Are Narratives Contagious? Modeling Narrative Diffusion Using Epidemiological Theories

Mayor Inna Gurung[0009−0003−5828−3189], Nitin Agarwal[0000−0002−5612−4753], and Ahmed Al-Taweel[0000−0002−9396−7696]

COSMOS Research Center, University of Arkansas, Little Rock, USA
{mgurung,nxagarwal,asaltaweel}@ualr.edu

Abstract. We live in a hyper-connected world where narratives travel at lightning speed, transforming individuals into groups and groups into mobs. Recent events illustrate how narratives on social media can be weaponized and spread rapidly to sow discord. These insidious threats, aimed at influencing beliefs and behaviors, must be viewed as modern weapons of cognitive warfare. While cyber capabilities are crucial for multi-domain operations (i.e., land, sea, air), today's challenges extend beyond traditional cybersecurity to include information and influence operations targeting individuals at the micro level and groups or societies at the macro level. This study hypothesizes that narratives spread like diseases and utilizes epidemiological theories to evaluate this hypothesis. Key research questions include: Do narratives spread like contagions? Can we identify and measure which narratives are more infectious? Comparing the SIR and SEIZ models, which compartmentalize the target audience into susceptible (S), infected (I), recovered (R), and/or skeptical (Z), this study aims to model narrative diffusion and objectively measure infection rates across various narratives. Data from YouTube, a prominent social media platform in the South China Sea region, is analyzed to examine competing narratives and validate the research. Ultimately, this research aims to inform science-driven decision-making by prioritizing narratives with intervention strategies tailored to individuals in specific compartments.

Keywords: Narrative Dissemination · Epidemiological Modeling · LLMs.

1 Introduction

Narratives, structured accounts of events or experiences, significantly influence social media by shaping beliefs and behaviors. When widely shared, they swiftly impact public opinion, especially on platforms like YouTube, where users collectively spend 720 million hours per day [17]. Studies on social media information dissemination have focused on rumors, facts, misinformation, and toxicity using epidemiological models [3, 10, 22]. However, these studies often rely on textual data and metadata, which may not fully capture the essence of posts.

Understanding narratives on YouTube is essential due to its role in shaping public opinion, political discourse, and societal behaviors. Researchers have highlighted the spread of anti-media populism in the Philippines by YouTube influencers, revealing its impact on political brokerage and governance [20]. Studies on deradicalization efforts have identified and addressed religiously intolerant Arabic videos, suggesting personalized interventions [2]. Analysis of COVID-19 information fatigue shows a shift from scientific to political discourses, affecting public health communication [19]. Additionally, platform biases favoring certain content emphasize the need to understand YouTube's influence [5, 18]. Research by [8, 6] indicates that while titles often sensationalize content, actual narratives can be less provocative. This underscores the importance of analyzing narratives beyond metadata.

Our study proposes a method to extract and track narrative spread on YouTube, focusing on its role as the third-most-popular social media platform in Southeast Asia. This is particularly relevant for our case study on the South China Sea Dispute. Using two epidemiological models, SIR and SEIZ, we aim to provide insights into how narratives propagate and influence audiences in digital environments, addressing the following research questions: **RQ1**: Do narratives spread like contagions? **RQ2**: Can we identify and measure which narratives exhibit higher infectiousness, thereby affecting a larger audience?

Background: The South China Sea Dispute involves multiple nations, including China, Vietnam, the Philippines, Malaysia, Brunei, and Taiwan, all claiming sovereignty over various islands, reefs, and maritime areas. This region holds strategic importance due to its vital shipping lanes, rich fishing grounds, and potential underwater oil and gas reserves. China's expansive claims, epitomized by the "nine-dash line," have sparked significant diplomatic and military tensions. Understanding narratives is critical as they shape perceptions, inform policy decisions, legitimize territorial claims, and mobilize international support. Key research by [7, 9, 21] delves into the geopolitical complexities, historical context, and the role of narratives in the South China Sea Dispute.

2 Related Work

This literature review synthesizes research on epidemiological models and their use in understanding epidemic dynamics and information diffusion across digital platforms. Compartmental models categorize populations into groups to study disease or information spread dynamics. The SI model classifies populations into susceptible and infected groups, evolving into the SIS model where infected individuals can revert to a susceptible state [11, 13]. The SIR model further partitions individuals into susceptible, infected, and recovered compartments [15]. The SEIZ model (susceptible, exposed, infected, and skeptic) simulates the spread of news and rumors on platforms like Twitter [10], with detailed descriptions in later sections. Similar models explore toxicity propagation in YouTube comments, showcasing their potential to mitigate harmful content spread [16].

Beyond epidemiological models, [24] provide a comprehensive survey of models, predictions, and recent advancements in understanding information diffusion across social networks. [14] investigate temporal dynamics in information spread, detailing the rise and fall patterns of cascades and their broader implications. [23] introduce Seismic, a self-exciting point process model designed to forecast tweet popularity, highlighting the self-reinforcing nature of information propagation. These studies illuminate various aspects of information cascade dynamics, offering foundational insights into how information spreads and shapes behavior within digital ecosystems. Despite advancements in understanding information diffusion on social media, adapting these models to the complex data structure of YouTube remains challenging. There is a significant gap in methodologies that can comprehensively capture and analyze diverse narratives across different digital platforms, including video narratives. Future research should focus on developing such methodologies to enhance predictive capabilities in digital ecosystem analysis.

3 Methodology

In this section, we detail our data acquisition process and outline the preprocessing steps like narrative extraction and clustering.

Data Collection: We utilized the YouTube Data API to gather data using keywords formulated by a subject matter expert. The keywords include territorial disputes, maritime claims, freedom of navigation, the nine-dash line, China-ASEAN relations, regional security, international law (UNCLOS), and the South China Sea. This process resulted in a dataset of 8,920 videos spanning from 2007 to 2023. Given the extensive dataset, we analyzed posting frequency by year to identify significant periods, discovering that 2022 had the highest data influx. Consequently, we selected 2022, which had 2,682 videos, for our research. Additionally, we obtained local transcripts from these videos and employed the Whisper API to convert audio to text for videos lacking transcripts, following the methodology suggested by [4].

Extraction of Narratives: After transcribing the videos, we extracted narratives using GPT-4o (omni). This involved adjusting several key parameters: temperature controlled the randomness of the output, with lower values making it more deterministic. Top-k sampling limited the pool to the top k probable tokens, while top-p (Nucleus) sampling focused on tokens whose cumulative probability exceeded p, ensuring high-likelihood selections. Max tokens defined the maximum length of the generated output and the prompt. These parameters were chosen based on insights from [8].

Clustering of Narratives:

Macro Narratives: For our study, we focused on the two clusters with the largest data points. These are the following:

Impact of Russia-Ukraine Conflict (Cluster 0): The Russia-Ukraine conflict has significantly impacted global economics, politics, and society. It has disrupted trade, caused market volatility, heightened geopolitical tensions, and led to sanctions on Russia and military support for Ukraine. Though not directly related to the South China Sea Dispute, many videos about the topic frequently

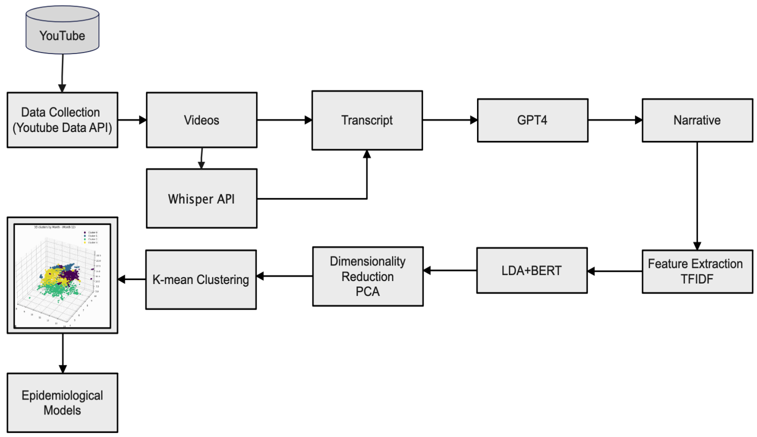


Fig. 1: The figure shows the methodology for acquiring YouTube data, extracting narratives with GPT-4, and clustering them into thematic groups.

reference the Russia-Ukraine conflict. These references often highlight how China perceives the conflict as an opportunity to test the waters for its expansion plans.

The US as the Aggressor in the South China Sea (Cluster 1): This narrative highlights America's growing influence in the region, particularly through the establishment of military bases near disputed areas. It implies that the US is interfering in the conflict, positioning itself as an aggressor.

4 Epidemiological Model

In this section, we use SIR and SEIZ models to analyze the narrative spread on YouTube about the South China Sea Dispute, and we consider fractional derivatives and integrals, followed by model comparisons.

Table 1: Redefining the SIR model to adapt to YouTube [12].

Scope	N	S	I	R
Epidemiology	Total population	Susceptible to disease	Currently infected	Removed (immune or deceased)
YouTube Data	All users	Users not supporting the topic	Users supporting the topic	Users initially supporting but now unsupporting

Adapting the SIR Model for YouTube Narratives: The SIR framework divides the population into three segments: Susceptible (S), Infected (I), and Recovered (R) [1]. This study presents a comprehensive framework for applying the SIR model to YouTube, examining how narratives spread across YouTube channels. By adopting the SIR compartment model [12], we redefine its terms to align with YouTube dynamics, as detailed in Table 1. The system of ordinary differential equations (ODEs) presented below represents the SIR model:

$$\begin{cases} \frac{dS}{dt} &= \frac{-\beta SI}{N}, \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \gamma I, \\ \frac{dR}{dt} &= \gamma I. \end{cases} \tag{1}$$

Equation 1 involves various parameters, and more detailed explanations are listed in Table 2. In regard to the temporal direction, we apply the Caputo

Table 2: Parameter definitions in the SIR model equation 1.

Parameter	Meaning	Cluster 0 Estimation	Cluster 1 Estimation
β	Narratives spreading rate	0.15	0.3
γ	Recovery rate	0.11	0.11
$\Re_0 = \frac{\beta}{\gamma}$	Basic reproduction number	1.36	2.7
$S(0) + I(0) + R(0) = N$	Population size	$\approx$ 10K	$\approx$ 10K

fractional-order derivative to the SIR formulation equation 1, resulting in the following system of ODEs.

$$\begin{cases} {}_0^C D_t^\alpha S &= \frac{-\beta SI}{N}, \\ {}_0^C D_t^\alpha I &= \frac{\beta SI}{N} - \gamma I, \\ {}_0^C D_t^\alpha R &= \gamma I. \end{cases} \tag{2}$$

Adapting the SEIZ Model for YouTube Narratives: The SIR model's limitation for modeling narrative spread on platforms like YouTube is its one-way transition from Susceptible to Infected. Social media interactions are multidirectional, involving sharing, commenting, and responses, making this approach too simplistic. To better depict online narrative dynamics, we use the SEIZ model, which includes Skeptic and Exposed components [10]. The SEIZ model comprises four compartments: Susceptible, Exposed, Infected, and Skeptic as shown in Table 3. When a Susceptible channel (S) encounters a narrative, it can quickly become an Infected channel (I) with probability *p*. Alternatively, it may analyze the narrative and move to Exposed (E) status with probability *(1-p)*.

Table 3: Redefining the SEIZ model to adapt to YouTube.

Scope	N	S	E	I	Z
Epidemiology	Total population	Susceptible to disease	Exposed but not yet infectious	Currently infected	Skeptic (immune or resistant)
YouTube Data	All channels	Channels that have not yet encountered the narrative	Channels exposed to the narrative but delayed in engagement	Channels actively posting about the narrative	Channels exposed but choosing not to engage

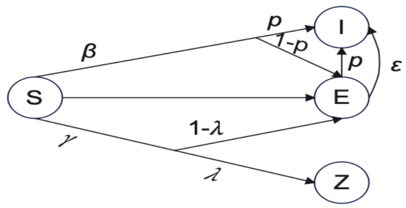


Fig. 2: SEIZ compartment model.

As depicted in Fig. 2, the SEIZ model rules can be summarized as follows:

$$\begin{cases} \frac{dS}{dt} &= -\frac{\beta SI}{N} - \frac{\gamma SZ}{N}, \\ \frac{dE}{dt} &= \frac{(1-p)\beta SI}{N} + \frac{(1-\lambda)\gamma SZ}{N} - \frac{\eta EI}{N} - \epsilon E, \\ \frac{dI}{dt} &= \frac{p\beta SI}{N} + \frac{\eta EI}{N} + \epsilon E, \\ \frac{dZ}{dt} &= \frac{\lambda\gamma SZ}{N}. \end{cases} \tag{3}$$

The system of ODEs in equation 3 involves various parameters, and more detailed explanations are listed in Table 4.

Table 4: Parameter definitions in the SEIZ model.

Parameter	Definition	Estimation (Cluster 0)	Estimation (Cluster 1)
β	S-I contact rate	0.15	0.3
γ	S-Z contact rate	0.08	0.11
ϵ	E-I contact rate	0.0005	0.001
p	Probability of S to I given contact with I	0.05	0.1
$1-p$	Probability of S to E given contact with I	0.95	0.9
η	Transition rate of E to I (Incubation rate)	0.0008	0.001
$1-\lambda$	Probability of S to E given contact with Z	0.98	0.99
λ	Probability of S to Z given contact with Z	0.02	0.01

Additionally, we apply the Caputo fractional to the SEIZ formulation equation 3, resulting in the following system of ODEs.

$$\begin{cases} {}_0^C D_t^\alpha S &= -\frac{\beta SI}{N} - \frac{\gamma SZ}{N}, \\ {}_0^C D_t^\alpha E &= \frac{(1-p)\beta SI}{N} + \frac{(1-\lambda)\gamma SZ}{N} - \frac{\eta EI}{N} - \epsilon E, \\ {}_0^C D_t^\alpha I &= \frac{p\beta SI}{N} + \frac{\eta EI}{N} + \epsilon E, \\ {}_0^C D_t^\alpha Z &= \frac{\lambda\gamma SZ}{N}. \end{cases} \tag{4}$$

Parameter Learning: Various parameters must be defined to apply the SIR and SEIZ models outlined in equations (1) and (3). These parameters include contact rates (β, γ, η), probabilities (p, λ), and a transition rate ϵ. Furthermore, to solve the ODEs associated with the models, initial values for each component must be specified: $S(t_0)$, $E(t_0)$, $I(t_0)$, and $Z(t_0)$. Since these parameters and the total population size (N), representing the sum of channels across all four compartments, are initially unknown, they need to be determined through evaluation. The experimental implementation was conducted in Python. This Python implementation utilized the functionality offered by the *scipy.optimize* module for performing a *least squares* fit. The systems of ODEs were solved using the *odeint* method, a fundamental component of the *scipy.integrate* module.

The adoption of these modules enabled efficient optimization and precise numerical integration, enhancing the robustness of the model fitting process. Utilizing the Nelder-Mead optimization method within the *scipy.optimize.minimize* function ensured the convergence of parameter estimation to optimal values. Optimal parameters were determined by minimizing the absolute difference between the Infected compartment (I) and corresponding YouTube data, denoted as $|I(t) - videos(t)|$. To quantify the number of videos in the Infected (I) component, we tracked cumulative counts from the start of narrative spread on January 1, organized by weeks, and used this data to define parameter optimization boundaries. Detailed results of this study are presented in the following sections.

5 Results

This section conducts numerical tests to validate theoretical results and address our research questions using the SIR and SEIZ models. Focusing on two macro-narratives outlined in Section 3, we analyze temporal aspects of YouTube data related to the South China Dispute. The experiments utilize a dataset spanning 52 weeks from January to December 2022, involving around 10,000 channels engaged with the narrative.

Fitting Data To Infected Component: In this section, we aimed to analyze channel transitions in different compartments of the SIR and SEIZ models for the YouTube dataset. A least squared criterion was applied as given by:

$$\Theta = \operatorname{argmin}_{\varphi} \left(\sum_{i=0}^{\mathrm{n}} \left(I_{cu}(t_i) - \underbrace{I_{vcu}(t_i)}_{\text{Videos}} \right)^2 \right), \tag{5}$$

Here, Θ denotes the vector containing estimated parameter values, φ represents the parameter space, t_{n} is the most recent date considered in the analysis, t_i denotes the date, $I_{cu}(t_i)$ signifies the cumulative incidence up to t_i, and $I_{vcu}(t_i)$ represents the cumulative incidence according to the model up to t_i. In order to evaluate the SIR and SEIZ models, Figures 4 and 5 show how well the SIR and SEIZ models fit two macro-narratives on the YouTube data, along with the respective 2-norm relative error

$$E_{-\mathrm{rel}} = \frac{\|I_{cu}(t_i) - I_{vcu}(t_i)\|_2}{\|I_{vcu}(t_i)\|_2}, \tag{6}$$

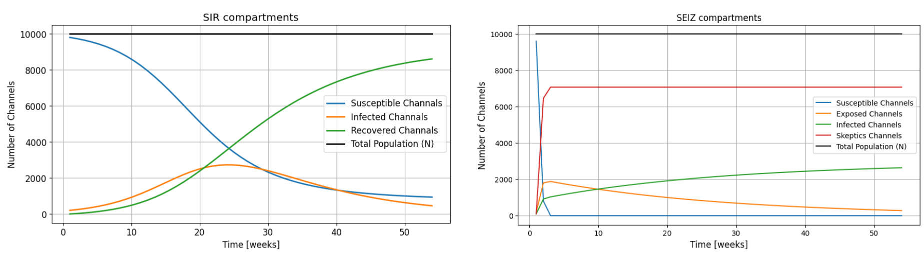
and the mean absolute error (MAE) described by:

$$MAE = \frac{\|I_{cu}(t_i) - I_{vcu}(t_i)\|_1}{n}. \tag{7}$$

The numerical examples indicate that the SEIZ model significantly outperforms the SIR model in accurately fitting YouTube data for the two narratives. The low error in the SEIZ model suggests a more precise representation of YouTube data in both cases, as compared to other models and shown in Table 5. Across these narratives, the SEIZ model consistently captures the initial spread on YouTube more effectively, a phenomenon attributed to a delay as individuals in the 'Exposed' category take time before sharing their videos.

Table 5: Comparative fitting error analysis of SIR and SEIZ Models in YouTube data.

Epidemiology Model	Cluster 0	Cluster 1
SIR	0.0606	0.0628
SEIZ	0.0217	0.0381



(a) SIR model predictions: US as the aggressor in the South China Sea.

(b) SEIZ model predictions: US as the aggressor in the South China Sea.

Fig. 3: Comparison of SIR and SEIZ model predictions for US aggression in the South China Sea.

Error Estimates: In this study, we apply the SIR and SEIZ mathematical models and present numerical performance metrics for the two macro-narratives outlined in the data section. Figures 4 and 5 demonstrate that the SEIZ model achieves lower error rates of 0.0217 and 34.7145, respectively, compared to the SIR model with error rates of 0.0628 and 105.5941, for YouTube data in L^2-norm and MAE at the final time. These results address **RQ1**, illustrating that narratives spread akin to contagions, as evidenced by our ability to adjust transmission and recovery rates to fit dissemination patterns with minimal error. Similarly, Figures 6 and 7 show that the SEIZ model produces lower error rates (0.0381 and 75.4) compared to the SIR model (0.0628 and 115.8670) for YouTube data in L^2-norm and MAE at the final time.

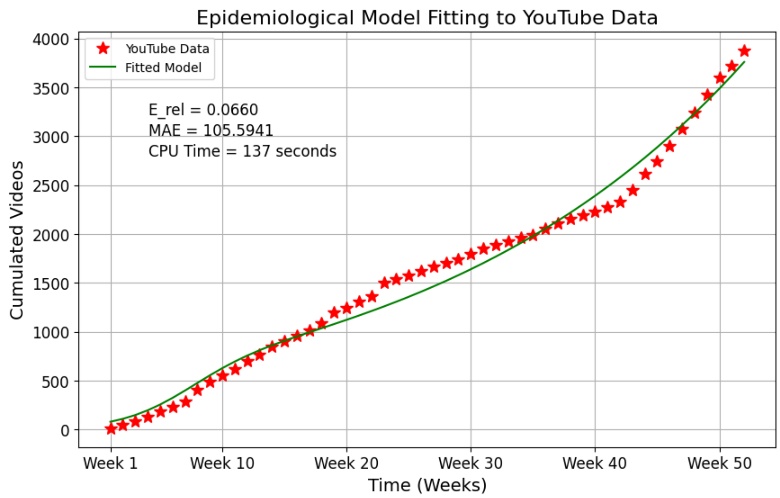


Fig. 4: SIR model predictions of the compartment sizes over time: analyzing the impact of the Russia-Ukraine Conflict on China's geopolitical strategy.

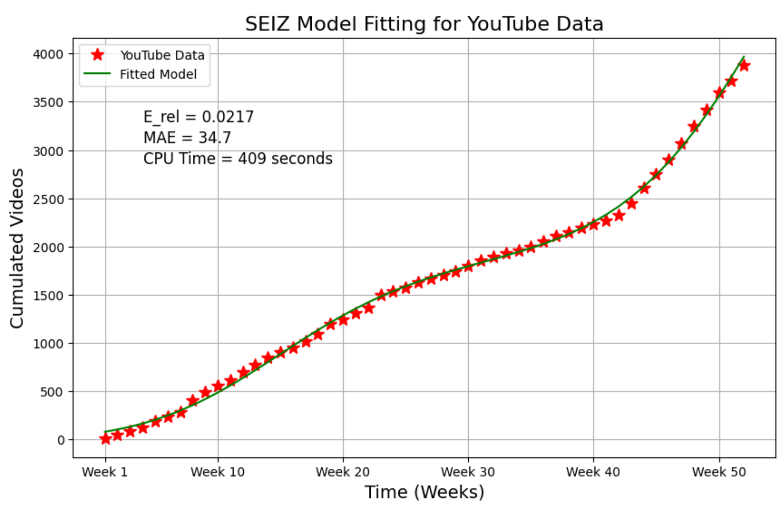


Fig. 5: SEIZ model predictions of the compartment sizes over time: analyzing the impact of the Russia-Ukraine Conflict on China's geopolitical strategy.

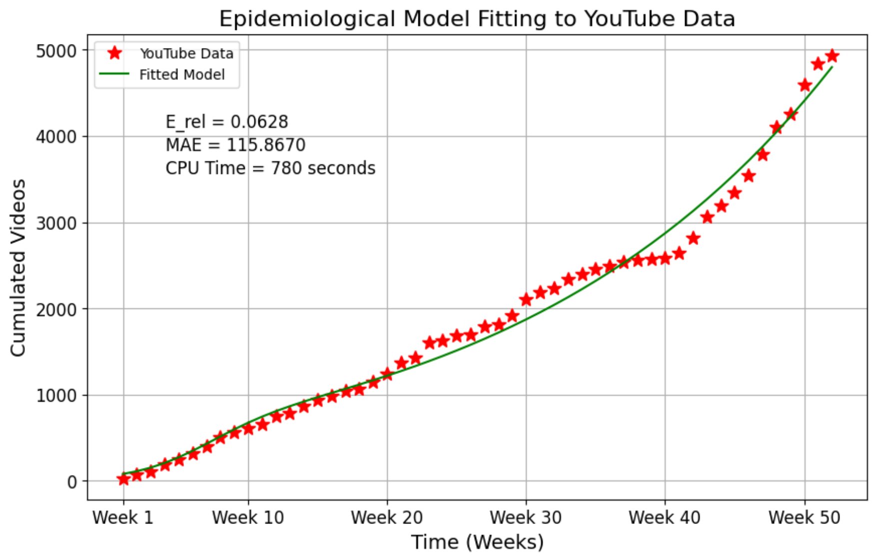


Fig. 6: SIR model predictions of the compartment sizes over time: US as the aggressor in the South China Sea).

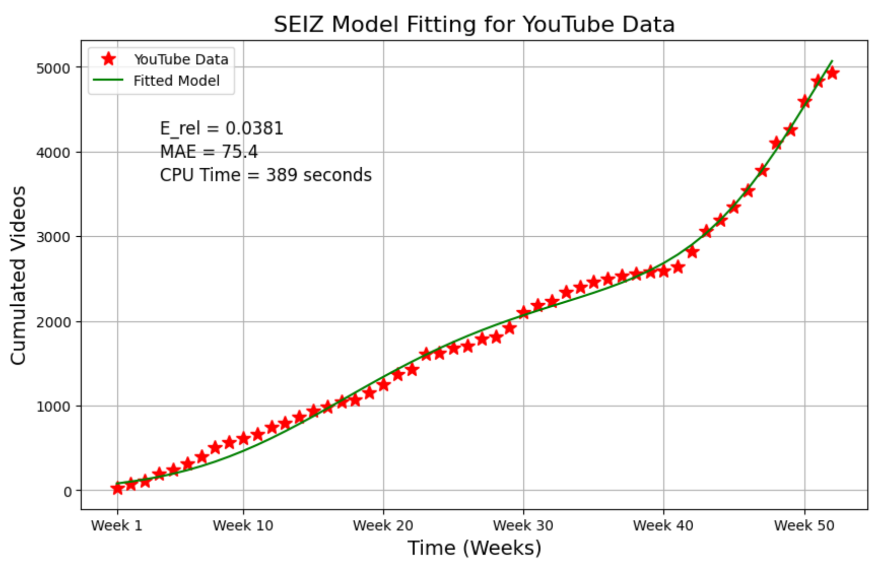


Fig. 7: SEIZ model predictions of the compartment sizes over time: US as the aggressor in the South China Sea.

Caputo Fractional Derivative: Many natural phenomena, like epidemiological dynamics, exhibit time-memory effects, revealing insights into nonlocal biological processes. Fractional derivative models, with their time-dependent kernels, effectively address these issues. Using the Caputo fractional derivative in this study is beneficial as it shares initial conditions with classical derivatives, avoiding the need for fractional initial values. Due to these benefits, we modify the SIR model equation 1 and SEIZ model equation 3 by incorporating the Caputo fractional time derivative as shown in equations 2 and 4. This test involves numerically simulating the spread of a narrative using Caputo fractional derivatives for SIR and SEIZ epidemic formulation equations (2) and (4), respectively. The simulation relies on the following initially reported information: $S(0) = 500, I(0) = 100, R(0) = 0, E(0) = 200$, and $Z(0) = 50$. We adopted the parameter values for SIR and SEIZ formulation equations (2) and (4), respectively, as outlined in Tables 2 and 4. The simulation was implemented using the Caputo fractional scheme in Python. In Figures 8 and 9, YouTube channels classified as exposed (E), infected (I), susceptible (S), skeptic (Z), and recovered (R) demonstrated specific trends for various α values, showing the behavior of population groups based on narrative status. Exposed and infected channels reached a peak and then stabilized, associated with an extended duration before stopping content posting. On the other hand, susceptible channels decreased while recovered channels increased, indicating shifts in YouTube's narrative spread dynamics, where channels may cease posting and later re-post content. Furthermore, to address our **RQ2**, we examined the rate of change from susceptible individuals to infected people to identify the more infectious narrative. The parameter β determines the infection rate. The parameter descriptions and estimations are presented in Table 4. In this case, the narrative portraying the US as the aggressor was the more infectious narrative, with a β value of 0.3, while the narrative concerning the Russia-Ukraine conflict was less infectious, with a β value of 0.15. **Analyzing the Reproduction Number $\Re_0$ with Model Parameters:** The basic reproduction number is described as the average number of secondary infections generated by one infected individual within a completely susceptible population. The primary application lies in assessing whether a new infectious disease has the potential to proliferate among a population. This test uses relevant contour plots to explain the impact of critical parameters in SIR model equation 2 and SEIZ model equation 3 on the reproductive number $\Re_0$. Figure 11 illustrates how $\Re_0$ changes with the effective contact rates β and γ. $\Re_0$ is higher when both β and γ decrease below 1 in the SIR model. Figure 11 shows how $\Re_0$ changes with the effective contact rates β and $1 - p$ in the SEIZ model. We observe that $\Re_0$ has a smaller value when both β and $1 - p$ decrease below 1.

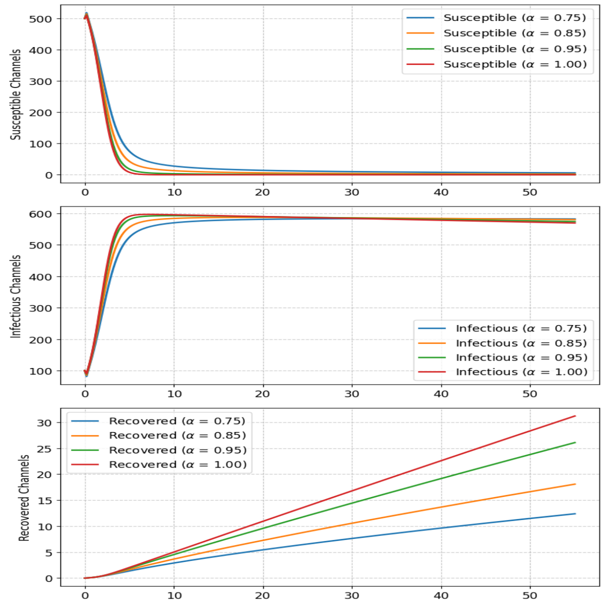


Fig. 8: Numerical simulation of the SIR model employing the Caputo differential operator with various fractional orders α: US aggression in the South China Sea.

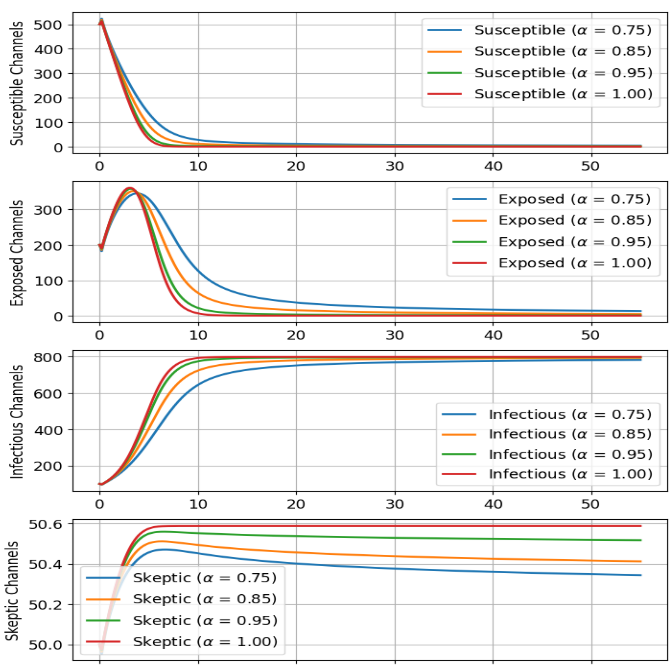


Fig. 9: Numerical simulation of the SEIZ model employing the Caputo differential operator with various fractional orders α: US aggression in the South China Sea.

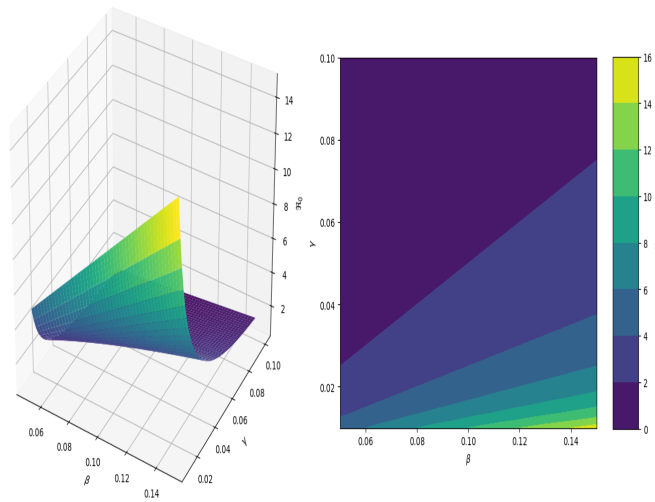


Fig. 10: Numerical simulation of the SIR model without fractional components: 3D surface plot of basic reproduction number $\Re_0$ (right) and contour plot of basic reproduction number $\Re_0$ (left).

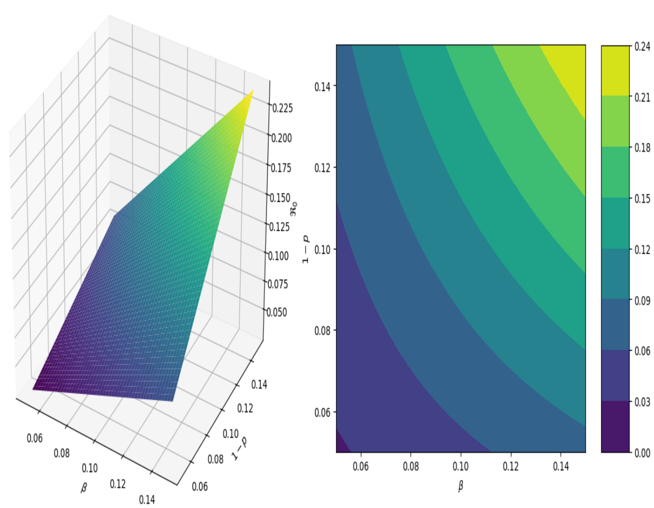


Fig. 11: Numerical simulation of the SEIZ model without fractional components: 3D surface plot of $\Re_0$ (right) and contour plot of $\Re_0$ (left).

6 Conclusion and Future Work

In this study, we analyzed YouTube narratives about the South China Sea Dispute to answer two research questions. We identified two main narratives, showing they spread like contagions, supported by the SEIZ model. We measured the infection rate β and found optimal values matching real-world data. This analysis helped us identify infectious narratives.We also explored narrative dynamics using the Caputo derivative, providing insights into stability and evolution. Both SIR and SEIZ models indicated control over narrative spread when transmission and recovery rates dropped below 1, with $\Re_0$ reflecting this potential.Our findings revealed that participants often deliberated before spreading narratives, especially those highlighting America's aggression, which showed higher infection rates. Channels like *South China Morning Post* and *Al Jazeera English* significantly influenced these narratives. In conclusion, our study highlights the contagion-like spread of narratives on YouTube and the effectiveness of SEIZ models with Caputo derivatives for analysis. This research aids policymakers in combating misinformation by targeting uninfluenced skeptics and creating counter-narratives, promoting informed digital environments and healthier public discourse.

Acknowledgments

This research is 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). 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.

References

[1] Saeed Abdullah and Xindong Wu. "An epidemic model for news spreading on twitter". In: *2011 IEEE 23rd international conference on tools with artificial intelligence*. IEEE. 2011, pp. 163–169.

[2] Nuha Albadi, Maram Kurdi, and Shivakant Mishra. "Deradicalizing YouTube: Characterization, Detection, and Personalization of Religiously Intolerant Arabic Videos". In: *Proceedings of the ACM on Human-Computer Interaction* 6.CSCW2 (2022), pp. 1–25.

[3] Luís MA Bettencourt, Ariel Cintrón-Arias, David I Kaiser, and Carlos Castillo-Chávez. "The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models". In: *Physica A: Statistical Mechanics and its Applications* 364 (2006), pp. 513–536.

[4] Mert Can Cakmak and Nitin Agarwal. "High-Speed Transcript Collection on Multimedia Platforms: Advancing Social Media Research through Parallel Processing". In: *2024 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*. 2024, pp. 857–860. DOI: 10.1109/IPDPSW63119.2024.00153.

[5] Mert Can Cakmak, Obianuju Okeke, Ugochukwu Onyepunuka, Billy Spann, and Nitin Agarwal. "Analyzing Bias in Recommender Systems: A Comprehensive Evaluation of YouTube's Recommendation Algorithm". In: *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*. 2023, pp. 753–760.

[6] Mert Can Cakmak, Obianuju Okeke, Ugochukwu Onyepunuka, Billy Spann, and Nitin Agarwal. "Investigating Bias in YouTube Recommendations: Emotion, Morality, and Network Dynamics in China-Uyghur Content". In: *Complex Networks & Their Applications XII*. Ed. by Hocine Cherifi, Luis M. Rocha, Chantal Cherifi, and Murat Donduran. Cham: Springer Nature Switzerland, 2024, pp. 351–362. ISBN: 978-3-031-53468-3. DOI: 10.1007/978-3-031-53468-3_30.

[7] M Taylor Fravel. "China's strategy in the South China Sea". In: *Contemporary Southeast Asia* (2011), pp. 292–319.

[8] Mayor Inna Gurung, Md Monoarul Islam Bhuiyan, Ahmed Al-Taweel, and Nitin Agarwal. "Decoding YouTube's Recommendation System: A Comparative Study of Metadata and GPT-4 Extracted Narratives". In: *Companion Proceedings of the ACM on Web Conference 2024*. WWW '24. Singapore, Singapore: Association for Computing Machinery, 2024, 1468–1472. ISBN: 9798400701726. DOI: 10.1145/3589335.3651913. URL: https://doi.org/10.1145/3589335.3651913.

[9] Bill Hayton. *The South China Sea: the struggle for power in Asia*. Yale University Press, 2014.

[10] Fang Jin, Edward Dougherty, Parang Saraf, Yang Cao, and Naren Ramakrishnan. "Epidemiological modeling of news and rumors on twitter". In: *Proceedings of the 7th workshop on social network mining and analysis*. 2013, pp. 1–9.

[11] KM Ariful Kabir, Kazuki Kuga, and Jun Tanimoto. "Analysis of SIR epidemic model with information spreading of awareness". In: *Chaos, Solitons & Fractals* 119 (2019), pp. 118–125.

[12] Julian Kauk, Helene Kreysa, and Stefan R Schweinberger. "Understanding and countering the spread of conspiracy theories in social networks: Evidence from epidemiological models of Twitter data". In: *Plos one* 16.8 (2021), e0256179.

[13] Ting Liu, Ping Li, Yan Chen, and Jie Zhang. "Community size effects on epidemic spreading in multiplex social networks". In: *PloS one* 11.3 (2016), e0152021.

[14] Yasuko Matsubara, Yasushi Sakurai, B Aditya Prakash, Lei Li, and Christos Faloutsos. "Rise and fall patterns of information diffusion: model and implications". In: *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*. 2012, pp. 6–14.

[15] Maria Marily Nika. "Synthedemic Modelling and Prediction of Internet-based Phenomena". In: (2014).

[16] Adewale Obadimu, Esther Mead, Maryam Maleki, and Nitin Agarwal. "Developing an Epidemiological Model to Study Spread of Toxicity on YouTube". In: *Social, Cultural, and Behavioral Modeling*. Ed. by Robert Thomson, Halil Bisgin, Christopher Dancy, Ayaz Hyder, and Muhammad Hussain. Cham: Springer International Publishing, 2020, pp. 266–276.

[17] Daniel Ruby. "YouTube Statistics—Insights And Infographics (2023)". In: *DemandSage, June* 6 (2023).

[18] Shadi Shajari, Mustafa Alassad, and Nitin Agarwal. "Characterizing Suspicious Commenter Behaviors". In: *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*. 2023, pp. 631–635. DOI: 10.1145/3625007.3627309.

[19] Chyun-Fung Shi, Matthew C So, Sophie Stelmach, Arielle Earn, David JD Earn, and Jonathan Dushoff. "From science to politics: COVID-19 information fatigue on YouTube". In: *BMC Public Health* 22.1 (2022), p. 816.

[20] Cheryll Ruth R Soriano and Marie Fatima Gaw. "Broadcasting anti-media populism in the Philippines: YouTube influencers, networked political brokerage, and implications for governance". In: *Policy & Internet* 14.3 (2022), pp. 508–524.

[21] Shogo Suzuki. "The rise of the Chinese 'Other'in Japan's construction of identity: Is China a focal point of Japanese nationalism?" In: *Identity Change and Foreign Policy*. Routledge, 2015, pp. 95–116.

[22] Marcella Tambuscio, Giancarlo Ruffo, Alessandro Flammini, and Filippo Menczer. "Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks". In: *Proceedings of the 24th international conference on World Wide Web*. 2015, pp. 977–982.

[23] Qingyuan Zhao, Murat A Erdogdu, Hera Y He, Anand Rajaraman, and Jure Leskovec. "Seismic: A self-exciting point process model for predicting tweet popularity". In: *Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining*. 2015, pp. 1513–1522.

[24] Fan Zhou, Xovee Xu, Goce Trajcevski, and Kunpeng Zhang. "A survey of information cascade analysis: Models, predictions, and recent advances". In: *ACM Computing Surveys (CSUR)* 54.2 (2021), pp. 1–36.