Exploring the Impact of Indian Revenues During COVID-19 Using Social Network Analysis

Suriyakrishnan Sathish, Aravinda Boovaraghavan, S. Hashwanth, and L. Jani Anbarasi

Abstract The second wave of the COVID-19 pandemic affected economy resulting in job loss for many throughout the world. The main objective is to analyze the Indian revenues before COVID-19, i.e., before December 2019 and the ongoing COVID-19 pandemic. A comparison is done between the revenues that are collected from various departments sectors for the years 2019–2020 and 2020–2021 to find out how the COVID-19 has affected the Indian economy as well as jobs for millions of people. Studying the network through social network analysis helps us to understand clearly how the Indian economy has changed during the COVID-19 period. The network is studied using an average path along with a weighted path to ascertain the small word characteristics. Centrality measures helps us to understand the characteristics of the dataset. The tools used to analyze them are designed using Gephi.

Keywords COVID-19 · Economy · Revenues · Indian economy · Social network analysis

1 Introduction

In recent times, we have heard a lot of increase in COVID-19 cases in India. This deadly disease not only affects human beings but also affects the whole economic system in India. Because of this pandemic, in the second wave of COVID-19, more than 15 million Indians have lost their employment. This has a significant impact on both the Indian economy and the lives of Indian citizens. The COVID-19 pandemic has had a significant economic impact in India where the growth declined to 3.1% in the fourth quarter of financial year 2020, according to the Ministry of Statistics. This decrease is primarily due to the impact of the Corona virus pandemic on the Indian economy. All the major sectors are affected due to the lockdowns as the factories, manufacturing companies, transportation, small-scale industry, construction, food

S. Sathish (✉) · A. Boovaraghavan · S. Hashwanth · L. J. Anbarasi
School of Computer Science and Engineering, Vellore Institute of Technology, Chennai, India
e-mail: suriyakrishnan.s2019@vitstudent.ac.in

A. Kumar et al. (eds.), *Proceedings of the International Conference on Cognitive and Intelligent Computing*, Cognitive Science and Technology,
https://doi.org/10.1007/978-981-19-2358-6_24

wastage increased due to affected supply chains, affecting small farmers and other departments. As a result, the daily wage workers, migrant workers, and household workers are severely hit in this pandemic.

Consumer demand, which was already slowing prior to the pandemic, has now collapsed, accounting for nearly 55% of the economy as household incomes and jobs have declined. As a result, it is important to understand the various reasons the economy has been affected. Similarly, it is seen that there is a major effect of COVID-19 on Indian revenue. There is no existing research for the Indian revenue data before and after COVID-19. So, an analysis was performed in this area and presented an outline between these two periods. Janssens et al. [7] investigated economic effects on low-income households in rural Kenya. An analysis using weekly financial household data was performed. World Development, 138, 105,280. According to the World Bank, the COVID-19 problem has the potential to force between 40 and 60 million people into extreme poverty, the majority of whom are in Sub-Saharan Africa [12]. According to another research, if income and consumption decrease by 20%, between 420 and 580 million people will be forced into poverty, reversing decades of decreasing poverty trends [11]. However, for most low-and middle-income nations, the immediate implications of the COVID-19 crisis at the household level are missing. The majority of the evidence comes from wealthy countries. In the United States, the lockdown policy reduced time spent outside the home [5], contributed to a significant drop in employment [9], and resulted in a significant drop in job vacancy postings [3].

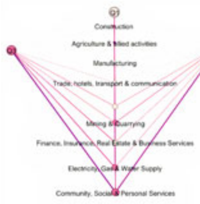
Social network analysis has always been a prominent field of research. With the increased popularity of online socializing, a great amount of data and information are gathered. The amount of information available is enormous, and Big Data [2, 6] has resulted in numerous issues. Furthermore, such information provides opportunities for researchers in a variety of sectors. Social network analysis is one such field. With social network analysis, you may learn a lot about how people interact and socialize with one another. Such social data can be used to track the user behavior of an individual as well as a community. Based on the metrics, an analysis is performed with the Indian data to study the effects of the economy before and after COVID-19.

The paper is structured as follows: Sect. 2 elaborates the data and the method used for analysis, Sect. 3 details the related work along with all the metrics used for evaluation. Section 4 details the proposed work followed by observations and discussions. Section 6 concludes the work with its future scope.

2 Data and Methods

In our research, two communities are formed out of which one community refers to the 3 quarters of 2019 (i.e.) Pre-COVID period and the other community refers to the 3 quarters of 2020. Data from 8 major departments is collected in order to form a network, and the same are segregated as per their year. This resulted in values for

Table 1 Data collected from various departments

List of departments analyzed	Generated dataset in network form
Agriculture and allied activities	
Finance, insurance, real estate and business services	
Trade, hotels, transport and communication	
Community, social and personal services	
Transport and communication [6, 11]	
Mining and quarrying	
Electricity, gas and water supply	
Mining and quarrying	

the eight departments in Pre-COVID-10 data as well as during COVID-19 data. The eight departments that are considered for this research are given in Table 1a.

From the main data, only the revenue data for the years 2019 and 2020 is collected. Then the data is segregated into separate csv files, forming the nodes, resulting in 2 datasets, 2019 and 2020. With the help of this data in the proposed work edges are manually formed, by connecting the 3 quarters of a year to each of the eight departments. The 3 quarters represent a specific month range in a year. Here, Quarter 1 denotes April to June, Quarter 2 denotes July to September, Quarter 3 denotes October to December, and finally, Quarter 4 denotes January to March. For the research process 2019–20 Quarter 1, 2, 3 is taken for pre-COVID-19 data and 2019–2020 Quarter 4 along with 2020–2021 Quarter 1 and 2 for During COVID-19 data. The data has been segregated as per the mentioned quarters and then it is linked from those quarters to the departments mentioned. The nodes representing the quarters are considered as the source, and then the target nodes are nodes that represent each department [1, 4, 8]. The graph type that is declared is an undirected graph and importing the datasets (both node and edge datasets) in such a form as shown in Table 1b.

3 Related Work

This section details the measure of modularity, Girvan-Newman Community Detection, average degree and weighted degree, which are used to measure the proposed work. To understand the economy in different quarters, degree, betweenness, and eigenvector centrality were performed. To comprehend the network's complexity and the adhesiveness of the nodes inside it, the clustering coefficient and average shortest path [10] were determined.

Modularity, which measures the structure of graphs, is used to assess the strength of a network's partition into modules. High modularity refers to dense connections between nodes within modules, and low modularity refers to weak connectivity among nodes and is commonly utilized in optimization techniques for identifying community. The modularity of the network model is computed as given in Eq. (1).

$$\mathrm{Mod} = \frac{1}{M}\sum_{i,j}\left[M_{ij} - \frac{M_i M_j}{M}\right]\delta_{c_i c_j} \tag{1}$$

M_i and M_j are the input and output of the nodes 'i' and 'j'. δ is the Kronecker's delta where $\delta_{c_i c_j}$= 1 refers that nodes 'i' and 'j' from same community or 0 otherwise.

By progressively eliminating edges from the original network, the Girvan-Newman algorithm identifies the communities. This identifies the edges that are most likely "between" communities, rather than measuring the vital edges. By deleting these edges, the groups are separated from one another, showing the network's underlying community structure.

Average degree refers to the average number of links present for each node. It refers to the number of links compared to the number of nodes in an undirected graph. The straightforward computation of the average degree is computed as follows (Eq. 2).

$$\text{Average Degree} = \frac{\text{Total Edge}}{\text{Total Nodes}} \tag{2}$$

The term "degree centrality" refers to how many alters an ego is linked with. The centrality of a node is determined by the number of neighbor nodes that are directly linked to it. It relates to the degree to which departments are involved in this study Eq. (3).

$$C_D(i) = \sum_{j=1}^{n} d_{ij} \tag{3}$$

Betweenness centrality assesses a node's ability to act as an intermediary in a network. The centrality of betweenness is expressed in Eq. (4) where g_{jk} is the number of the shortest distance links between j and k and $g_{jk}(i)$ is the number of times the path crosses these two nodes.

$$C_B(i) = \sum_{j<k}^{n} \frac{g_{jk}(i)}{g_{jk}} \tag{4}$$

The nodes which influences more in a network is identified using eigenvector or eigen centrality. It shows which node has had the most impact on the other nodes.

The eigenvector centrality (also known as eigen centrality) of a node in a network is a measure of its influence.

Closeness centrality is used to determine a node's reach in a network. The minimal hops required to move from one node to the other is the connectedness of a pair of nodes in undirected and unweighted networks. The closeness centrality of a node u is defined as follows in Eq. (5)

$$C_u\left(u\right) = \frac{n-1}{\sum\limits_{\forall v} d\left(u, v\right)} \tag{5}$$

4 Proposed Work

In this research, two datasets before and after COVID-19 are collected and compared across both networks to determine the impact of COVID-19 on our Indian revenue. From the results achieved, it can be seen that there is a strong effect of COVID-19 on many industries and departments. The nodes are the three quarters and the 8 departments. In both the segregated data, the departments are the same except for the quarters. The analysis and graph generation were performed using Gephi. The clustered column graph in Table 2a shows department vs revenues before the pandemic and during the pandemic. In Table 2b, it shows the revenue which was collected before the pandemic is slightly higher than in the COVID-19 period.

The social network model is generated for both the datasets to gain knowledge about the revenue status during the pre-COVID-19 period. The edge ranking, node color, and node size based on calculated values are defined. Edge ranking is based on weight , so the more the weight, the darker the edge is visible. Also, for node size, it is defined on the basis of weight. The minimum size is given as 1, and the maximum size of the node is given as 4. The node color is generated with respect to modularity. Once the average modularity is computed, the node color is displayed based on the calculated value. The higher the value of modularity, the darker the node color. With the help of these, the nodes and edges are defined so that a clear view of the network model is described. During the pre-COVID-19 period, Finance, Insurance, Business services hold the highest revenue values, followed by others. Table 3a, b shows the network model generated for pre-COVID-19 data.

Table 2 Revenue generated before and after COVID-19

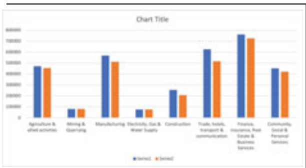
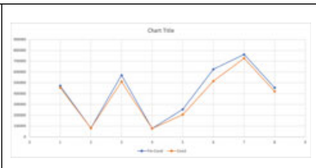
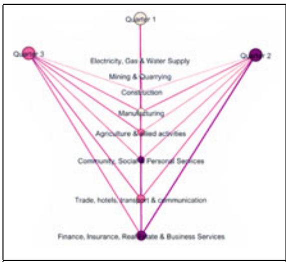
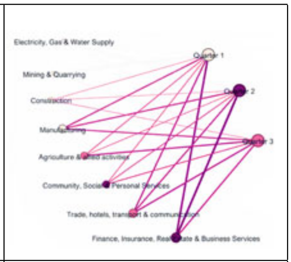
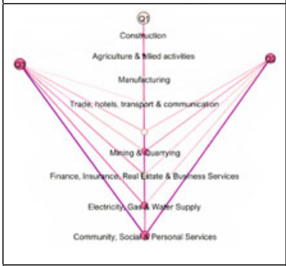
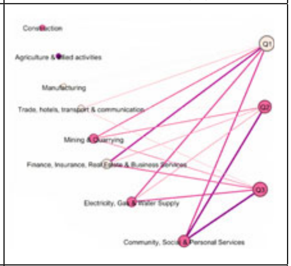
		
(a)	(b)	(c)

Table 3 Social network model for pre-COVID-19 and during COVID-19 data

	
Pre COVID-19 (a)	Pre COVID-19 (a)
	
During COVID-19(a)	During COVID-19(b)

The graph is generated from the dataset using Gephi, to attain knowledge about the revenue status during the COVID-19 period. The same procedure is followed for edge ranking, node color and node size which is based on calculated values. Edge ranking is based on weight, so the more the weight, the darker the edge is visible. Based on average modularity the node color is displayed based on the calculated value. The higher the value of modularity, the darker the node color. With the help of these, the nodes and edges can be defined so that it gives us a clear view of the network. It is seen that during the COVID period, Community, Social, and Personal services hold the highest revenue value, followed by Electricity, Gas and Water Supply. So if listed down in ascending order, like the higher revenue value node holds the first position followed by the next value, results in: Community, Social and Personal Services followed by Electricity, gas and Water Supply followed by Finance, Insurance, Real Estate and Business services followed by Mining and Quarrying followed by Trade, hotels, transport and communication, followed by Manufacturing, followed by Agriculture and Allied activities, and construction. Table 3a, b shows the network model generated during COVID-19 period. So, from the two formed graphs of both the pre-COVID-19 and during COVID-19 period. It can be easily seen that there is a tremendous effect on the revenue table. Top departments have been shattered during this COVID period. A comparison of the revenue data between pre-COVID-19 and during COVID-19 data is shown in Table 4.

Table 4 Metrics used to evaluate Pre COVID-19 and During COVID-19 data

Measures	Pre Covid-19	During Covid-19
Average degree	4.364	4.364
Average weighted degree	1,800,740.364	1,636,174.182
Eigen vector centrality	0.0038	0.0038
Graph density	0.436	0.436
Modularity	0.031	0.032
Betweenness centrality	9.333333	9.333333
Closeness centrality	0.833333	0.833333

From the table, the effect of COVID on the departments can be visualized where Finance sectors had a greater impact. Pre-covid, it was the hottest department with respect to revenue, but this covid had a greater effect on that department. It has been pushed down from first place to third place. At the same time, it can be seen that Community, Social, and Personal Services have been raised during this COVID period. It simply shows the fatal truth of how people were affected by the COVID wave, both financially and personally. If personal expenses increase, then financial problems arise. After this, it can be seen that Electricity, Gas and Water supply have the greatest increase. Due to quarantine, it can definitely said that more electricity will be needed and more water supply will be needed. So there has been massive growth in the revenue of these departments, which made them get top profit during this COVID period. Also, the trade, hotels, transportation, and communication sectors has experienced a significant drop in the lane. Mining and Quarrying may seem to have increased, but by checking with values, it is seen that there has been no change in their revenue status.

5 Observation and Discussion

Basic parameters like average degree and average weighted degree can be used to analyze the network's topology. The average degree of the network trained for the Pre-COVID-19 is 4.364, and the average weighted degree is 1,800,740.364. The average degree of the network constructed during COVID-19 is measured to be 4 and the average weighted degree of the same is measured as 1,406,244.727. The values clearly shows the downfall of the economy between these periods. The diameter can show us the largest shortest path between the pair of the nodes in the graph. The diameter values obtained for Pre-COVID-19 and during COVID-19 shows a good difference. Modularity refers to the network's partition. The modularity is obtained by randomizing the values with resolution as 1 and including the weights of the edge. The modularity obtained is 0.29 and 3 communities are formed for Pre-COVID-19 data. The modularity obtained for during COVID data is 0.037 with the same number of communities. The values clearly depicts that the economy fall during the period of COVID-19. Similarly, the communities were analyzed using the Girvan-Newman scheme, and only one community was formulated for both the datasets. The centrality and the density measures doesn't show much difference in the values obtained for the network model.

6 Conclusion

The COVID-19 pandemic has changed all our lifestyles. It is necessary to adapt ourselves to the "new normal". This research will help us in understanding the problems of a new recession and economic crisis in a variety of ways. A comprehensive social-economic development strategy that includes sector-by-sector plans to help businesses succeed with stable and long-term business models during the pandemic. The social network model was created for different sectors. These sectors are analyzed for both pre-COVID-19 and during COVID-19 data. This social network data is analyzed using various metrics in the proposed work. The modularity, average path, centrality measures obtained for both the duration shows a clear impact for both pre-COVID-19 and during COVID-19 data.

References

1. A. Agrawal, Sustainability of airlines in India with Covid-19: challenges ahead and possible way-outs. J. Rev. Pricing Manag. **20**(4), 457–472 (2021)
2. D. Dhanya, R. Vidhya, N. Modigari, A. Davamani, L. Anbarasi, in *Consistency and Reliability of Services in Cloud Environment*, pp. 747–751 (2017). https://doi.org/10.1109/CTCEEC.2017.8455064
3. E. Forsythe, L.B. Kahn, F. Lange, D. Wiczer, Labor demand in the time of Covid-19: evidence from vacancy postings and UI claims. J. Publ. Econ. **189**, 104238 (2020)
4. B. Goswami, R. Mandal, H.K. Nath, Covid-19 pandemic and economic performances of the states in India. Econ. Anal. Policy **69**, 461–479 (2021)
5. S. Gupta, T.D. Nguyen, F.L. Rojas, S. Raman, B. Lee, A. Bento, K.I. Simon, C. Wing, *Tracking public and private responses to the Covid-19 epidemic: evidence from state and local government actions*. Technical Report (National Bureau of Economic Research, 2020)
6. S.L. Kumar et al., Predictive analytics of Covid-19 pandemic: statistical modelling perspective. Walailak J. Sci. Technol. (WJST) **18**(16), 15583–14 (2021)
7. S. Kumar, V. Maheshwari, J. Prabhu, M. Prasanna, P. Jayalakshmi, P. Suganya, B.A. Malar, R. Jothikumar, Social economic impact of Covid-19 outbreak in India. Int. J. Pervasive Comput. Commun. (2020)
8. V. Kumar, Indian tourism industry and Covid-19: present scenario. J. Tour. Hosp. Educ. **10**, 179–185 (2020)
9. L. Montenovo, X. Jiang, F.L. Rojas, I.M. Schmutte, K.I. Simon, B.A. Weinberg, C. Wing, in *Determinants of Disparities in Covid-19 Job Losses*. Technical Report (National Bureau of Economic Research, 2020)

10. N. Prabhakar, L.J. Anbarasi, Exploration of the global air transport network using social network analysis. Soc. Netw. Anal. Min. **11**(1), 1–12 (2021)
11. A. Sumner, C. Hoy, E. Ortiz-Juarez, in *Estimates of the Impact of COVID-19 on Global Poverty*. No. 2020/43, WIDER Working Paper (2020)
12. World Bank, in *Monitoring Global Poverty* (2020)