A qualitative geographical information system interpretation of mobility and COVID-19 pandemic intersection in Uttar Pradesh, India

Alok Tiwari, Mohammed Aljoufie
Department of Urban and Regional Planning, Faculty of Architecture and Planning, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia

Abstract
Mobility of individuals and their physical social networks are the root causes for the spread of current coronavirus pandemic. We propose here a method of visualizing the spatial and chronological aspects of the spread of this virus based on geographical information systems (GIS) and Gephi graphs. For this approach we used qualitative data from newspaper reports and prepared layouts varying from macro to micro scales that show that this approach can enrich traditional GIS approaches, thereby assisting mobility planners and policymakers.

Introduction
The coronavirus disease emerging in China in 2019 (COVID-19) has rapidly and dramatically engulfed the entire world which led the World Health Organization (WHO) to declare it a pandemic on 11th March 2020 (WHO, 2020a). COVID-19 has altered the lives of people around the globe and forced governments into taking extraordinary actions including social distancing, wearing face masks and temporal lockdowns in efforts to control and curtail the spread of the disease (De Vos, 2020). These countermeasures have transmuted mobility of individuals and their mindsets (de Haas et al., 2020). However, there is a lack of studies aimed at understanding how exactly mobility affects COVID-19 transmission at the micro and macro levels. Given the main challenges of limited documentation and data, there are reasons to believe that geographical information systems (GIS) might be assisted by Gephi (https://gephi.org), a relatively new software that enables investigation of global connectivity and examination of network distributions and traffic.

The first Indian COVID-19 patient, a medical student who had studied in Wuhan, China and returned back, was identified in Thrissur, Kerala State in southern India on 30 January 2020 (Raghunath, 2020). However, the first case in Uttar Pradesh State occurred in its Ghaziabad District on 4 March 2020 and concerned a person who had been abroad in the recent past (Livemint, 2020). In the next several weeks, the disease spread all over India. This paper attempts to document and visualize the mobility of individual carriers of the coronavirus and discusses how to delay its spread by incorporating a new type of data into crisis planning. Gao (2008, 2020) has outlined how complementary data and tools can augment traditional GIS. We demonstrate here a promising approach that should be useful for mobility planners, who use GIS as a comprehensive tool for evidence-based policymaking. Firstly, we discuss the application of GIS in tracking and reversal of infectious diseases. Secondly, we review the evolution and amalgamation of qualitative GIS in several fields of society and economy with special reference to its potential to support understanding the spread of the disease at the intersection of mobility and health crisis.

Geographical information system and infectious diseases
GIS tools can visualize specific relationships, such as indicating the originating point of a disease, the risk factors involved and the sources of contamination. The tradition of mapping diseases has helped policymakers and planners in finding remedial interventions to maximize health and wellbeing of people (Rushton, 2003; Gao et al., 2008). Public health professionals have for many years seen traditional mapping as a part and parcel of their research and policymaking. For example, Dr John Snow, working in London, UK altered the world of epidemiology forever by his revolutionary cholera map in 1854 pointing out the source of the infection by establishing a robust link between the outbreak of cholera cases and the sources of drinking water supply (Brody, 2003). Likewise, mapping of the Italian plague epidemic in 1964 visualized the association between infectious diseases and place (Koch, 2005). In fact, the gradual evolution of GIS has further developed the relationship into making GIS a vital means to trace and fight with contamination (Kamel Boulos and Geraghty, 2020).

In the present scenario, GIS professionals have created several solutions to manage COVID-19 outbreaks including the map-based dashboard of Johns Hopkins University Center for Systems Science and Engineering (2020), HealthMap (2020) from Boston Children’s Hospital, USA and information from WHO (2020b).
These interactive dashboards present location-based information on COVID-19 cases in near real-time based on official sources; however HealthMap also reports informal data sources including social media, Google News and expert-managed estimates (Hossain and Househ, 2016). Early on, social distancing was suggested as a countermeasure to COVID-19 pandemic (Chen et al., 2020b) because close social contacts boost the spread of the virus causing this disease, while China generated Big Geospatial Data based on the movement of people through a geo-social application aimed at ‘close contact detection’. The application was launched to record real-time mobility of people making sure whether or not the client has had any public contact with a confirmed or suspected COVID-19 patients in the recent past (BBC, 2020). Similarly, the Environmental Systems Research Institute (ESRI) in the U.S. developed a proximity-tracing tool to locate closeness of events and monitor time and location of individuals present within a specified immediacy to another person (ESRI, 2020). These applications have proven beneficial, but only to some extent since their final outcomes have not been made widely available until very recently. Regardless of these and other initiatives, it must be accepted that surveillance of the COVID-19 pandemic has been limited. Gardener in the ‘Story Map team’ at ESRI (2020) contends that ‘it is especially challenging to collect good data at a fine spatial resolution, which is what most people want to know, and without having travel data in real time that captures these altered mobility patterns, it is hard to assess what the geographic risk profile will look like moving forward’.

**Qualitative geographical information system**

Qualitative GIS combines qualitative elements with spatial traits to track an option comprising an exclusive insight that illustrates how the spatial framework is built within the community. It operates through a mixed-method approach to perceive a variety of depictions of a spatial framework placing it with a perpetual and modern perspective by recognizing multiple skills (Teixeira, 2018). Recent advancements in geo-visualization software and web-based mapping have eliminated some of the barriers posed by traditional GIS approaches that prohibit the incorporation of several types of qualitative analysis (Taylor et al., 2020). New GIS techniques have enabled analysis and inclusion of cultural attributes, perceptions, opinions and attitudes concerned with engagement, participation and understanding (Egiebor and Foster, 2019; Mukherjee, 2019). For instance, ESRI (Redlands, CA, USA) has introduced a tool called ‘Story Map’ by which researchers can generate effective and fascinating tales by harmonizing interactive maps, texts and other multimedia materials including mental maps, sketches, photographs, audio and video (ESRI, 2019). A study by Garnett and Kanaroglou (2016) contends that qualitative GIS is quite an innovative approach in research methodology aimed at analysis and visualization of qualitative data in a distinct geography. Methodological developments in recent years have proved that GIS is more than a quantitative tool (ESRI, 2019); applications have led to innovative exploration of what has been called ‘affective geo-visualization’, which is a methodology that deals with the intersection of human emotions, stress, logical reasoning and the geographical context (Kemper, 2014). Additionally, qualitative GIS has been shown to be useful for harnessing non-numerical information examined in non-mathematical setups (Dennis, 2006). In contrast to conventional GIS that requires choosing spatial layers to disseminate a single straightforward narrative, qualitative GIS attempts to represent intense contextual information of space in a range of formats (Monmonier, 2005). Consequently, qualitative GIS has the potential to enrich policymaking by appending empirical intellectual data offering the addition of a kind of organic spatial layer. Initial studies in the field prompts the qualitative approach that in fact is exceedingly ingrained into the type of evidence gathered through qualitative sources including in-depth interviews, field diary, observations, sketches and photographs (Jung and Elwood, 2010). Qualitative GIS can also link analyses based on discourse or content analyses (Cope and Elwood, 2009). In later years researchers have explored possibilities in qualitative geo-visualization by scrutinizing geo-tagged social media data. For instance, Jung (Jung, 2015) proposed the idea of ‘code cloud’, which can identify key topics produced by digital qualitative research, while Andris (2016) found that location-based data from the social media networks offer a novel way of studying human connections that shape spaces through mobility, communication and sustaining relations. Moreover, we note that qualitative GIS is augmenting both evidence and analysis resulting from traditional GIS approaches. We feel that this approach is crucial in relation to the COVID-19 pandemic which is believed to spread by social contacts and mobility.

Based on the information given above, we decided to attempt a spatial demonstration of the role of mobility and social networks in the spread of COVID-19 derived from qualitative GIS. Although we acknowledge that geospatial visualization on the interaction between infectious diseases and space is growing, we argue that capturing data on mobility patterns during the COVID-19 pandemic requires several new sources of data; this is a difficulty, which can be overcome by newspaper-based narratives on mobility data at the local level. The overall aim of this research was to deliver qualitative evidence and methodology of geo-visualization for the mobility planners in Uttar Pradesh State of India during the COVID-19 pandemic.

**Materials and methods**

**Study location and context**

With 231.5 million inhabitants (Registrar General and Census Commissioner of India, 2020), Uttar Pradesh State is one of the most populated regions in the world comprising 75 districts, which serve as administrative divisions. Known for its geographical variation and cultural diversity, this state is situated in the northern part of India and well connected via international flights with a number of destinations. Major airports exist in the districts of Lucknow, Varanasi, Gorakhpur, Kanpur and Allahabad (Sharma, 2019). These districts are also well linked with metropolitan cities in the country through rail and road networks. Additionally, eight districts in the western part of Uttar Pradesh (Meerut, Ghazibad, Gautam Budh Nagar, Bulandshahr,Baghat, Harpur, Shamli and Muzzafarnagar) are part of the National Capital Region (NCRPB, 2017) of India, which all enjoy high transportation connectivity through New Delhi, the capital city. We note that the mobility pattern changed dramatically during the COVID-19 lockdown when huge numbers of internal and external migrants made it a priority to return back their native places (Press Trust of India, 2020a). According to the Community Mobility Report prepared by Google (2020), the pandemic reduced the mobility of people in Uttar...
Pradesh drastically in several sectors; by 58% for retail and recreational services; by 46% for parks; and by 40% for public transport.

Qualitative design

We used newspaper reports and visualized its spread by Gephi software (https://gephi.org).

A recent study, Chen et al. (2020a) adopted a contents analysis based on newspaper narratives to investigate the impact of COVID-19 on tourism. Following this initiative, we retrieved a total of 945 newspaper articles from 5 March 2020 to 15 June 2020. Afterwards we sketched each newspaper narrative forming a kind of map that is not to scale (Taylor et al., 2020). We added the approximate latitudes and longitudes for the originating and destination locations for the international, interstate, and local movements of COVID-19 patients, and the spread of infection to their contacts. However, although we observed that Indian newspapers were reporting the COVID-19 crisis widely, a proper spatial context was often missing. Hence, although we also used Google to collect news related to the mobility of people during the pandemic. The keywords used for the analysis were the following: ‘first coronavirus case’; ‘travel history of coronavirus patients’; and ‘coronavirus patient contact’. This process was repeated for each of the 75 districts of Uttar Pradesh.

We used the Gephi (Cherven, 2013) approach to visualize the data collected based on graph theory (Novović et al., 2020). The graphs are usually created with nodes (or vertices) and edges, where the former are group(s) of objects depicting units in a dataset, and the latter links between the nodes. In this paper we used individuals (marked by their background or profession) and cities as the nodes (Cherven, 2013). Gephi is an open-source software for network visualization and analysis that not only reveals patterns and trends but also examines the interconnections among diverse physical, informational, biological, cognitive, and social networks. Gephi combines built-in functionalities and flexible architecture and displays large graphs in real-time accelerating exploration. The network visualization adds to statistical analysis in a complementary way to discover, extract and classify patterns in network structures.

In this research we used two types of Gephi layouts in combination: Fruchterman-Reingold-layout that is force-directed algorithms emphasizing the complementarities of a directed network graph with a high level of simplicity, flexibility, and good quality results; and Geo-layout, which focuses on geographical repartition. Geo-layout as done by Jacomy (2019) in combination with the Fruchterman-Reingold approach carried out with Gephi software (2017) was preferred for geo-visualization, as this methodology has previously proved apt in aggregating mobility data in strategies to delay the spread of COVID-19 (Buckee et al., 2020).

We firmly believe that combination of both these layouts provide better results due to their unique features. It is noteworthy that the Geo-layout algorithm uses geocoded attributes (latitude and longitudes) and standard projections like Mercator. The Fruchterman-Reingold layout, on the other hand, is a kind of forced directed graph (named after its developers) that positions edges in a manner which tries to keep minimal crossing and equal length of edges as much as possible (Fruchterman and Reingold, 1991). We opted to use both above mentioned layouts in combination because the spread of COVID-19 involves spatial mobility that can be easily visualised with the Geo-layout (Gephi, 2017). However, crossing of multiple edges makes it messy, also this layout does not allow further manipulation of graphs for better aesthetic visualisation. So, to supplement the Geo-layout we used the Fruchterman-Reingold layout which offers simple, stylist, efficient and conceptually intuitive graphs.

Quantum GIS or QGIS (Graser, 2016) was deployed to generate latitudes and longitudes data for the above locations while simple choropleth maps were generated to understand the spread of COVID-19 across districts as given by official information by the Ministry of Health and Family Welfare.

Results

Newspaper reports confirm that the initial spread of the COVID-19 pandemic in Uttar Pradesh was low and only 10 districts had positive patients during the first quarter of the year; however by 13 May 2020, there were patients in all of its 75 districts (Chaturvedi, 2020). We briefly describe six cases to show how the mobility of infected travelers contributed to the spread of this infection to others who did not travel.

Case #1

A shoe factory owner returned to his home in Agra District from an international exhibition in Milan, Italy in the first week of March 2020, later to find out that he had become positive for COVID-19. His wife tested positive two days after his return and during the following days, four other family members also tested positive (Agarawal, 2020). Next, one of his factory workers developed symptoms and was subsequently found positive for the virus. Thus, within a week the coronavirus infected seven people with no travel history from just one person, which clearly illustrates that the role of the social contact network (Figure 1).

Case #2

In the second week of March 2020, a lady physician from Toronto, Canada visited her relatives in Lucknow. She was asymptomatic and passed the thermal screening at the airport without problems but developed symptoms after a few days and subsequently tested positive for COVID-19. She was then kept in isolation at the King George Medical University (KGMU). A few days later, however, her father-in-law, mother-in-law, a cousin, son and a medical doctor at KGMU were detected as COVID-19 positive (Press Trust of India, 2020b). In this case, one traveller acted as a carrier of coronavirus for five stationary people (Figure 2).

Case #3

In the last week of April 2020, wholesaler of pharmaceutical products travelled back from Europe to Varansi District and was found to be positive for COVID-19 after a few days at home (Srivastava, 2020). Within the next 2 weeks the COVID-19 infection jumped to 11 more individuals including 4 family members, 4 workers, 2 nearby pharmacy owners and 1 customer (Figure 3).

Case #4

In the last week of March 2020, a businessman travelled approximately 1200 km from Amravati, Maharashtra Province (in the middle of India) to his family in Meerut District by train and car. Within a week he himself, with his wife, two brothers-in-law, and a sister-in-law, four of whom without a travel history (Bhatia, 2020), were infected (Figure 4).
Figure 1. Spread of COVID-19 to Agra District, Uttar Pradesh by an international traveller. Case #1 - Agra: A) Fruchterman-Reingold layout; B) Geo-layout.

Figure 2. Spread of COVID-19 infection to Lucknow District, Uttar Pradesh by an international traveller. Case #2 - Lucknow: A) Fruchterman-Reingold layout; B) Geo-layout.
Case #5
In the third week of May 2020, a migrant worker returned to Azamgarh District by truck from Mumbai, Maharashtra State. He fell ill and did not survive. After his death at the Government Medical College & Super Facility Hospital Chakrpanpur, it was revealed that he had been infected. Within a few days, a total of 4 individuals were detected to also have the infection (Press Trust of India, 2020b): the driver who had earlier transported the deceased, his wife, the ambulance driver and a laboratory technician at the medical facility (Figures 5 and 6).

Next, we chronologically visualised the transmission of COVID-19 from major cities in Canada, Europe, Middle East and...
the U.S., which were directly connected to the place of origin of this pandemic in Uttar Pradesh; and also with the cities that were indirectly connected (Dubai and Istanbul). With respect to how the infection entered Uttar Pradesh, it was clear it first got footholds in 13 key districts. Taking the chronological aspect into account, the majority COVID-19 cases were reported in the second, third and fourth week of March 2020, including Agra (from Milan) and Ghaziabad (from Tehran) on 5 March 2020, Lucknow (from Toronto) and Noida (from Rome) on 12 March 2020, Kheri (from Istanbul) and Moradabad (from Paris) on 19 March 2020, Kanpur (From San Francisco) on 24 March 2020, while the districts Azamgarh, Shamli, Baghpat and Jaunpur were not infected until April 2020 (Figure 7). As can be seen in Figure 8, the virus spread rapidly after 30th March 2020. Soon afterwards, it reached all dis-

Figure 5. Spread of COVID-19 to Azamgarh District, Uttar Pradesh by a migrant worker. Case #5 - Azamgarh: A) Fruchterman-Reingold layout; B) Geo-layout.

Figure 6. Index cases in Uttar Pradesh districts caused by international travel (Fruchterman-Reingold layout).

Figure 7. Index cases and their datelines in Uttar Pradesh districts caused by international travel (Geo-layout).
Discussion

We assembled newspaper reports regarding 224 COVID-19 cases and visualized the network of 120 travellers displaying international, national, and local mobility histories, and their 104 social contacts without any travel history in Fruchterman-Reingold layout and in Geo-layout. Although not surprising, our displays confirmed the primary spread of COVID-19 between close contacts, e.g., family members, other relatives, work associates and common social contacts.

The increased transmission seen after the end of March from the original 13 districts, now having become COVID-19 hotspots, can be explained by inter-district mobility as individuals commonly had to travel short distances due to their business, service, or medical needs. This is displayed by a Geo-layout in Figures 10 and 11. Importantly, while the first Indian COVID-19 patient transferred from China to southern India in January 2020 (Raghunath, 2020), the main spread in India did not occur until March and April. Indeed, the Gephy illustrations shown here support the pos-
sibility that the wave of COVID-19 infections that eventually engulfed India might have originated through travel from Europe, Middle East or Canada rather than from China.

Our work displays that expediency of Gephi visualization as a qualitative GIS tool for the interpretation of mobility and COVID-19 transmission. Through the example of Uttar Pradesh, this research also reinforces findings of a previous study that suggests hypermobility as a cause of COVID-19 outbreak (Wu et al., 2020). Especially, the partition and rank features of Gephi demonstrate that Delhi and Mumbai are important external hotspots, and Agra an internal one in Uttar Pradesh, which is shown with a larger number of labels and unique colours to better visualize how the edges connect the nodes (Figures 12 and 13). The partition feature allows us to find communities in a complex network with unique colour labels for individual nodes and edges, while the Rank feature allows us to keep nodes and their respective labels at different sizes as shown by the different Gephi graphs (Figures 6-13). For instance, the size of the Delhi node in Figure 9 is larger than that of Mumbai because of its higher rank, which is due to the larger number of connections in the network. Similarly, the edges and the nodes are presented in different colours to help distinguish the connections. Likewise, the labels indicating Mumbai and Delhi in Figure 12 are larger in size compared to the locations they are connected with. Moreover, it is also possible to vary the width of

Figure 9. Index cases in Uttar Pradesh Districts caused by inland travel (Fruchterman-Reingold).
edges based on their assigned weights (if any); however we did not use this feature here. In addition the colour, size and labels for the nodes, and the colour and width of edges can also be changed manually as per the aesthetic requirements required.

Apparently, the argument behind using two different types of layouts is not noticeable in Figures 1-5 because they used very few data points. However, the difference is clearly seen in Figures 6-13 where the Geo-layout visualisations (Figures 7, 8, 11 and 13) are rather messy (though pertinent due to the need to show all spatial attributes), while the Fruchterman-Reingold layouts (Figures 6, 9, 10 and 12) are elegant, conceptually-intuitive and sober. Moreover, the Geo-layout approach does not allow further manipulation in visualisation, while the Fruchterman-Reingold layout does. Together, this information supports remedial actions to tackle mobility challenges during and after COVID-19 outbreaks. The qualitative data based on geo-visualization shown here can help planning specific lockdowns and other countermeasures avoiding blanket treatment (de Haas et al., 2020). For instance, areas characterized as harbouring international travellers and emigrants at some point in time might potentially be identified as a COVID-19 hotspot, and mobility restrictions enforced without delay. In this way, Gephi could be employed as a qualitative GIS tool, thereby informing and assisting mobility planners and policymakers to take effective decisions when faced with an epidemic or a pandemic.

Figure 10. Index cases of COVID-19 in some of the Uttar Pradesh Districts caused by inter-district travel (Fruchterman-Reingold).

Figure 11. Chronology of the emergence of index COVID-19 cases in Uttar Pradesh Districts (Geo-layout).
Conclusions

This research contributes to existing research of epidemic-induced geo-visualization by captivating and displaying qualitative information on mobility and physical social networks in connection with COVID-19 outbreaks. We have proposed a method of qualitative visualisation of spatial layers with the help of Gephi software which is based on graph theory and illustrated the same in case of COVID-19 spread in Uttar Pradesh in India due to varying mobility of individuals. Next, we show that even qualitative contents based on newspaper reports and narratives can enrich traditional GIS by visualizing otherwise invisible social networks. We also recommend future research using this method in other countries to understand the intersection between mobility and spread of a pandemic.

Figure 12. Visualizing connections binding together COVID-19 cases in Uttar Pradesh (Fruchterman-Reingold).
Figure 13. Visualizing COVID-19 cases in Uttar Pradesh (Geo-layout).

References


ed-pharmacist-doing-home-deliveries-causes-surge-in-COVID-cases/412390/

StoryMap team ESRI, 2020. Mapping the novel coronavirus pandemic, ESRI. Available from: https://storymaps.arcgis.com/stories/4fde0d03d3a34aa485de1fb0d2650ee0


