



HAL
open science

Investigating the Dynamics of Polarization in Online Discourse During COVID-19 Pandemic

Samrat Gupta, Gaurav Jain, Amit Anand Tiwari

► **To cite this version:**

Samrat Gupta, Gaurav Jain, Amit Anand Tiwari. Investigating the Dynamics of Polarization in Online Discourse During COVID-19 Pandemic. 20th Conference on e-Business, e-Services and e-Society (I3E), Sep 2021, Galway, Ireland. pp.704-709, 10.1007/978-3-030-85447-8_58 . hal-03648117

HAL Id: hal-03648117

<https://inria.hal.science/hal-03648117>

Submitted on 21 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Investigating the Dynamics of Polarization in Online Discourse during COVID-19 Pandemic

Samrat Gupta¹, Gaurav Jain¹, Amit Anand Tiwari²

¹ Indian Institute of Management, Ahmedabad, India

² Indian Institute of Management, Rohtak, India

Abstract. The socio-cultural polarization induced due to information and communication technology because of the selective online exposure during COVID-19 has been a major cause of concern around the globe. In this paper, we use random network theory-based simulation technique to investigate the temporal dynamics of opinion formation on YouTube videos. Our findings reveal that as the pandemic unfolded, the degree of polarization in the online discourse has increased with time. This study is significant for understanding that online discourse on sociocultural issues can lead to polarization particularly in crisis situations such as a pandemic and exacerbate the social divide.

Keywords: Polarization, COVID-19 pandemic, Online opinion formation, Consensus formation, Social media

1 Introduction

In December 2019, an outbreak of infectious disease, later named as coronavirus disease 2019 (COVID-19) and declared a pandemic, was first reported to WHO¹. This unsettling pandemic has been generating a massive amount of data across the world each day. This data generation is a result of people using social media platforms to search the information and express their views related to the disease and to reduce isolation usually associated with the anxiety as well as long-term distress during the pandemic [1].

In the data generated due to COVID-19 crisis, there is a fair share of misinformation that is intentionally or unintentionally disseminated [2,3,4,5]. The socio-cultural polarization added to the turbulence even in such difficult times when the global agenda is (and should be) to fight an invisible enemy (coronavirus) and protect lives [6]. This socio-cultural polarization has given rise to *echo chambers* wherein individuals are by choice exposed only to the opinions, beliefs, and attitudes that are consistent with their own, leading to reinforcement of their views and behavior. However, the degree of polarization on a topic tends to change with time, as the users' interest and focus shifts from one topic to the other. This change in the interest of users around various topics related to COVID-19 over time motivates us to study the dynamics of opinion formation in online discourse using a network theoretic approach. We use a simulation technique based on the principles of evolution of opinions in an adaptive random network [7] to evaluate

¹<https://www.businessinsider.in/slideshows/miscellaneous/a-comprehensive-timeline-of-the-new-coronavirus-pandemic-from-chinas-first-covid-19-case-to-the-present/slidelist/74721133.cms#slideid=74721140>

the formation of consensus with time among users who comment on a particular YouTube video. To the best of our knowledge, this work is one of the earliest attempts for investigating the temporal dynamics of consensus formation in online discourse related to COVID-19 pandemic. This has practical and policy implications that we highlight in the discussion section.

2 Background

Information and communication technology offer an opportunity to easily access any kind of information. The propensity to surround ourselves with similar others who share our perceptions intensifies feudal mindsets and produce "echo chambers" that damage the quality of online discourse on social media platforms [8]. This selective exposure to like-minded information fueled by recommendation algorithms on social media platforms leads to increased polarization which is a cause of growing concern as users don't get exposed to diverse opinions. Consider the causes and the consequences of polarization due to the events as listed below:

1. The treatment of daily wage migrant workers in India during the lockdowns has been widely discussed over various social media platforms. The government and the opposition were polarized in their views on how effectively the government handled the relief work for the migrant workers².
2. Brazil's president's incendiary response to the pandemic wherein he encouraged street protests, attacked officials who imposed quarantines, lambasted the media as well as the judiciary, and fired his health minister led to polarizing debates in Brazil³.

These events indicate that the socio-cultural polarization of opinions about information/misinformation related to sociocultural activities may lead to violence, killings, civil protests, and can influence the response of governments. This socio-cultural polarization during COVID-19 has prevented the governments and healthcare agencies around the world from making concentrated efforts towards an effective response to the crisis.

3 Methodology: Simulation on Networks

The study of opinion dynamics in adaptive networks has gained significant attention for its profound prospects in understanding human social behavior [9]. Prior research has used various methods to understand the issues related to the characterization of the collective social behavior of individuals such as cultural information dissemination and the dynamics of opinion formation [7,9,10,11]. We borrow from this stream of work to model the evolution of users' opinions as random binary encounters occur within a network and users update their opinion whenever the difference between the polarity of their opinions is below a certain threshold value of tolerance parameter.

The initial interaction between the users who comment on a YouTube video at a particular time corresponds to an Erdős–Rényi random network model i.e. it is considered as a random graph constructed by connecting a pair of nodes among the N users commenting on a YouTube video with probability p . In such a graph denoted by $G(N, p)$ the number of possible undirected

² <https://www.youtube.com/watch?v=K-51rP9drwU>

³ <https://carnegieendowment.org/2020/04/28/polarization-and-pandemic-pub-81638>

connections without self-loops will be ${}^N C_2$. Since each of these connections has a probability of occurrence p , the expected number of connections will be ${}^N C_2 \times p$, and on average, a node in the network will have $p \times N$ acquaintances which is the degree of a node.

For N comments posted by users on a YouTube video at a given time t , let us denote the respective polarity scores as s_i where $i = 1, 2, 3, 4 \dots N$. One user is randomly chosen at each step and its polarity score is compared with that of the rest of the users in the network. The absolute difference of their polarity scores is compared with the tolerance limit (TL) and if the difference is within the TL, mean of all the differences is considered for adjusting the polarity score of the chosen user as per the following logic:

$$\text{Polarity score of user } i = \text{Convergence parameter} \{ \text{Polarity score of user } i + \text{Convergence parameter} (\text{Sum of the differences of the polarity score of user } i \text{ with each user where the difference is less than the tolerance level}) \}$$

Using this methodology gives an insight into the number of iterations taken to form a consensus on different time-varying subsets of comments on a YouTube video. The delayed consensus formation indicates higher polarization and a faster consensus formation suggests less polarization in online discourse.

4 Experiments and Results

We focused on YouTube videos related to COVID-19 posted mainly by various news and healthcare agencies covering the spread of the pandemic. The extracted and cleaned data as described above was subjected to the sentiment analysis process to determine the polarity score of each of the comments. The data along with polarity score was subjected to the opinion formation simulation described in Section 3.

Appendix 1 shows a consolidated view of the experimental results on all the COVID-19 related YouTube videos that we considered for this study. These experimental results establish that the degree of polarization increases with time on diverse topics related to the COVID-19 pandemic. This lack of consensus could be attributed to the severity and sensitivity of the issues and the absence of an unequivocal governmental response to these issues. Overall, the results confirm that it becomes difficult to reach a consensus on diverse issues pertaining to COVID-19 over time due to increasing polarization and reinforcement of ideological homophily around these issues.

5 Discussion and Conclusion

One of the interesting observations from this study is that the consensus formation among polarized users takes longer as the public discourse's recency on a COVID-19 related video increases. The results of experiments indicate a gradual increase in iterations to form a consensus among the users' opinions over time. This observation indicates an increased polarization around COVID-19 among the users who posted comments on related videos as time passes.

Though the prior work on polarization around sociocultural issues such as controversial movies indicates a decrease in polarization with time [12] our findings suggest that polarization

during the COVID-19 pandemic has increased with time. This increasing polarization indicates long-lived echo chambers which are detrimental to society, especially during pandemics. One of the practical implications of this work is to minimize the societal divide and prevent the damage that social media-induced polarization can cause during crisis situations.

However, this work is not devoid of limitations. The first limitation is related to the sentiment analysis methodology employed in this work as Azure sentiment analysis methodology does not consider the contextual polarity of a word. The second limitation of this work is the gap of information not captured in social media. The third limitation is that we have considered the comments only in the English language to avoid any misinterpretation due to subtleties associated with other languages that authors were not familiar with.

This study could open several directions for future work. One would be to extend the scope of this to cover polarization around black lives matter and climate change. Secondly, the investigation and analysis of polarization using big data around COVID-19 related posts, images, videos on other social media platforms (such as Facebook and Twitter) might reveal interesting insights and help in ascertaining and theorizing the evolution of polarization during COVID-19 [13]. Finally, another future research direction could be to address issues pertaining to information asymmetries or spreading false information in social media created by large and strong players with specific political intent.

| Videos | Time | #Iterations for convergence |
|--|---------------------|-----------------------------|
| The lockdown: One month in Wuhan | Last week of March | 15 |
| | Last week of April | 16 |
| | First week of May | 33 |
| | Second week of May | 71 |
| Bill Gates makes a prediction about when coronavirus cases will peak | First week of June | 18 |
| | First week of July | 19 |
| | Second week of July | 20 |
| | Third week of July | 23 |
| | Fourth week of July | 25 |
| Coronavirus: How the deadly epidemic sparked a global emergency | First week of May | 31 |
| | First week of June | 31 |
| | First week of July | 33 |
| How wildlife trade is linked to coronavirus | First week of May | 12 |
| | First week of June | 12 |
| | First week of July | 16 |
| | Second week of July | No consensus |
| | Third week of July | No consensus |
| | Fourth week of July | No consensus |
| Corona Virus - Covid 19 Short Film End of the world? | First week of May | 15 |
| | First week of June | 17 |
| | First week of July | 17 |
| | Second week of July | 23 |

| | | |
|--|---------------------|--------------|
| | Fourth week of July | 30 |
| What is the coronavirus? | First week of March | 13 |
| | First week of April | 14 |
| | First week of May | 14 |
| | First week of June | 38 |
| | First week of July | 40 |
| The Coronavirus Explained & What You Should Do | First week of March | 14 |
| | First week of April | 14 |
| | First week of May | 15 |
| | First week of June | 20 |
| | First week of July | 23 |
| Locked down India struggles as workers flee cities | First week of March | 13 |
| | First week of April | 35 |
| | First week of May | No consensus |
| | Last week of May | No consensus |

Appendix 1: Experimental results of consensus formation on COVID-19 related YouTube videos

References

1. Gonzalez-Padilla, D. A.: Social media influence in the COVID-19 Pandemic. *Int. braz j urol.vol.46 sup1.1 Rio de Janeiro. ISSN 1677-6119* (2020).
2. Apuke, O. D., & Omar, B.: Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics and Informatics, 101475* (2020).
3. Hart, P. S., Chinn, S., & Soroka, S.: Politicization and Polarization in COVID-19 News Coverage. *Science Communication, 1075547020950735* (2020).
4. Hatcher, W.: A Failure of Political Communication Not a Failure of Bureaucracy: The Danger of Presidential Misinformation During the COVID-19 Pandemic. *The American Review of Public Administration, 50(6-7), 614-620* (2020).
5. Kumar, S., Kar, A. K., & Ilavarasan, P. V. (2021). Applications of text mining in services management: A systematic literature review. *International Journal of Information Management Data Insights, 1(1), 100008*
6. Qureshi, I., Bhatt, B., Gupta, S., & Tiwari, A. A.: Causes, Symptoms and Consequences of Social Media Induced Polarization (SMIP) (2020).
7. Kozma, B., & Barrat, A.: Consensus formation on adaptive networks. *Physical Review E, 77(1), 016102* (2008).
8. Garimella, K., Gionis, A., Parotsidis, N., & Tatti, N.: Balancing information exposure in social networks. *Advances in Neural Information Processing Systems, 4663-4671* (2017).
9. Ju, C., Wang, J., & Shi, J.: The intelligent method of public opinion polarisation modelling and simulation analysis based on multi-dimensional and multi-level evaluation. *Behaviour & Information Technology, 1-13* (2020).
10. Jin, C., Li, Y., & Jin, X.: Political opinion formation: Initial opinion distribution and individual heterogeneity of tolerance. *Physica A: Statistical Mechanics and its Applications, 467, 257-266* (2017).

11. Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G.: Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(01n04), 87-98 (2000).
12. Amendola, L., Marra, V., & Quartin, M.: The evolving perception of controversial movies. *Palgrave Communications*, 1(1), 1-9 (2015).
13. Kar, A. K., & Dwivedi, Y. K. (2020). Theory building with big data-driven research—Moving away from the “What” towards the “Why”. *International Journal of Information Management*, 54, 102205.