

Modeling geographical spread of COVID-19 in India using network-based approach

Ankush Kumar

Submitted to: JMIR Public Health and Surveillance
on: May 05, 2020

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 4
Supplementary Files..... 13
 Figures 14



Modeling geographical spread of COVID-19 in India using network-based approach

Ankush KumarPhD,

Corresponding Author:

Ankush KumarPhD,

Phone: +917986301011

Email: ankush.science@gmail.com

Abstract

Background: COVID-19 pandemic is a global concern, due to its high spreading and alarming fatality rate. Mathematical models can play a decisive role in mitigating the spread and predicting the growth of the epidemic. India is a large country, with a highly variable inter-state mobility, and dynamically varying infection cases in different locations; thus, the existing models, based solely on the aspects of growth rates, or generalized network concepts, may not provide desired predictions. The internal mobility of a country must be considered, for accurate prediction.

Objective: This study aims to propose a framework for predicting the geographical spread of COVID-19 based on human mobility, by incorporating migration and transport statistics. The motivation of the research is to identify the locations, which can be at higher level COVID -19 spread risk, during migrants transfer and transportation activities.

Methods: We use reported COVID-19 cases, census migration data, and monthly airline data of passengers.

Results: We discover that spreading depends on the spatial distribution of existing cases, human mobility patterns, and administrative decisions. In India, the mobility towards professional sites can surge incoming cases at Maharastra and Karnataka, while migration towards the native places can risk Uttar Pradesh and Bihar. We anticipate that the state Kerala, with one of the highest cases of COVID-19, may not receive significant incoming cases, while Karnataka and Haryana may receive the challenge of high incoming cases, with medium cases so far. Using airline passenger's data, we also estimate the number of potential incoming cases at various airports. The study predicts that the airports located in the region of north India are vulnerable, whereas in the northeast India and in some south India are relatively safe.

Conclusions: A model is developed for systematically understanding the effect of migration and transport on the spreading of COVID-19, and predetermining the hotspots on real time basis. Through the model, we identified the airports and states that are at higher level of COVID-19 risk. The study can guide policymakers in prior planning of transport and estimate the required medical and quarantine facilities to minimize the impact of COVID-19.

(JMIR Preprints 05/05/2020:19868)

DOI: <https://doi.org/10.2196/preprints.19868>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to all users.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in http://www.jmir.org/preprint/19868

Original Manuscript

Original Paper

Modeling geographical spread of COVID-19 in India using network-based approach

Ankush Kumar, Ph.D.

Corresponding Author: Ankush Kumar, Patiala Punjab, India, 147001

Email: ankush.science@gmail.com

ORCID: 0000-0001-6263-3143

Abstract

Background: COVID-19 pandemic is a global concern, due to its high spreading and alarming fatality rate. Mathematical models can play a decisive role in mitigating the spread and predicting the growth of the epidemic. India is a large country, with a highly variable inter-state mobility, and dynamically varying infection cases in different locations; thus, the existing models, based solely on the aspects of growth rates, or generalized network concepts, may not provide desired predictions. The internal mobility of a country must be considered, for accurate prediction.

Objective: This study aims to propose a framework for predicting the geographical spread of COVID-19 based on human mobility, by incorporating migration and transport statistics. The motivation of the research is to estimate the incoming infected passengers in different locations and identifying the locations, which can be at higher level of risk.

Methods: We use reported COVID-19 cases, census migration data, and monthly airline data of passengers.

Results: We discover that spreading depends on the spatial distribution of existing cases, human mobility patterns, and administrative decisions. In India, the mobility towards professional sites can surge incoming cases at Maharastra and Karnataka, while migration towards the native places can risk Uttar Pradesh and Bihar. We anticipate that the state Kerala, with one of the highest cases of COVID-19, may not receive significant incoming cases, while Karnataka and Haryana may receive the challenge of high incoming cases, with medium cases so far. Using airline passenger's data, we also estimate the number of potential incoming cases at various airports. The study predicts that the airports located in the region of north India are vulnerable, whereas in the northeast India and in some south India are relatively safe.

Conclusions: A model is developed for systematically understanding the effect of migration and transport on the spreading of COVID-19, and predetermining the hotspots on real time basis. Through the model we identified the airports and states that are at higher level of COVID-19 risk. The study can guide policymakers in prior planning of transport and estimate the required medical and quarantine facilities to minimize the impact of COVID-19.

Keywords: COVID-19 Modeling, COVID-19 in India, Disease spreading, Migration, Transport, COVID-19 Hotspots, Network Science, Coronavirus.

Introduction

COVID-19 is an actively spreading pandemic in the whole globe and is an unprecedented challenge for the healthcare, economy, and lifestyle of the community.[1-3] Countries are striving hard to mitigate the spread of COVID-19 by various strategies: banning gathering, closing schools, stopping transports, locking down cities, imposing curfews, and sealing locations, and still not able to effectively contain it. The need of the hour is to get location by location risk assessment so that timely preventive measures can be taken. Researchers have systematically studied various aspects related to COVID-19, such as the role of isolation of cases and their contacts, [4] obtaining infected population from the death counts, [5,6] and calculating optimum duration [7] and effectiveness of lockdown period. [8-9] The available research in this area is primarily on analyzing growth in the number of infectious cases in the local community. [7, 10, 11] These models mainly use non-linear fittings on time series of reported cases in a particular region and SIR derived models to estimate the time evolution of epidemics in that region.

Human mobility and transport also play critical roles in the spread of COVID-19, adding seeds of disease-transmission. However, there is a limited effort in the literature to model the impact of human mobility on the spread of COVID-19, particularly within a country. Chinazzi et al. and Wells et al. studied the importance of the travel ban of China and important border policies. [12-13] Paster et al. used a Long short-term memory (LSTM) based neural network to predict the risk category of a country. Pujari et al. attempted multicity model and assumed the fraction of population reaching a neighbor is inversely proportional to its degree connections. [14]

COVID-19 currently affects almost all developed and developing countries; India is one of the countries with significant COVID-19 cases. India is currently in stage 2 of epidemics, and strict plans and steps are required to prevent it from entering stage 3 or higher. India is a large country, with diverse cultures, languages, jobs, and educational opportunities, resulting in distinct and complex connection patterns between different locations. Thus, a generic mathematical network approach [15-17] such as a small world and scale-free model cannot be employed for disease-spread analysis for India; important geographical aspects of human-mobility should also be incorporated in a model. To tackle it, herein, we propose a network-based framework for modeling COVID-19 risk at different geographical locations by using migration and airflow based real data. The proposed model can be used for policymaking, regulating transport, and predicting future hotspots.

Methods

The model consists of dividing the space (say country) into various components, which can be states, districts, or cities depending on the available data. Consider there are N components with the population of individual components P_i and infected numbers I_i .

The probability, p_i of an infected person in the population P_i can be written as

$$p_i = \frac{I_i}{P_i} \quad (1)$$

Now, let, T_{ij} is the number of transported individuals from i to j component, then, a certain number of infected individuals would also be transferred from location i to j as follows:

$$C_{ij} = \frac{I_i}{P_i} T_{ij} \quad (2)$$

Thus, the total number of incoming cases at destination (j) from all connected components (i) can be written as

$$C_j = \sum_{i=1}^N C_{ij} \quad (3)$$

To model COVID-19 spread in India, here we use the states and union territories as components and their reported positive cases as the number of people infected. Ideally, the precise number of infected people in the population should be included, however, one can only have the information of the recorded cases. Mobility between two states relies on the intermixed community; migration may be a strong indicator of this. Thus, to analyze the potential inter-state mobility, inter-state migration data is obtained from the census 2011. [18] To obtain native to professional (work or education) place-based migration, only the data of male entries are considered to avoid significant marriage-related migration.

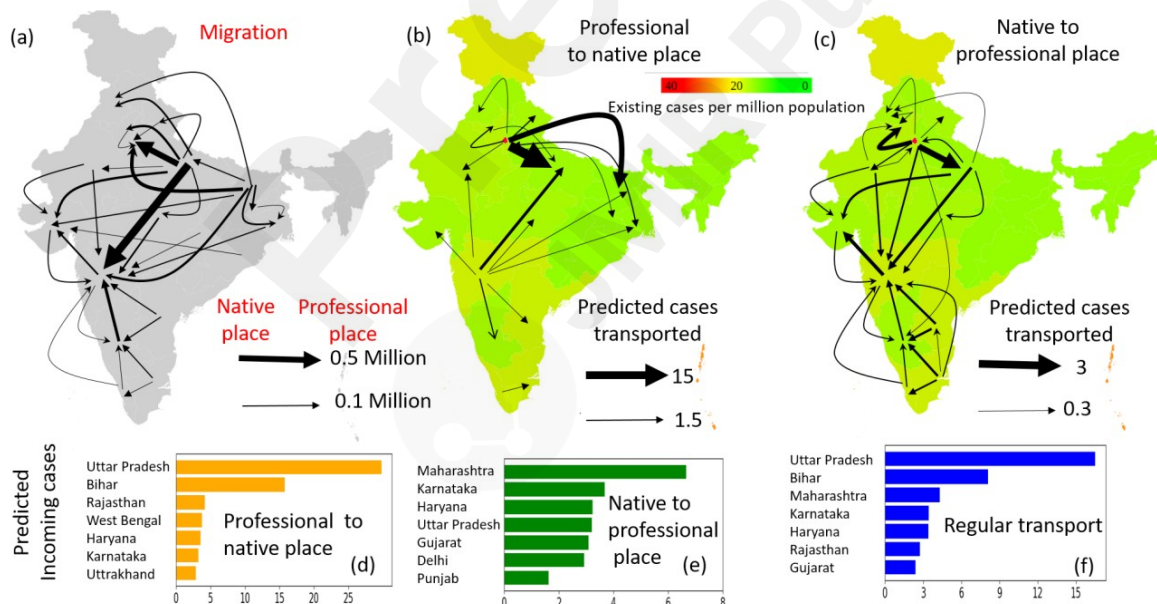
Next, we examine the spread of COVID-19 via transport through the airline's network. To obtain the mobility of passengers via airflow, we use publicly available monthly airline data of 2017, provided by the Ministry of Civil Aviation. [19]

Results

Figure 1a represents the migration map between different states with arrow-profile; here, the arrow width represents the migration number and the arrow direction is from native to the professional place. As seen from the arrow widths, the migration nature is unsymmetrical with high variability. The highest migration takes place between Uttar Pradesh to Maharashtra, Uttar Pradesh to Delhi, Bihar to Delhi and Karnataka to Maharashtra. In reference to COVID-19, different scenarios of mobility may arise: the mobility can be (a) towards the native site, (b) towards the professional place, or (c) nearly equal in both the directions. As an example, few days before the declaration of lockdown, the migrants might prefer to move towards their native places; on the closing of the lockdown period, the migrants may flow back to their professional sites and on the usual days the transport can be nearly equal in both the directions. The color map of figures 1b and 1c are depicting the total number of reported cases normalized with per million population. Throughout this article, we have considered the cases reported on April 10, 2020, and one can easily replace it with current data before the analysis and migration factor can also be reduced based on the situation. The case density (in per million population) is highly diverse, with a maximum at Delhi (40), Andaman and Nicobar Islands (29), Chandigarh (17), Jammu, Kashmir and Ladakh (14), Kerala (10), Tamil Nadu (10) and Maharashtra (10). Figure 1b represents the flow of people towards their native places (case a), which is a question of national concern as tweeted by Environment and Tourism Minister Aaditya Thackeray "Right from the day the trains have been shut down, the State had requested trains to run for 24 hours more so that migrant labor could go back home. CM Uddhav Thackeray ji raised this issue in the PM- CM Video Conf as well as requesting a roadmap for migrant labor to

reach home.” However, the preliminary calculations (Figure 1b and 1d) reveal that this transport may transfer a significant number of infected individuals to Uttar Pradesh and Bihar and can multiply further in the packed trains. It may be noted that, if the same scenario would occur in any other state having lower cases, instead of Maharashtra, then the problem might not be that much challenging. Now, we study case (b), i.e. the flow of people is from native to professional sites, which usually occurs at the ending of lockdown. Figure 1c and 1e show that Maharashtra, Karnataka, Haryana, Uttar Pradesh could receive maximum infected individuals. Strikingly, Karnataka, Haryana and Uttar Pradesh were not having high existing cases, yet could receive significant cases during the ending of lockdown. Further, Haryana and Uttar Pradesh may receive high cases from the hotspot Delhi, while Karnataka may receive infected cases from Andhra Pradesh, Telangana, Kerala, and Tamil Nadu. On the other hand, Kerala, which emerged as one of the first hotspots, with several international cases, may not receive significant national cases due to fewer migrants. Now we study case(c), importantly, it represents the natural connectedness between two locations, and in turn, decides the regular human mobility.. After a certain interval from lockdown, the flow would become equal from both sides due to regular movements; the transport, in this case is the average of both side migrant flow (Figure 1a). The analysis of regular flow shows that Uttar Pradesh, Bihar, Maharashtra, and Karnataka can be at more risk, once the transport is resumed. By the above discussion, we conclude that the spreading depends on existing cases, connectivity between locations and social scenarios, and is somewhat predictable by the proposed model.

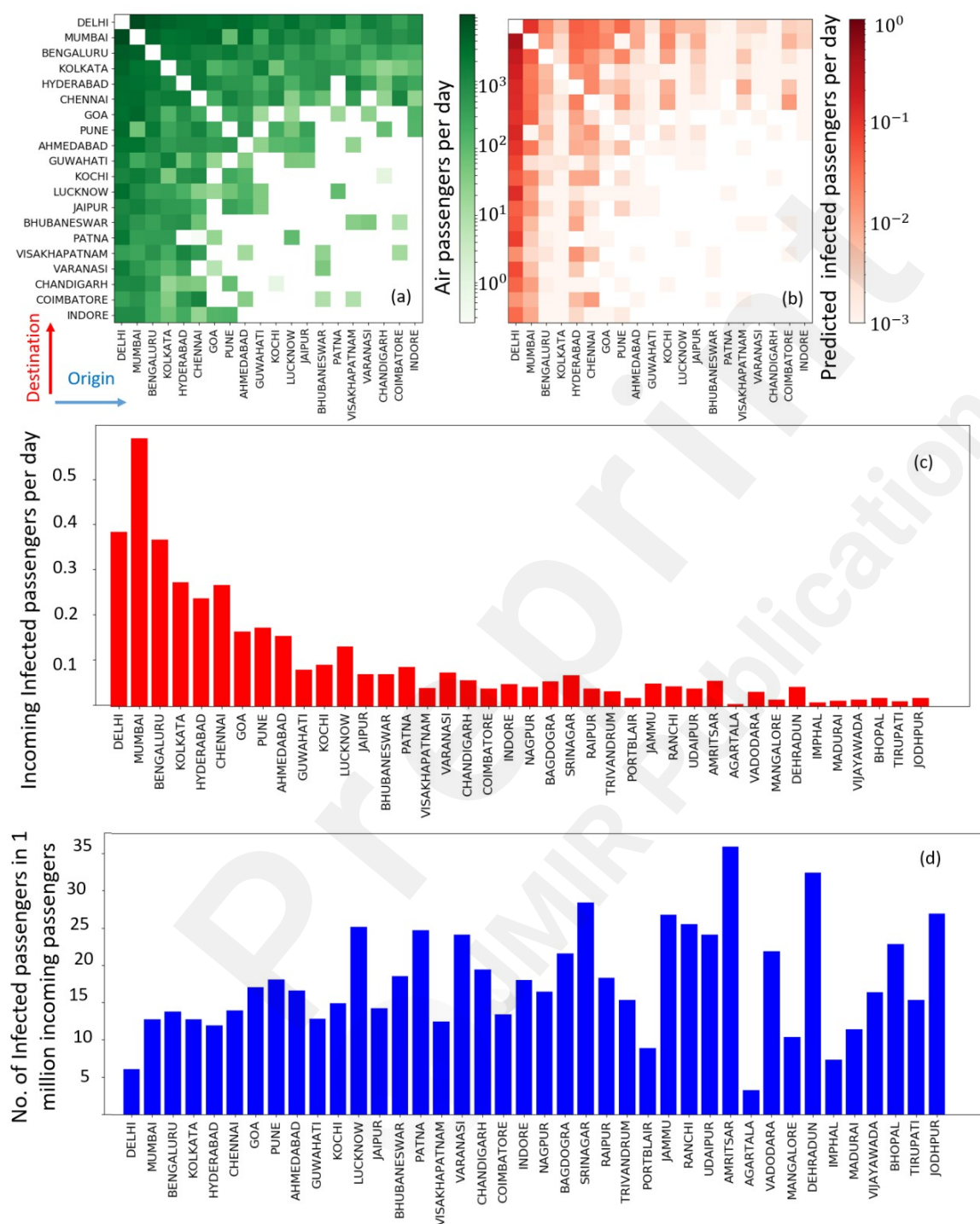
Figure 1: Arrow map showing migration profile between different states. The direction of the arrow is from native to the professional place, and the arrow width represents the migration population. (b-c) Color map of reported COVID-19 cases per million population (as per April, 10, 2020). (b) The arrows represent the estimated number of infected individuals from professional to native places and (c) native to the professional place. List of states with maximum incoming infected individuals during (d) professional to native place transport, (e) native place to professional place transport, and (f) during regular transport.



Next, we examine the spread of COVID-19 via transport using the airline's network as example. Figure 2a represents the pairwise passenger flow among different airports. The predicted flow of infected passengers is calculated based on our proposed model using Equation 3, incorporating the population and COVID-19 data of the state, where the particular airport is located. The predicted infected passengers Figure 2b, the cities with passengers from hotspots like Delhi, Mumbai can

receive significant cases. Figure 2c shows the predicted number of incoming infected passengers at 40 busiest airports. Mumbai, Delhi, and Bangalore would be receiving maximum incoming infected individuals, which is expected because of their high traffic. To normalize this effect, the total number of infected passengers is divided by the total number of passengers. The airports Amritsar, Dehradun, Srinagar, Jodhpur, Lucknow, Jammu, Patna, Varanasi can significantly receive a higher number of normalized infected individuals; it may be noted that all of these airports are situated in north India. Also, it is interesting to know that most of the North-East airports (Imphal, Agartala, and Guwhati) and a few south Indian airports (Vishakhapatnam and Madurai) are relatively at the lower level of danger, due to less connection with hotspots like Delhi and Mumbai. The work can help in identifying which airports could be shut down, which could remain operational and estimating the required number of quarantine and medical facilities.

Figure 2: (a) Pairwise number of passengers traveling between various airports. Only data of 20 airports (out of 75) are shown for visualization. (b) The predicted number of infected passengers traveling between different airports. (c) The predicted number of total incoming infected passengers at different airports (d) and normalized incoming infected passengers at different airports.



Discussion

Through the above results, we infer that the incoming cases based on national migration cases are not going well with the current cases. The passenger's air-transport analysis and regular transport based on migration map (case(c)), both predict that Uttar Pradesh, Bihar, and Karnataka could receive a significantly higher number of cases. Maharashtra with higher current cases already can receive additional incoming cases. Several unpredictable local factors and strictness of administration may

affect the active cases dynamically, thus the analysis should be performed on a day to day basis, once the transport is resumed. Note that the airport traffic flow is an indicator of connectedness between two locations, thus results by analyzing train or road network may also provide similar inferences. The developed model and platform can effectively be used for real-time risk monitoring of different locations. Therefore, future research will give optimal techniques for managing airflow, with minimum case-spread and maximum airflow. The model can also be used for border policy purposes, in an international situation that differs dynamically. For instance, hotspots shifted dramatically within the first four months of 2020, hotspots began from China in January, impacted Iran in February, Italy, and Europe in March, and the US in April. Not that, India's foreign mobility is very complex from one state to another. As a result, higher cases of COVID-19 in UAE can have a greater impact on Kerala, whereas the higher cases in Canada can have a greater impact on Punjab. The model may provide a systematic approach for controlling international flights in a rapidly evolving COVID chart. The model is generic and has the potential to study more magnified regions such as state or districts; mobile location data, bus, or train passenger's network data can be used for the purpose. Future work can investigate the time evolution aspect of COVID-19 in different locations.

Conclusions

The present work proposes a network-based model for predicting the spread of COVID-19, incorporating human mobility through knowledge on migration and air-transport. We found that migration towards native places may result in higher incoming cases in Uttar Pradesh and Bihar, while migration to the professional sites can surge incoming cases in Karnataka, Maharashtra, and Haryana, and daily flows are likely to endanger both sets of sites. Airports situated in North India are relatively at higher risks as compared to northeast and south. The model could be utilized to estimate the number of quarantine and medical facilities for incoming cases. Additionally, it will help to control the spread of COVID-19 by closing specific routes with a higher risk of pandemic spread.

References

1. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of travel medicine*.2020
2. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next?. *The Lancet*.2020.
3. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention.*Jama*.2020
4. Hellewell J, Abbott S, Gimma A, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *The Lancet Global Health*.2020.
5. Prakash MK, Kaushal S, Bhattacharya S, Chandran A, Kumar A, Ansumali S. A minimal and adaptive prediction strategy for critical resource planning in a pandemic. *medRxiv* doi:10.1101/2020.04.08.20057414.2020.
6. Gupta S, Shankar R. Estimating the number of COVID-19 infections in Indian hotspots using fatality data. *arXiv preprint arXiv:2004.04025*.2020.
7. Singh R, Adhikari R. Age-structured impact of social distancing on the COVID-19 epidemic in India. *arXiv preprint arXiv:2003.12055*.2020.
8. Kaur T, Sarkar S, Chowdhury S, Sinha SK, Jolly MK, Dutta PS. Anticipating the novel coronavirus disease (COVID-19) pandemic. *medRxiv* doi: 10.1101/2020.04.08.20057430.2020.
9. Sardar T, Nadim SS, Chattopadhyay J. Assessment of 21 days lockdown effect in some states and overall India: a predictive mathematical study on COVID-19 outbreak. *arXivpreprint arXiv:2004.03487*.2020.

10. Ranjan R. Predictions for COVID-19 outbreak in India using Epidemiological models. medRxiv doi:10.1101/2020.04.02.20051466.2020.
11. Bastos SB, Cajueiro DO. Modeling and forecasting the Covid-19 pandemic in Brazil. arXiv preprint arXiv:2003.14288.2020.
12. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science*. 2020; 368(6489): 395–400.
13. Wells CR, Sah P, Moghadas SM, et al. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proceedings of the National Academy of Sciences*. 2020; 117(13):7504–7509.
14. Pujari BS, Shekatkar SM. Multi-city modeling of epidemics using spatial networks: Application to 2019-nCoV (COVID-19) coronavirus in India. medRxiv doi:10.1101/2020.03.13.20035386.2020.
15. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. Epidemic processes in complex networks. *Reviews of modern physics*. 2015;87(3):925.
16. Newman ME. Spread of epidemic disease on networks. *Physical review E*. 2002;66(1):016128.
17. Castellano C, Pastor-Satorras R. Cumulative Merging Percolation and the epidemic transition of the Susceptible-Infected-Susceptible model in networks. *Physical Review X*. 2020;10(1):011070.
18. <http://censusindia.gov.in/2011census/populationenumeration.html>, note=Accessed: 2020-04-04.
19. <https://data.gov.in/resources/city-pair-wise-domestic-passenger-traffic-statistics-month-february-2017>, note = Accessed: 2020-04-04.

Supplementary Files

Figures