Planes, Trains, and Afflictions

Agent-based modeling for the spread of disease through transportation networks

Frank Acquaye
Ho Lum Cheung
Dimas Muñoz-Montesinos
Elie Wanko

Faculty of Computer Science
Higher School of Economics
Moscow

27 JUN 2020

Table of Contents

Motivation

Why were certain areas of NYC more affected by COVID-19?
How were the subways involved?

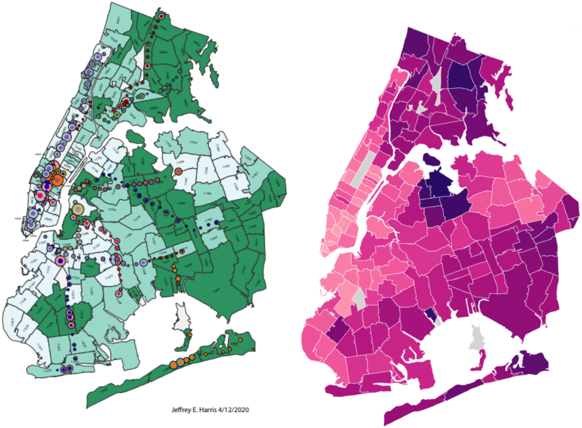
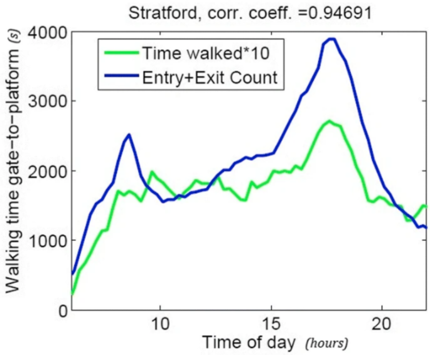
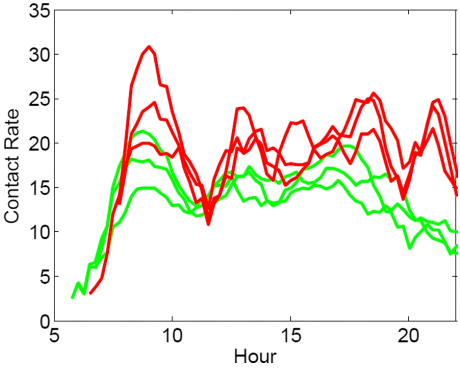


Figure: Left: Case Rate And Subway On April 12, 2020 [1]
Right: NYC COVID Case Rate As Of June 23, 2020 [2]

Prior Research (London Underground)

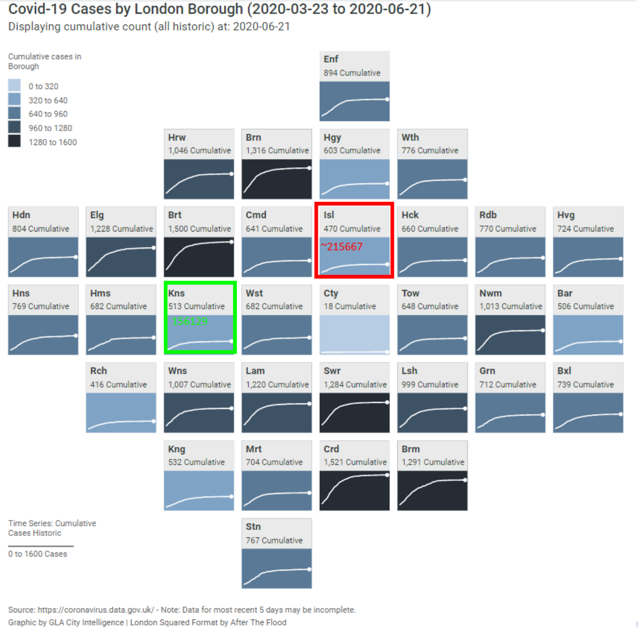
Data: Oyster (Card), CASA (Timetable), PHE Data (Demographics)





Results show a correlation between the use of the underground and ILI cases in London, specifically they show that higher numbers of ILI cases arise in those boroughs where the population spend more time in the Underground and/or incur in a higher number of contacts when travelling. [3]

COVID in London



Prior Research (Singapore Buses)

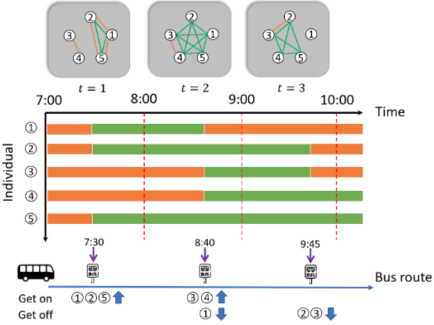
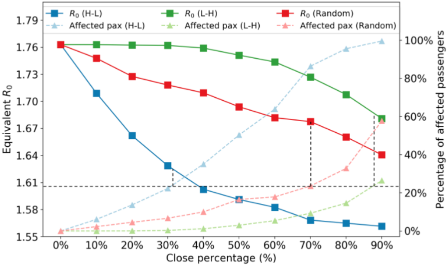
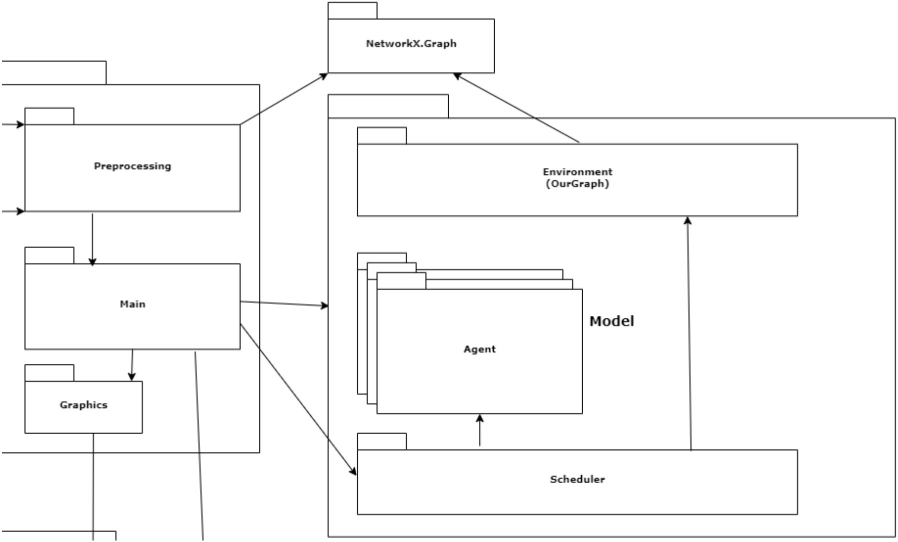


Figure 1: Network representation of a five-passenger system



The direct contact in trains is, however, difficult to obtain from smart card data because the transactions are recorded at the station level [4]

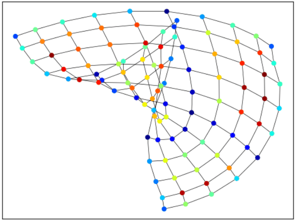
ABM Framework

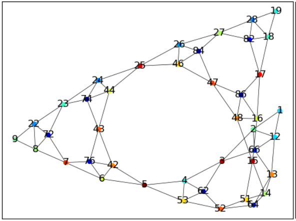


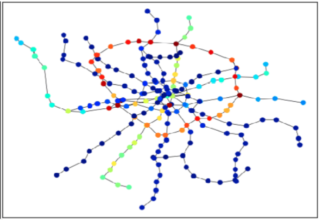
Components: Python, MESA, Networkx

Base Classes: Transportation Model, SEIR Agent

Simple Geometries



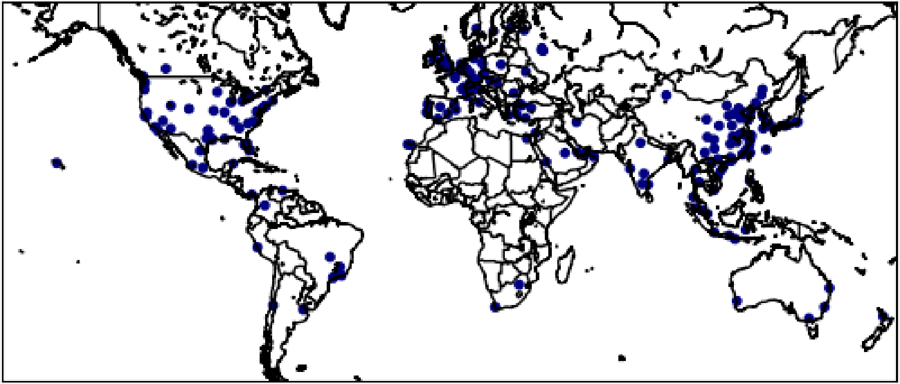




- Grid
- Sierpinski's Triangle
- Moscow

World Airline Network (Passenger Flow)

- Investigate similarities and differences with subways
- Make framework more robust



World Airline Network (Passenger Flow)

Main Idea: Detailed route data can help model spread of diseases.

- Regional / Domestic / EU Domestic / International Airports
- Regional / Domestic / EU Domestic / International Air Routes

Problem: This data not widely available. Solutions [5]:

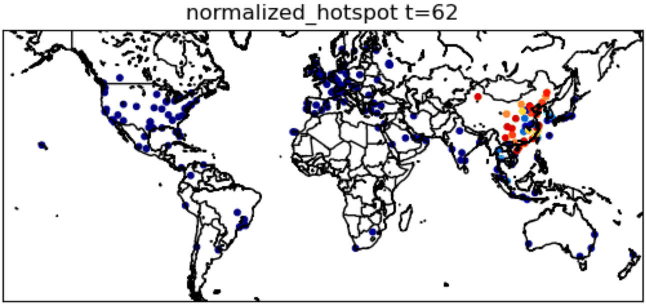
1. Commercial Providers
2. Open Data
3. Statistical Modeling

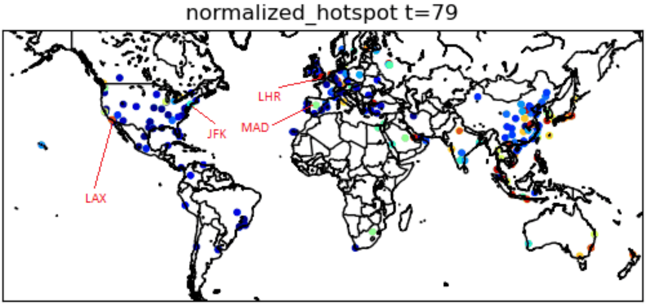
World Airline Network (Passenger Flow)

- List of Airports - OpenFlights [6]
- List of Air Routes - OpenFlights [6]
- Airport Passenger Flow Numbers - [7]
- Top Air Routes [8]
- Other Air Routes

1. Process airports, air routes
2. Add all known passenger flow numbers
3. Add all known airport and air route labels
4. Classify remaining airports and air routes
5. Interpolate missing air route data

World Airline Network Visualization (gif)





Madrid Commuter Trains (Central Hubs)

How important are central transportation hubs to disease spread?

Data Sources:

- Madrid Renfe GTFS: https://crtm.maps.arcgis.com/home/item.html?id=1a25440bf66f499bae2657ec7fb40144
- Madrid Renfe Turnstile data: https://data.renfe.com/dataset/volumen-de-viajeros-por-franja-horaria-madrid

Madrid Commuter Trains

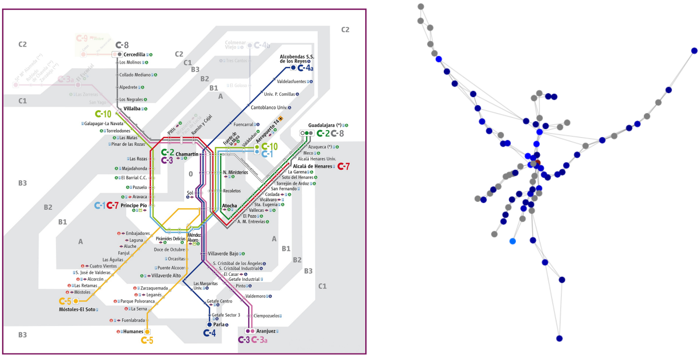


Figure: Madrid Cercanias map 2020 (urban trains)

Madrid Commuter Trains

Method: Agent Based Model (ABM)
Agents: passengers
Flow based on weights w which depend on the turnstiles data: *in-weight* and *out-weight* (incoming and outcoming passengers respectively).

Madrid Commuter Trains - Passengers flow

At each time step t...

N passengers join the network randomly:

1. Join at a random time.
2. Origin station depending on the *in-weight*.
3. Platform choice based on the weight of the next platforms.
4. The passengers wait in the platform until a train arrives.

Passengers who are in a train:

- Leave the network (destination arrival) depending on the *out-weight* of the station.
- Leave the train and go to another platform depending on the weights of each platform.

Madrid Commuter Trains - Parameters

In order to simulate the spread of an infection:

- Passengers (in a day): $500,000$.
- Infected passengers: $50,000$ (10% of them).
- Infection probability: 0.005.

On average, an infected passenger will infect to other 5.04 people.

Madrid Commuter Trains - Results

Figure: Top 10 stations with more infections

Madrid Commuter Trains - Results

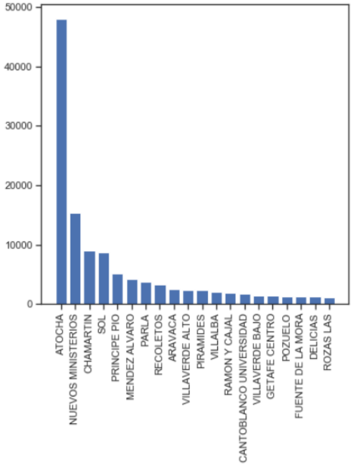


Figure: Top 20 stations with more infections

Madrid Commuter Trains - Results

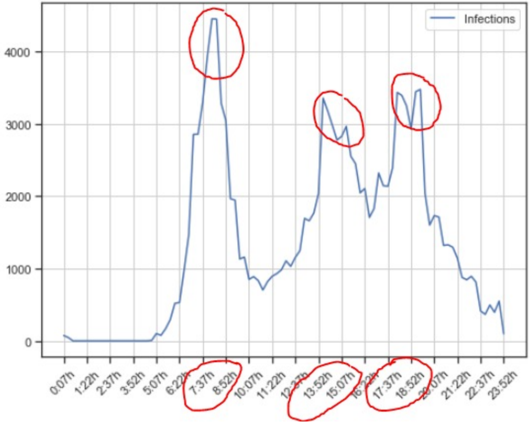


Figure: Infections in the whole network during a day.

Madrid Commuter Trains - Results

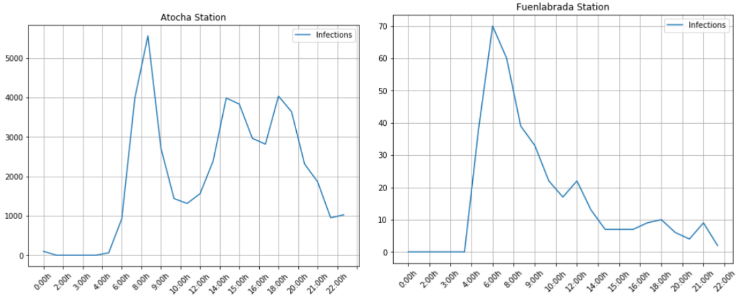
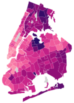


Figure: Infections in a central hub (left) and a station in a suburb (right).

NYC Demographics And COVID Timeline 

NYC residents working in Manhattan primarily travel by subway. This is also true for residents of the Bronx, Brooklyn, and Queens [9]

- March 9 - Mayor holds press conference and notes that there have been 16 confirmed cases. (106 cases)
- March 12 - Mayor declares a local state of emergency. (687 cases)
- March 15 - Schools officially close. (2,986 cases)

Borough - A geographical region. NYC has 5 boroughs.
MODZCTA - Modified Zip Code Tabulation Areas. $\propto$ postal codes.

Subway Systems

Complex, Station, Line, Route

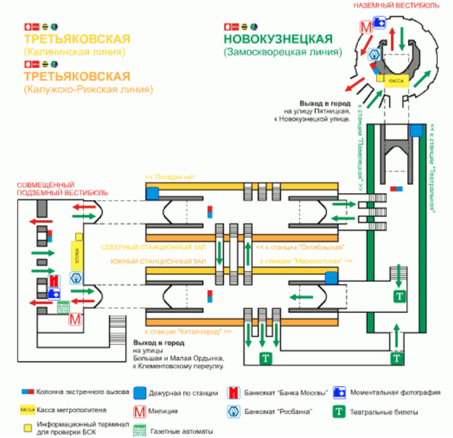


Figure: Floor Plan of Novokuznetskaya Metro Station [10]

NYC Subway Data

Stations, Map, Turnstiles (GTFS not considered)

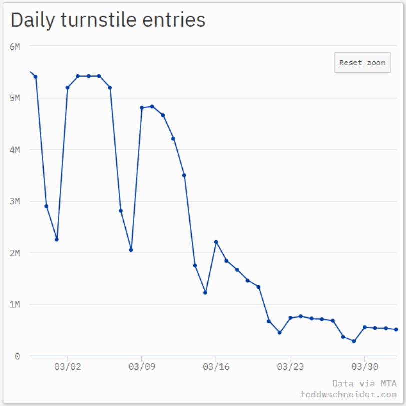


Figure: NYC Subway daily turnstile entries for March 2020 [11]

```
167,167,A32,IND,8th Av - Fulton St,W 4 St,M,A C E...
167,167,D20,IND,6th Av - Culver,W 4 St,M,B D F M...
```

NYC Subway Modeling

Algorithm 1 Simulation of Disease Spread on Subways

```
for i = 1; i < TIMESPAN; i++ do
    Check conditions (i, number of infected) to see if we should deploy COUNTERMEASURES
    for Station in SubwayModel.Environment.Nodes do
        Calculate 'Local Exposure' from infected and commute time.
        Calculate 'Route Exposure' from infected on the same route.
        Calculate 'General Exposure' due to city-wide infected.
        Update 'Exposure' at station based on above conditions
    end for
    for Agent in SubwayNetwork.Agents do
        Get 'Exposure' At Location
        Get City-wide COUNTERMEASURES
        Get Percentage of commuters
        Calculate SEIR beta and gamma based on conditions
        Update SEIR numbers
    end for
end for
```

Subway Agent - SEIR Agent with additional exposure based on location.
Subway Graph - Nodes $\propto$ Stations, Edges $\propto$ **Lines, Complex**, Route Lookup.

NYC Model Parameters, Hyper-parameters

	Contact Rate (β)	Latent Rate (α)	Removal Rate (γ)	Start Trigger
Default	**1.75**	**0.2**	**0.5**	t = 0
Isolation Modifier	0.25	1	2	I >5000
Recommendation Modifier	0.67	1	1.5	I >500
Awareness Modifier	1 to 0.25	1	1	I >500

Other Parameters: Defiance, Global Exposure, Commuter Ratio By Station, Awareness Increase Rate, Borough Commuting Modifier...

Compartmental Modeling Results (NYC)

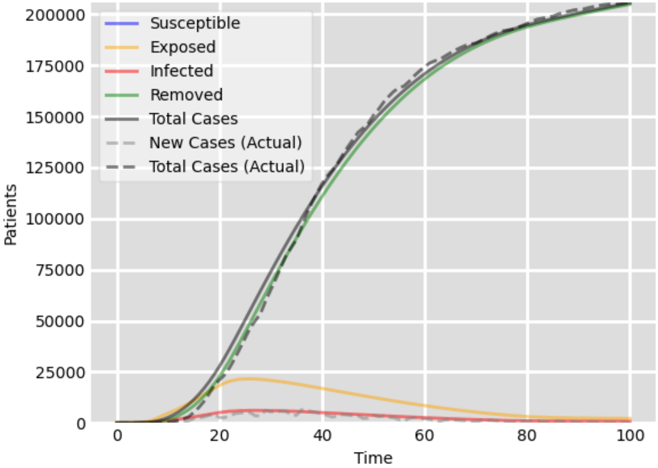
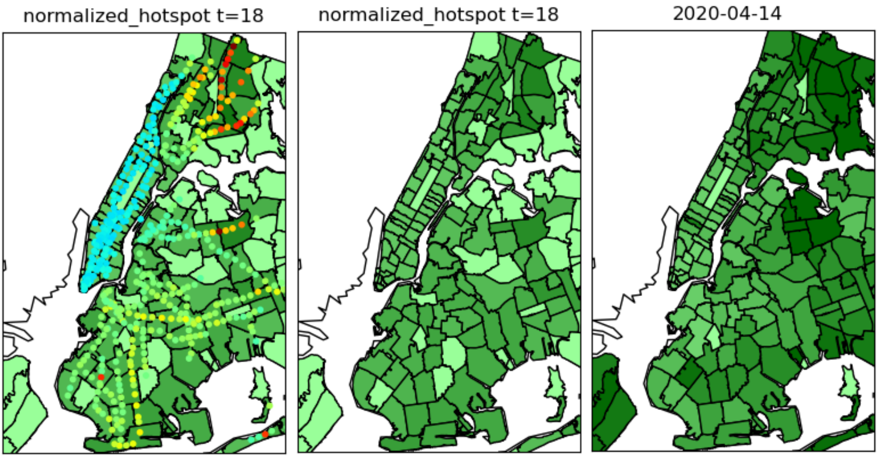


Figure: SEIR Fitting to NYC Case Total. *MAPE* ($t \geq 30$)*: 0.0145*

Results by MODZCTA (NYC)



Scale(0 - 0.11 infections/person, 0 - 0.11 infections/person, 0 - 0.20 cumulative cases/person)

$MAPE\ (t \geq 32, Stations \geq 1 = 0.311\ )$

New York City Timelapse

Discussion

- More granular passenger flow and population data.
- New York City is not a closed system.
- Control for population density, income, and other factors.
- Simplify away the feature bloat

Conclusions

All models are wrong, but some are useful -George Box

- Determining passenger flow is a difficult problem for all transportation networks at all granularities.
- Rush hour and hubs contribute strongly to disease propagation
- Commute time is correlated to prevalence of ILIs
- Subway routes are correlated to prevalence of ILIs

Acknowledgements (and Questions)

Frank Acquaye - WAN, Passenger Flow
Ho Lum Cheung - NYC, Organization, Research, Testing
Dimas Muñoz-Montesinos - Madrid, Hotspots
Elie Wanko - Theory, Consulting
Project on Github

References

- Jeffrey E. Harris.
  The subways seeded the massive coronavirus epidemic in new york city.
  http://web.mit.edu/jeffrey/harris/HarrisJE_WP2_COVID19_NYC_24-Apr-2020.pdf.
- Covid-19: Data.
  https://www1.nyc.gov/site/doh/covid/covid-19-data.page.
- Lara Goscé and Anders Johansson.
  Analysing the link between public transport use and airborne transmission: mobility and contagion in the london underground.
  Environmental Health, 17(1), 2018.
- Yu Shen Clarence Tam Daqing Li Yafeng Yin Jinhua Zhao Baichuan Mo, Kairui Feng.
  Modeling epidemic spreading through public transit using time-varying encounter network.
  Preprint, ResearchGate.
  https://www.researchgate.net/publication/340541290_Modeling_Epidemic_Spreading_through_Public_Transit_using_Time-Varying_Encounter_Network.
- Liang Mao, Xiao Wu, Zhuojie Huang, and Andrew J. Tatem.
  Modeling monthly flows of global air travel passengers: An open-access data resource, Sep 2015.
- https://openflights.org/.
- https://aci.aero/news/2019/03/13/preliminary-world-airport-traffic-rankings-released/.
- https://www.routesonline.com/news/29/breaking-news/286313/busiest-routes-in-the-world-the-top-100/.
- The ins and outs of nyc commuting.
  https://www1.nyc.gov/assets/planning/download/pdf/planning-level/housing-economy/nyc-ins-and-out-of-commuting.pdf.
- Plan of metro station novokuznetskaya.
  http://www.karta-metro.ru/stations/96/428/.
- New york city subway usage.
  https://toddwschneider.com/dashboards/nyc-subway-turnstiles/.