposed compartments to the Skeptics compartment, which refers to users who decide not to tweet/share the anti-hashtags.

Additionally, we will employ stance detection methods to detect "for," "anti," and "neutral" stances on a topic over time. Given that stances can evolve, this approach enables us to perform a dynamic analysis using our longitudinal data.

Lastly, we are working on refining our methodology to operationalize it for other social media platforms such as Reddit, YouTube, TikTok, Instagram, and Telegram. This will not only provide a comparative evaluation of various platforms vis-à-vis the diffusion characteristics of COVID-19 polarizing viewpoints on online social networks but also help make our approach generalizable. We will also assess the impact of bots on the diffusion of opposing viewpoints across these platforms.

Acknowledgment

The second author's contributions to this study are funded in part by the U.S. National Science Foundation (OIA-1946391, OIA-1920920), U.S. Office of the Under Secretary of Defense for Research and Engineering (FA9550-22-1-0332), U.S. Army Research Office (W911NF-23-1-0011, W911NF-24-1-0078), U.S. Office of Naval Research (N00014-21-1-2121, N00014-21-1-2765, N00014-22-1-2318), U.S. Air Force Research Laboratory, U.S. Defense Advanced Research Projects Agency, Arkansas Research Alliance, the Jerry L. Maulden/Entergy Endowment at the University of Arkansas at Little Rock, and the Australian Department of Defense Strategic Policy Grants Program (SPGP) (award number: 2020-106-094). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations. The researchers gratefully acknowledge the support.

6. References

Abdullah, S., & Wu, X. (2011). An epidemic model for news spreading on twitter. *2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*, 163–169.

Bettencourt, L. M. A., Cintrón-Arias, A., Kaiser, D. I., & Castillo-Chávez, C. (2006). The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and Its Applications*, *364*, 513–536.

Castiello, M., Conte, D., & Iscaro, S. (2023). Using Epidemiological Models to Predict the Spread of

Information on Twitter. *Algorithms*, *16*(8). https://doi.org/10.3390/a16080391

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., … Scala, A. (2020). The COVID-19 social media infodemic. *Scientific Reports 2020 10:1*, *10*(1), 1–10. https://doi.org/10.1038/s41598-020-73510-5

Colizza, V., & Vespignani, A. (2007). Invasion threshold in heterogeneous metapopulation networks. *Physical Review Letters*, *99*(14), 148701.

Cunha, E., Magno, G., Comarela, G., Almeida, V., Gonçalves, M. A., & Benevenuto, F. (2011). Analyzing the Dynamic Evolution of Hashtags on Twitter: a Language-Based Approach. *Proceedings of the Workshop on Language in Social Media LSM 2011*, (June), 58–65.

DALEY, D. J., & KENDALL, D. G. (1965). Stochastic Rumours. *IMA Journal of Applied Mathematics*, *1*(1), 42–55. https://doi.org/10.1093/imamat/1.1.42

Daley, D. J., Kendall, D. G., DJ, D., & DG, K. (1964). Epidemics and rumours. *Nature*, *204*(4963), 1118.

Escalante, R., & Odehnal, M. (2020). A deterministic mathematical model for the spread of two rumors. *Afrika Matematika*, *31*(2), 315–331. https://doi.org/10.1007/S13370-019-00726-8

Isea, R., & Lonngren, K. E. (2017). A New Variant of the SEIZ Model to Describe the Spreading of a Rumor. *International Journal of Data Science and Analysis*, *3*(4), 28–33.

Jain, S., Sharma, V., & Kaushal, R. (2016). Towards automated real-time detection of misinformation on Twitter. *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2015–2020.

Jin, F., Dougherty, E., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on twitter. *Proceedings of the 7th Workshop on Social Network Mining and Analysis*, 1–9.

Kabir, K. M. A., Kuga, K., & Tanimoto, J. (2019). Analysis of SIR epidemic model with information spreading of awareness. *Chaos, Solitons & Fractals*, *119*, 118–125.

Kauk, J., Kreysa, H., & Schweinberger, S. R. (2021). Understanding and countering the spread of conspiracy theories in social networks: Evidence from epidemiological models of Twitter data. *PLoS ONE*, *16*(8 August), 1–20. https://doi.org/10.1371/journal.pone.0256179

Kereselidze, N. (2019). Mathematical and computer RAI models of false information dissemination. *WSEAS Transactions on Computer Research*, *7*, 56–63.

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., … others. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096.

Li, H. O.-Y., Bailey, A., Huynh, D., & Chan, J. (2020). YouTube as a source of information on COVID-19: a pandemic of misinformation? *BMJ Global Health*, *5*(5), e002604.

Maleki, M., Arani, M., Buchholz, E., Mead, E., & Agarwal, N. (2021). Applying an Epidemiological Model to Evaluate the Propagation of Misinformation and Legitimate COVID-19-Related Information on Twitter. *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, 23–34. https://doi.org/10.1007/978-3-030-80387-2_3

Maleki, M., Arani, M., Mead, E., & Agarwal, N. (2023). Measuring the Interference Effect of Bots in Disseminating Opposing Viewpoints Related to COVID-19 on Twitter Using Epidemiological Modeling. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2023-Janua*, 2296–2305.

Maleki, M., Mead, E., Arani, M., & Agarwal, N. (2021). *Using an Epidemiological Model to Study the Spread of Misinformation during the Black Lives Matter Movement*. https://doi.org/10.48550/arxiv.2103.12191

Managing the COVID-19 infodemic: Promoting healthy... - Google Scholar. (n.d.). Retrieved February 20, 2022, from https://scholar.google.com/scholar?hl=en&as_sdt=0%2C4&q=Managing+the+COVID-19+infodemic%3A+Promoting+healthy+behaviours+and+mitigating+the+harm+from+misinformation+and+disinformation&btnG=

Obadimu, A., Mead, E., Maleki, M., & Agarwal, N. (2020). Developing an Epidemiological Model to Study Spread of Toxicity on YouTube. *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*, 266–276.

reports, S. M.-S., & 2016, undefined. (n.d.). Six susceptible-infected-susceptible models on scale-free networks. *Nature.Com*.

Rodrigues, H. S., & Fonseca, M. J. (2016). Can information be spread as a virus? Viral marketing as epidemiological model. *Mathematical Methods in the Applied Sciences*, *39*(16), 4780–4786.

Shu, K., Wang, S., Lee, D., & Liu, H. (2020). Mining Disinformation and Fake News: Concepts, Methods, and Recent Advancements. *ArXiv Preprint ArXiv:2001.00623*.

Tasnim, S., Hossain, M. M., & Mazumder, H. (2020). Impact of rumors and misinformation on COVID-19 in social media. *Journal of Preventive Medicine and Public Health*, *53*(3), 171–174.

Van Bavel, J. J., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., … others. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, *4*(5), 460–471.

Woo, J., & Chen, H. (2016). Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog. *SpringerPlus*, *5*(1), 1–19. https://doi.org/10.1186/S40064-016-1675-X

Works-Sprogvidenskabeligt, J. R.-J. of L., & 2020, undefined. (2020). SEIZ Matters: Modelling the spread of concepts on Twitter. *Tidsskrift.Dk*, *5*(1).

Xiong, F., Liu, Y., Zhang, Z., Zhu, J., & Zhang, Y. (2012). An information diffusion model based on retweeting mechanism for online social media. *Physics Letters A*, *376*(30–31), 2103–2108.

Zhao, L., Wang, J., Chen, Y., Wang, Q., Cheng, J., & Cui, H. (2012). SIHR rumor spreading model in social networks. *Physica A: Statistical Mechanics and Its Applications*, *391*(7), 2444–2453.

Zhao, L., Wang, Q., Cheng, J., Chen, Y., Wang, J., & Huang, W. (2011). Rumor spreading model with consideration of forgetting mechanism: A case of online blogging LiveJournal. *Physica A: Statistical Mechanics and Its Applications*, *390*(13), 2619–2625.