Strategizing COVID-19 Lockdowns Using Mobility Patterns

Olga Buchel[1], Anton Ninkov[2], Denise Cathel[1], Yaneer Bar-Yam[1], Leila Hedayatifar[1]

[1] New England Complex Systems Institute, 277 Broadway St, Cambridge, MA, USA
[2] Faculty of Information and Media Studies, University of Western Ontario, Canada

During the COVID-19 pandemic, countries and states/provinces have tried to keep their territories safe by isolating themselves from others by limiting non-essential travel and imposing mandatory quarantines for travelers. While large-scale quarantine has been the most successful short-term policy, it is unsustainable over long periods as it exerts enormous economic costs. Countries which have been able to partially control the spread of COVID-19 are thinking about policies to reopen businesses. However, pandemic experts strongly warn against reopening too soon. Thus, it is urgent to consider a flexible quarantine policy that balances these demands. Here, we have designed a multi-level quarantine process based on the mobility patterns of individuals and the severity of COVID-19 contagion in different areas. By identifying the natural boundaries of social mobility, policy makers can impose travel restrictions that are minimally disruptive of social and economic activity. The dynamics of social fragmentation during the COVID-19 outbreak are analyzed by applying the Louvain method with modularity optimization to the weekly mobility networks. In a multi-scale community detection process, using the locations of confirmed cases, natural break points as well as high risk areas for contagion are identified. At the smaller scales, for communities with a higher number of confirmed cases, contact tracing and associated quarantine policies is increasingly important and can be informed by the community structure.

Introduction

The emergence and global spread of the 2019 novel coronavirus (SARS-CoV-2 or COVID-19) has resulted in a global health emergency. With a high level of observed contagiousness [1] and a lack of proven medical treatment, the situation is becoming increasingly dire as the virus moves across the globe. Public health stakeholders are racing to find adequate methods for intervention as the outbreak spreads [2,3]. It is challenging to determine where the next outbreak will be and how to prevent and control it. Analyzing data about positive tests and location of current patients plays a critical role in public health agencies response [4]. In most cases, quarantine policies and data related to the COVID-19 outbreak are based on arbitrary borders such as state or county boundary lines [5-10]. While these boundaries may serve constituents well in meeting certain social needs of their communities (e.g. infrastructure, taxes), this is not the most effective way to analyze data for anticipating disease outbreaks.

For the purposes of examining the spread of COVID-19 in the US, mobility patterns can be characterized in three overarching concepts: short distance (grocery shopping, walking), medium distance (travel for work or fun), and long distance (travel to other cities for vacation, visiting families). Travel can be thought of as occurring in “bubbles” of progressively larger geographical scales. National travel bubbles includes the common movement of individuals traveling far from one region of the country to another. This type of mobility pattern was quickly identified as risky and attempts to limit this type of travel within the country were put in place. For example, in the very beginning of the North American outbreak in March, 2020, a group of university undergraduate students from around the country gathered on Florida beaches for the traditional spring break. During this gathering, local transmission of COVID-19 was detected and the spread of the disease to various other regions of the country occurred [11]. Soon after, most universities were closed and airline travel was reduced. Local travel bubbles, on the other hand, include regions with close proximity where there are more frequent mobility patterns. Local bubbles are prevalent in places such as the Northeast Megalopolis [12,13], where there are numerous cities and communities all continuously connected to one another. In this region, many individuals may live in one city/state (e.g. Philadelphia), work in another (New York City), and vacation in another (New Jersey coast). While these regions are separated by multiple administrative boundaries, they could still be considered to be in the same bubble.

In this paper, we utilize a well-known community detection algorithm to analyze the dynamics of movement behaviors and study the fragmentation patterns from the weekly aggregated mobility networks in the US. The recent availability of large-scale datasets derived from bank transaction records, landline, mobile and social media has greatly improved our ability to study social systems [15-17]. Geo-located data sources enable direct observation of social interactions and collective behaviors with unprecedented detail. Networks of human mobility [18-20] have revealed the existence of geo-located communities, or patches that exist at multiple scales from town to city, state, and national scales [21]. People in these patches have similar movement patterns and, in a self-organized manner, mostly do not cross the borders of their communities. Borders of patches are subject to vary by changes in the mobility preferences. The applied quarantine policies and lockdowns on large scales have changed the mobility patterns over the past months. Studying the changes and fragmentation patterns allow us to quantify the effectiveness of the policies and define the risk of the areas based on the mobility of individuals.

Methods and Materials

Data:

COVID-19 datasets: We use daily time series data from Johns Hopkins University COVID-19 Data Repository. These datasets provide cumulative counts for confirmed cases at the level of counties for the US. By adding the number of active confirmed COVID-19 cases to the map, we can define risk exposure for the communities.

Mobility datasets: We are collecting multiple data sets to extract the mobility networks in the US. In March 2020, technology companies that gather geo-located information on individuals started sharing anonymized mobility data to help researchers stop the spread of COVID-19. Each of these data sets covers aspects of an individual's movements. We combine mobility datasets from SafeGraph, Twitter, and Facebook.

SafeGraph dataset: The original data comes in CSV format, grouped by days. The file sizes range from 1 to 2.5 GB. Each file describes individual census blocks and lists links with weights (number of links) to other census blocks that occurred on a specific day. We first separate all these relationships and describe them as individual objects. Each relationship has a source, a target, the date, and weight of interaction. Daily dataframes are combined into weekly dataframes, grouped and their relationships are summed. Each census block in each relationship is augmented with central points derived from groups of census blocks that make up a census block group.

Mobility Network: In the mobility network, nodes represent a lattice with cells overlaid on a map of the US. Here, we consider the census block groups as the nodes of our network. Edges represent the movement of an individual from one location (node) to another one. Here, edges' weight represents the number of people who travel between the two census block groups. This network aggregates the heterogeneities of human mobilities in a large-scale representation of social collective behaviors [22].

Community Detection Algorithm: We analyze social fragmentation by applying the Louvain method [23] with modularity optimization [24] to the mobility network. Communities refer to the regions in which nodes are more connected to each other than the rest of the network. In the Louvain method, in an iterative process, nodes move to the neighboring communities and join them to maximize modularity (M). Modularity is a scalar value $-1<M<1$ that quantifies how distant the number of edges inside a community is from those of a random distribution. Values closer to 1 represent better detected communities. Due to the existence of multiple local minima in the Louvain algorithm, some variation in the assignment of nodes may occur between algorithm runs [24, 25]. To quantify the stability of detected communities and identify areas in which communities overlap with each other, we generate an ensemble of multiple realizations and analyze the borders of patches in all the realizations [21].

Results and Discussion:

After the first cases came to the US through international travel, the spread of COVID-19 occurred rapidly through patients with or without symptoms at the time of transmission. COVID-19 has an incubation period that typically varies between 1 and

14 days. Movement of asymptomatic individuals in public increases the risk of the disease in the areas they visit. The greater the number of infected people, the higher the risk of spreading COVID-19. Unfortunately, due to the delays in applying preventative policies across the US, many areas have seen a large number of cases. By tracking the location of recent active cases, we can define the risk exposure for various areas. While doing the analysis in lower resolutions, country or province/state scale, can provide an aggregate view of the world situation, doing the analysis this way will mean the loss of many important details and information. For example, a state may appear be in a good position when it comes to COVID-19, but when we zoom into the state, we may see that there is high risk in a particular area. Alternatively even in states with a high number of cases, some areas may have none. The higher the resolution we can provide, the better we can define the local risk levels. Figure 1 shows the severity of the pandemic at the county level for the US. If a county does not have any cases it is shown in dark green. There is a shift from dark green to dark red showing the highest risk of exposure in the counties with over 20,000 cases (see the legend of the figure). Green areas are the safest areas in the past 14 days. Metropolitan and urban areas are the most risky areas as the population density in those areas is greater.

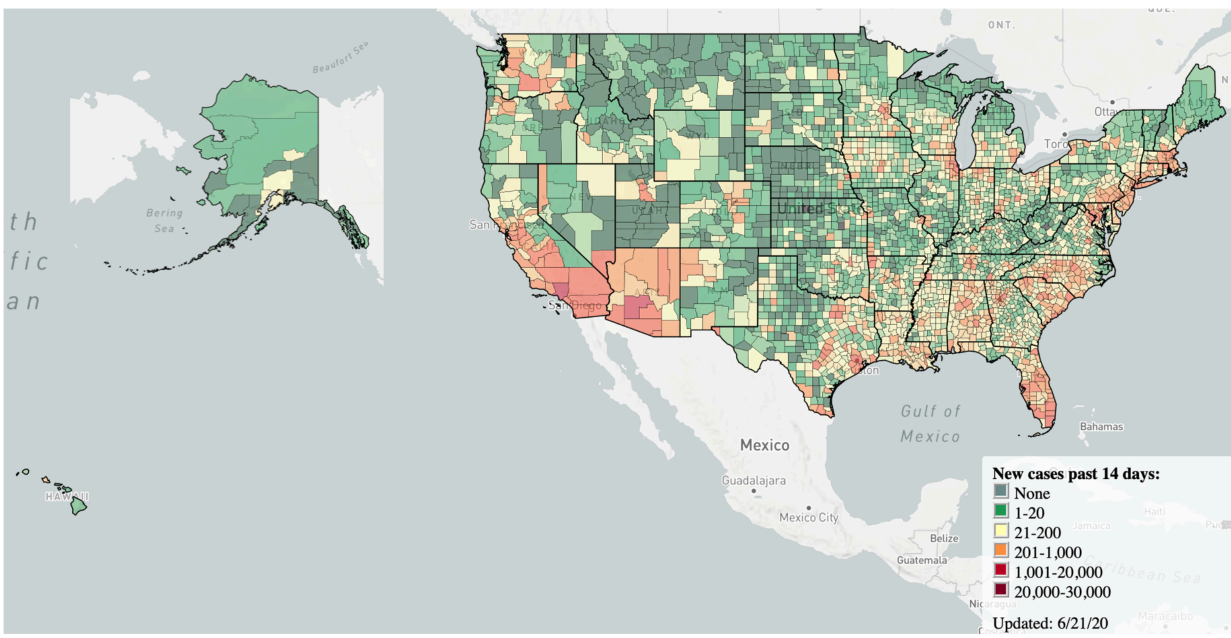

Figure 1: COIVD-19 cases in the counties of the US in 14 days ending 6/21/2020.

To apply the preventive policies more carefully, it is important to define the geographical patches based on actual mobility of individuals. During the global spread of COVID-19, many countries have isolated themselves from high risk countries by closing international flights and land borders. Inside the countries, including in the US, state governments closed public areas (e.g. work places, universities, schools and shopping centers), and asked people to wear masks. At the beginning of the outbreak, people were also asked to abstain from going outside unless for essential or emergency needs. These actions helped to reduce the unnecessary movements and the spread of the disease. To see the impact of these policies, in Figure 2, we compared the mobility networks of the US on January 19-26 and April 5-11, and applied the Louvian algorithm to extract the fragmentation pattern. Nodes with the same color belong to the same community.

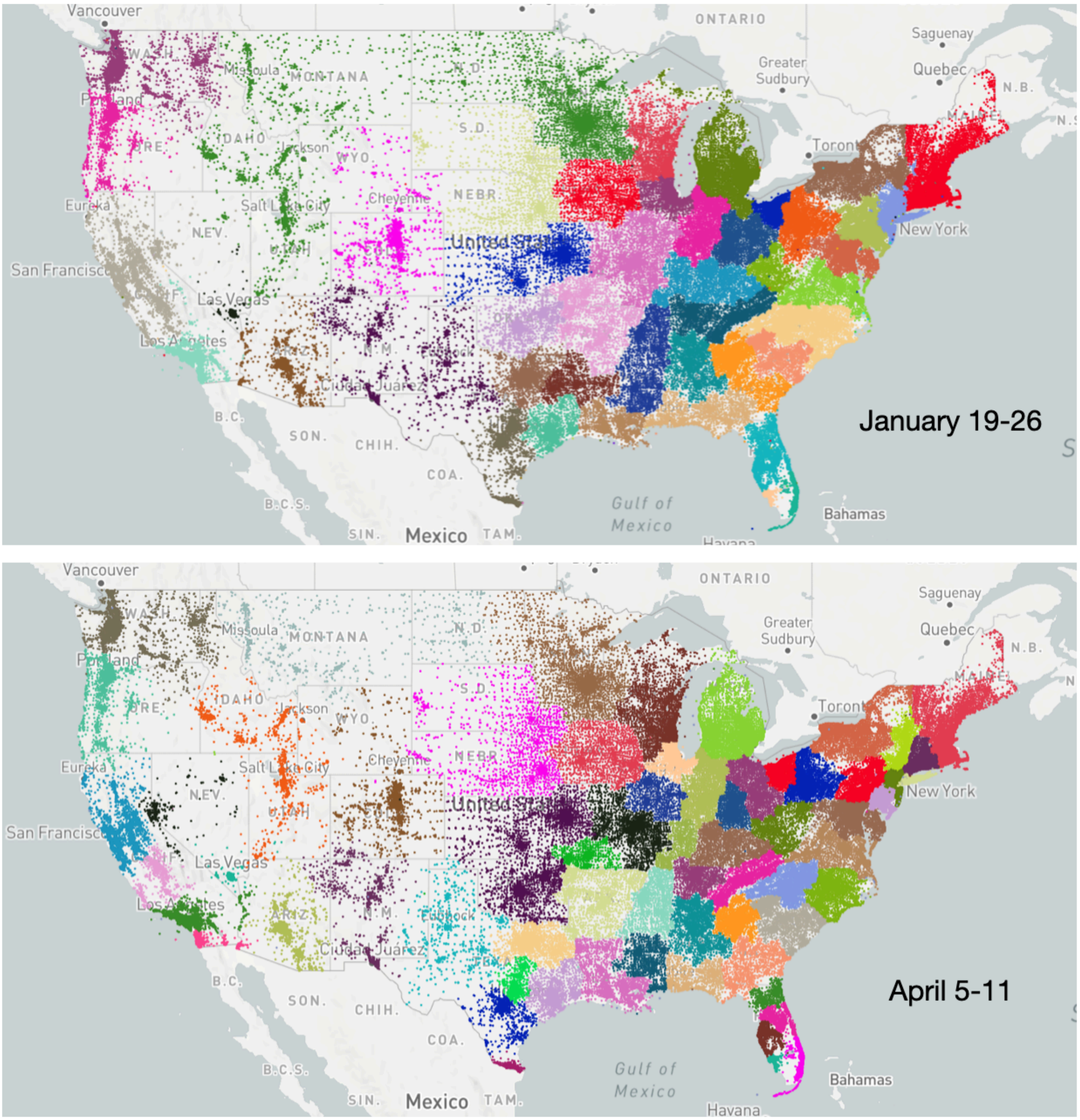

Figure 2: First level of the communities from the mobility network in the US at January 19-26 (upper panel) and April 5-11 (bottom panel).

Geo-communities are important because they indicate where there is more travel between geographic areas, and identify regions and boundaries where limiting travel is easier. Early in the outbreak travelers were limited from the epicenter in NY to other states. Recent regulations seek to regulate travel from the current outbreak areas into the NY, NJ and CT area [26]. As shown in the figure, some of the larger communities in January are split into smaller communities representing the changes in mobility. For example, California, with two main communities including Los Angeles and San Francisco, was divided into four communities in April. Communities in Florida are also divided into smaller communities. This same phenomenon happened for New York and New Jersey. These communities are among the most infected areas in the US.

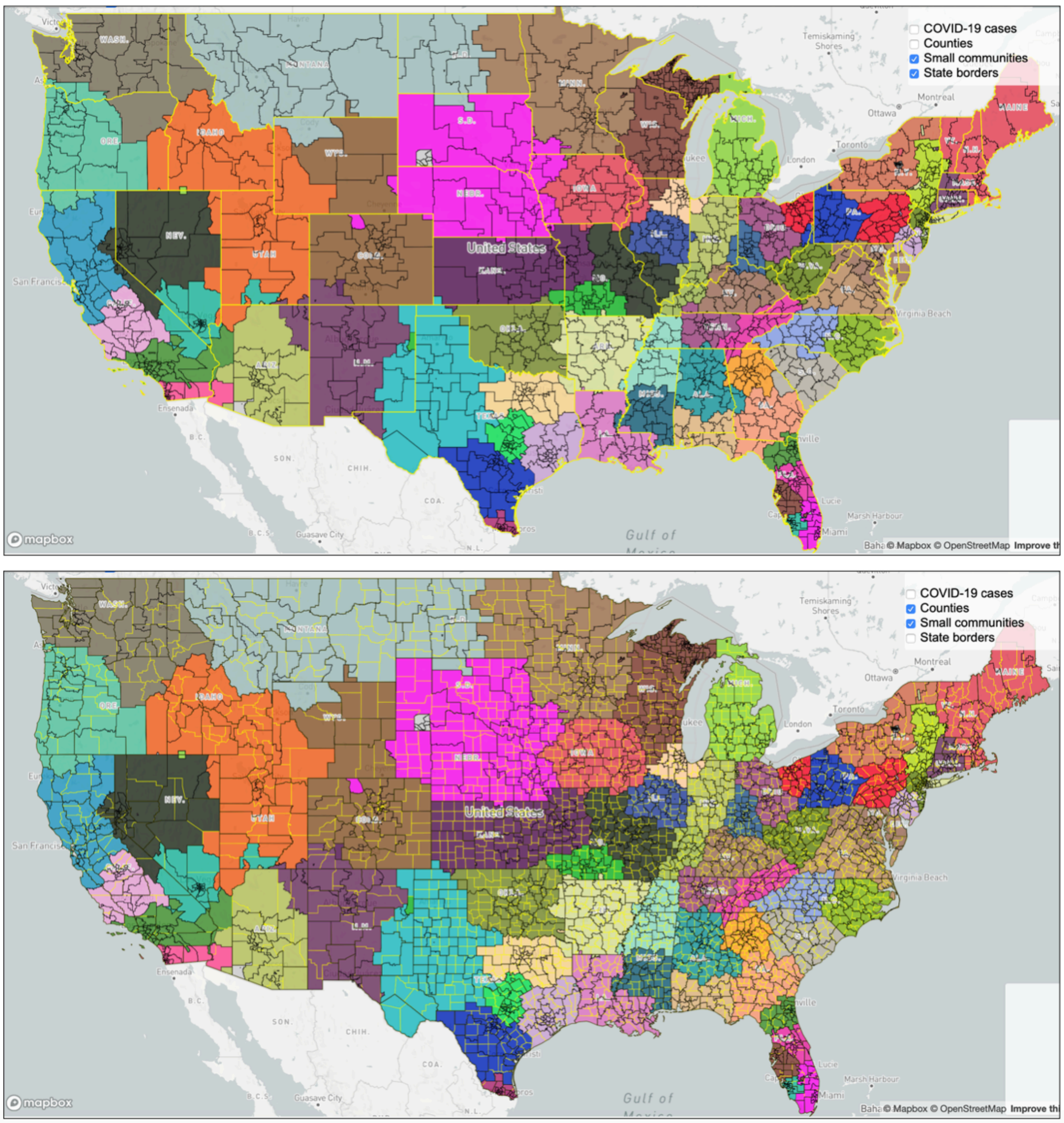

Figure 3: First and second levels of communities in the US during April 5-11. Colors show first level communities. Black lines show finer scale communities. Upper panel shows state boundaries in yellow. Bottom panel shows county boundaries in yellow.

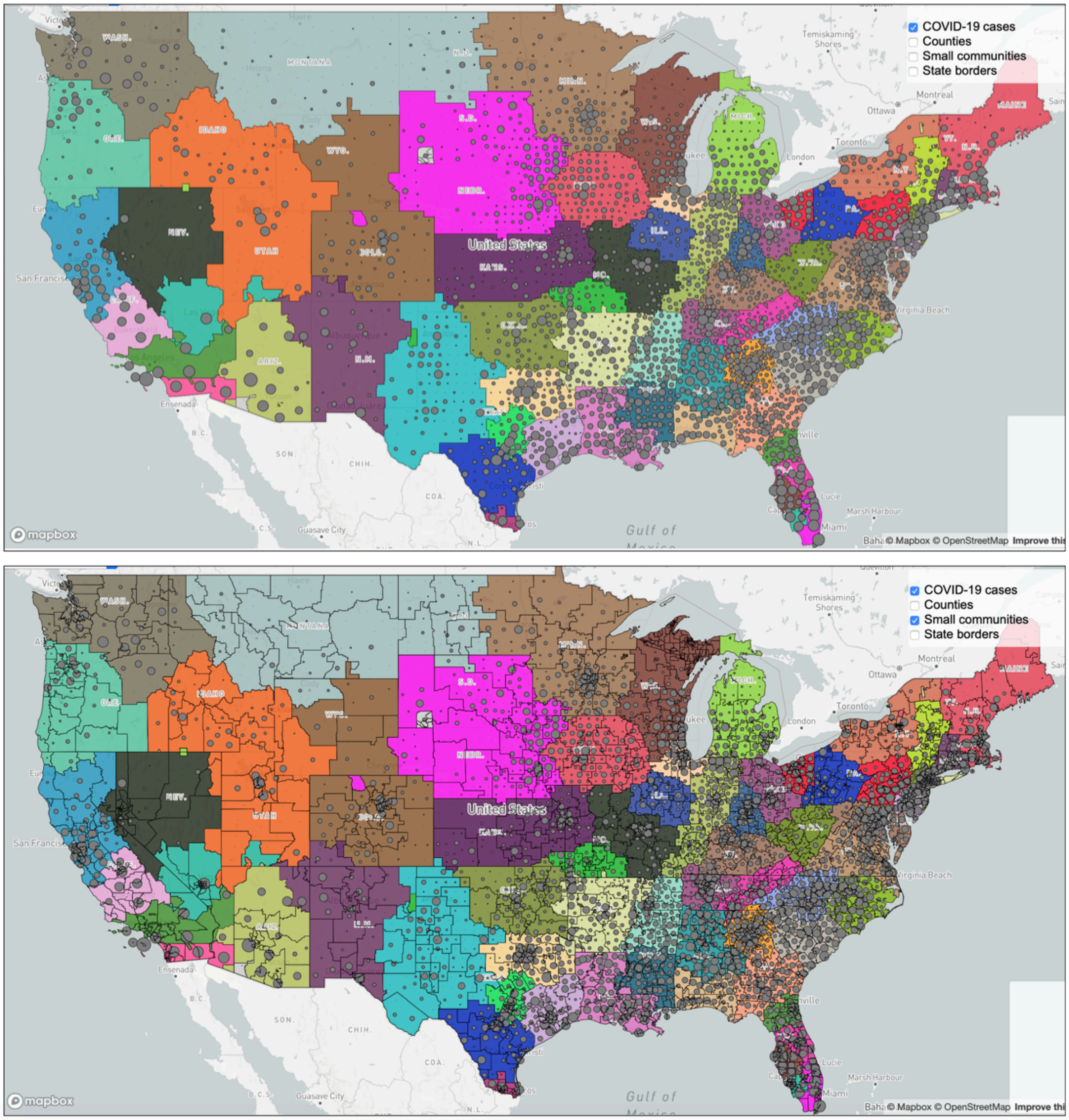

Figure 4: First and second levels of communities in the US at April 5-11 with number of COVID cases at the county scale (gray circles at the centroid of the counties).

In Figure 3, we show the communities at two different scales on April 5-11. Larger communities are shown by different colors and sub-communities are separated from each other with black lines. The second level reveals the substructure of the communities inside each of the larger communities, exposing the mobility patterns in more detail. In the upper panel, state borders are shown by yellow color and in the bottom panel, county borders are shown. From these two panels, it is clear that in some areas community borders align with the social border but in other areas they deviate significantly.

In Figure 4, we add indicators of the number of COVID-19 cases on top of the communities map. COVID-19 cases at the county level are placed in the centroid of the counties. While all large communities are highly infected, some of the sub-communities have much less infected areas, demonstrating they are safer places and have a better potential to reopen earlier than the higher risk areas.

By zooming into the map (see Figure 5(a)), we observe five interesting facts:

- Areas with no mobility data: There are some urban areas that do not share any mobility data, like a Native American community in New York state [27], see Figure 5(b).
- Isolated communities: Some parts of a community are geographically disconnected from the original community. This occurs for university areas and vacation areas for larger cities.
    - Universities: New York State is the home of many universities. These universities attract people from many different geo-clusters. In Figure 5(c), we see the example of Cornell University and SUNY Cortland, two Universities located in central New York State that are sub-communities for the New York City community. This corresponds with a 2014 investigation [28] which estimated that 65% of all students at Cornell from New York State came from that region of the state.
    - Vacationers: There are vacation spots that individuals from metropolitan regions of one community go to that are in the middle of other communities. These regions are known for their nice outdoor space and somewhat close proximity to the city they are connected to. This phenomenon creates isolated communities in the middle of other communities. Multiple reports have mentioned this issue [29, 30]. In New York City, the Catskill Mountains are one of these escapes, Figure 5(d), while for Philadelphia the Poconos serves the same purpose, Figure 5(e).
- Specific sub-communities within other sub-communities: All around the US, there are some areas in which people mostly interact with each other rather than their nearby urban areas. University campuses are good examples of such sub-communities. As shown in Figure 5(f), within the community of upstate/western New York there are specific sub-communities with connection to university campuses. On the left side of the figure, the smaller shape south of Rochester is Rochester Institute of Technology, while on the right side of the figure has 2 smaller areas in Syracuse, both associated with Syracuse University. These universities are large and attract many students from the upstate/western New York region.
- Communities that cross borders: Examples include the area of Philadelphia and southern New Jersey, see Figure 5(g). The light purple region highlights the Philadelphia cluster. This geo-cluster includes a multi-state region in north

eastern Maryland, northern Delaware, southern New Jersey, and south eastern Pennsylvania. The southern part of the Jersey Shore is a popular travel destination for people from Philadelphia, and the areas in Delaware and Maryland appear to be extensions of the greater Philadelphia area.

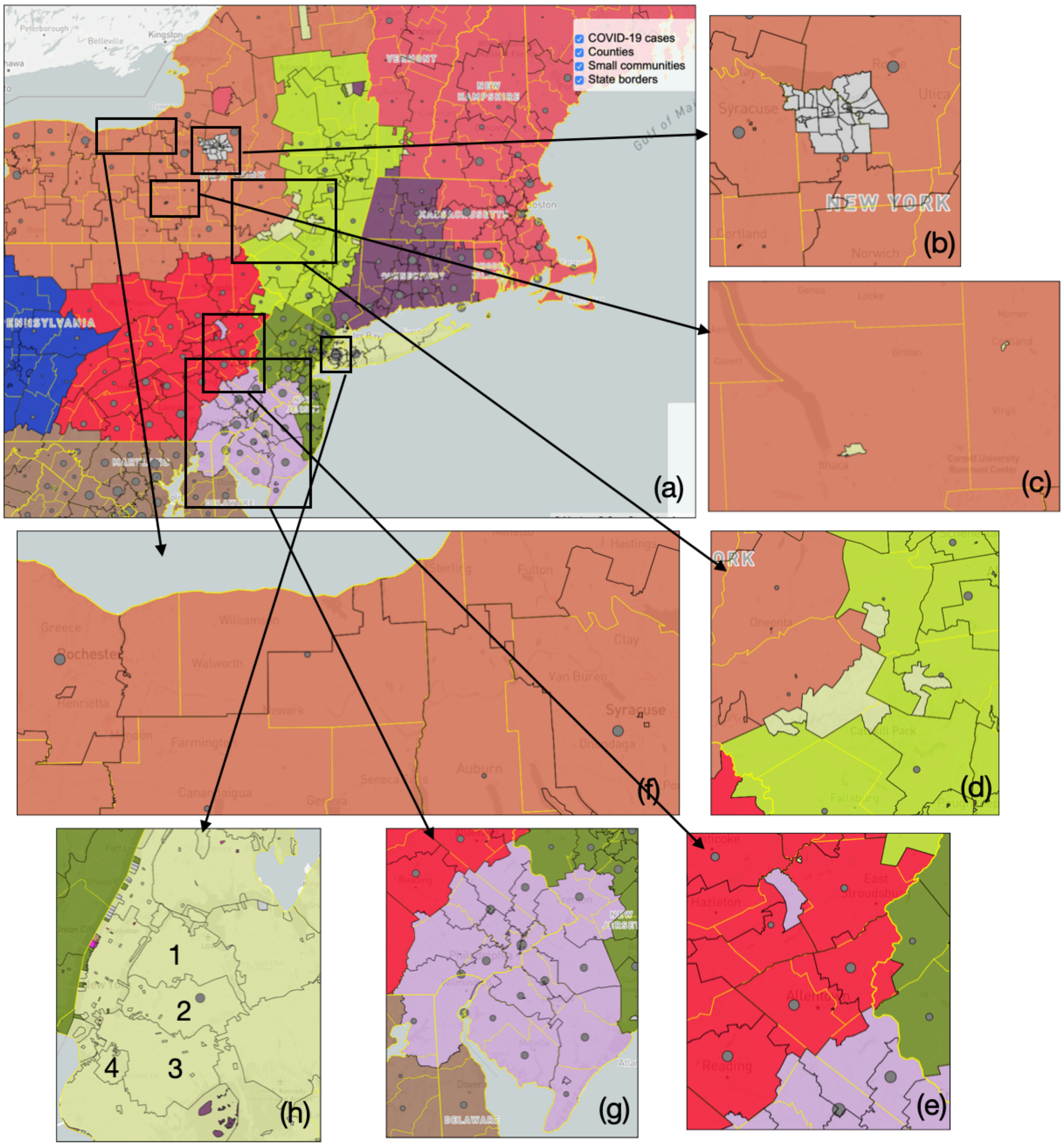

Figure 5: Zoom into the northeast of the US.

- Sub-communities in City areas: Racial and income differences, city infrastructure and transportation can be reasons for community formation in city areas. In New York City, these communities are shown by light green in Figure 5(a). Brooklyn and Queens have defined sub-communities, see Figure 5(h), that are necessary to investigate. Sub-community 1 includes areas of Queens (Long Island City, Astoria, Sunnyside, Woodside, Jackson Heights, Elmhurst, Corona). Sub-community 2 includes parts of northern Brooklyn (Williamsburg, Greenpoint, Maspeth, Middle Village, Rego Park, Forest Hills, Bushwick, Ridgewood, Glendale). Sub-community 3 includes central Brooklyn (Clinton Hill, Bedford-Stuyvesant, Fort Greene, Prospect Heights, Crown Heights, Flatbush, and Canarsie). Sub-community 4 includes areas around Prospect Park (Park Slope, Greenwood Heights, Kensington, Windsor Terrace, Prospect Lefferts Gardens). The public transportation that supports them is different for each area (within Brooklyn division as well). Prospect Park area of Brooklyn is the most wealthy (#4, smallest subsection). The border between #4 and #3 on the map is basically the wealth divide [31]. The racial divide between #3 and #4 is striking on this map as well [32].

References

[1] Sanche, S., Lin, Y. T., Xu, C., Romero-Severson, E., Hengartner, N., & Ke, R. (2020). Early Release-High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2.

[2] Hellewell, J., Abbott, S., Gimma, A., Bosse, N.I., Jarvis, C.I., Russell, T.W., Munday, J.D., Kucharski, A.J., Edmunds, W.J., Sun, F. and Flasche, S., 2020. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. The Lancet Global Health.

[3] Sun, K., Chen, J. and Viboud, C., 2020. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. The Lancet Digital Health.

[4] World Health Organization. "Coronavirus disease 2019 (COVID-19): situation report, 72." (2020).

[5] https://gis.suffolkcountyny.gov/portal/apps/opsdashboard/index.html#/76a26a0c83634266aa9efc35bd4f1975, https://erieny.maps.arcgis.com/apps/opsdashboard/index.html#/dd7f1c0c352e4192ab162a1dfadc58e1, https://covid19.ulstercountyny.gov/dashboard/, https://socpa.maps.arcgis.com/apps/opsdashboard/index.html#/7bd218bc8be04b209c0b80a83fc2eba5

[6] https://www.nj.gov/health/cd/topics/covid2019_dashboard.shtml

[7] https://www.mass.gov/doc/covid-19-dashboard-april-21-2020/download

[8] https://www.latimes.com/projects/california-coronavirus-cases-tracking-outbreak/

[9] https://chesco.maps.arcgis.com/apps/opsdashboard/index.html#/bce5af8a6f454ee78e00b5adc67f4c4a

[10] https://coronavirus.illinois.gov/s/county-map

[11] Mangrum, D., & Niekamp, P. (2020). College Student Contribution to Local COVID-19 Spread: Evidence from University Spring Break Timing. *Available at SSRN 3606811*.
[12] Gottmann, J. (1957). Megalopolis or the urbanization of the northeastern seaboard. *Economic geography*, *33*(3), 189-200.
[13] Morrill, R. (2006). Classic map revisited: the growth of megalopolis. *The Professional Geographer*, *58*(2), 155-160.
[14] Thiemann, C., Theis, F., Grady, D., Brune, R., & Brockmann, D. (2010). The structure of borders in a small world. *PloS one*, *5*(11).
[15] Lazer, D., Pentland, A., Adamic, L., Aral, S., BarabÃ¡si, A.L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M. and Jebara, T., 2009. Computational social science. *Science*, *323*(5915), pp.721-723.
[16] Gonzalez, M.C., Hidalgo, C.A. and Barabasi, A.L., 2008. Understanding individual human mobility patterns. *nature*, *453*(7196), pp.779-782.
[17] Calabrese, F., Dahlem, D., Gerber, A., Paul, D., Chen, X., Rowland, J., Rath, C. and Ratti, C., 2011, October. The connected states of america: Quantifying social radii of influence. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing* (pp. 223-230). IEEE.
[18] Menezes, T., & Roth, C. (2017). Natural scales in geographical patterns. *Scientific reports*, *7*, 45823.
[19] Amini, A., Kung, K., Kang, C., Sobolevsky, S., & Ratti, C. (2014). The impact of social segregation on human mobility in developing and industrialized regions. *EPJ Data Science*, *3*(1), 6.
[20] Xiang, F., Tu, L., Huang, B., & Yin, X. (2013, December). Region partition using user mobility patterns based on topic model. In *2013 IEEE 16th International Conference on Computational Science and Engineering* (pp. 484-489). IEEE.
[21] Hedayatifar, L., Rigg, R. A., Bar-Yam, Y., & Morales, A. J. (2019). US social fragmentation at multiple scales. *Journal of the Royal Society Interface*, *16*(159), 20190509.
[22] Reichardt, J. and Bornholdt, S., 2006. Statistical mechanics of community detection. *Physical review E*, *74*(1), p.016110.
[23] Blondel, V.D., Guillaume, J.L., Lambiotte, R. and Lefebvre, E., 2008. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, *2008*(10), p.P10008.
[24] Sales-Pardo, M., Guimera, R., Moreira, A.A. and Amaral, L.A.N., 2007. Extracting the hierarchical organization of complex systems. *Proceedings of the National Academy of Sciences*, *104*(39), pp.15224-15229.

[25] Good, B.H., De Montjoye, Y.A. and Clauset, A., 2010. Performance of modularity maximization in practical contexts. *Physical Review E*, *81*(4), p.046106.
[26] https://www.cnn.com/2020/06/24/us/new-york-coronavirus-travel-restriction/index.html
[27] https://www.oneidaindiannation.com
[28] https://ithacavoice.com/2014/11/percentage-cornell-students-come-downstate-ny/
[29] https://www.nytimes.com/2020/03/25/nyregion/coronavirus-leaving-nyc-vacation-homes.html
[30] https://whyy.org/articles/corona-discount-as-rentals-advertise-seclusion-poconos-become-per-capita-covid-leader/
[31] http://storymaps.esri.com/stories/2016/wealth-divides/index.html
[32] https://www.nytimes.com/interactive/2015/07/08/us/census-race-map.html