Subways, Social Networks and the Coronavirus

Robert McKee
Apr 19 · 4 min read

Updated: May 11, 2020

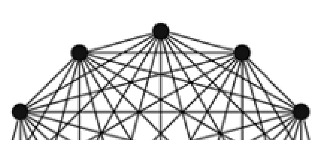
Much has been made of the mathematical model from the University of Washington's Institute for Health Metrics and Evaluation (IHME) that originally predicted as many as 2.2 million U.S. deaths from the coronavirus. IHME has revised that number downward to about 60 thousand deaths, but the *Washington Post* still has called it "America's most influential coronavirus model."

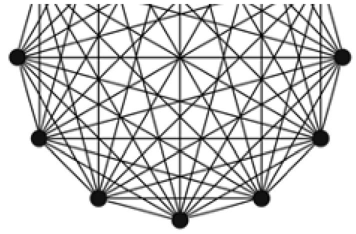
Fierce debate surrounds the model's inputs and predictions, but what if its underlying premise is incorrect?

The gist of the argument is this: the IHME and other traditional models assume what French mathematician David Madore calls "perfect mixing: any individual is equally likely to infect any other individual." But does this reflect reality?

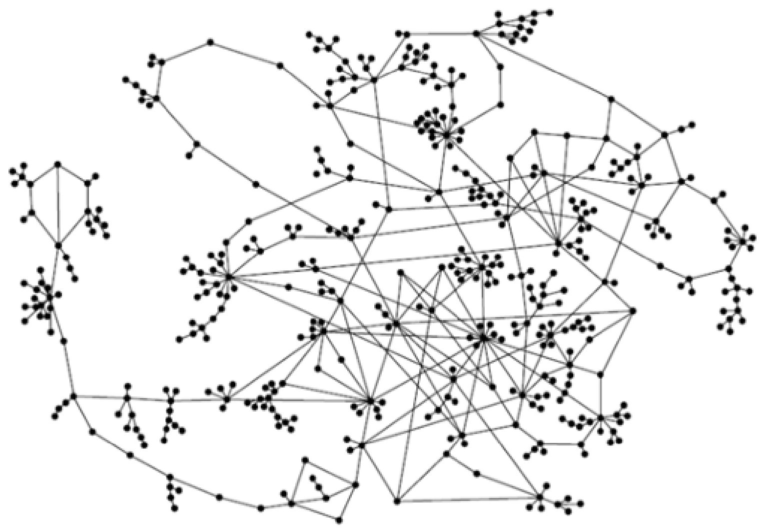
Another school of thought — given far less media and public policy attention — borrows from research into social networks on the web. These networks are called "graphs" and are visualized as "nodes" or "vertices" connected by lines. The visualizations often show people and their connections (e.g., their Facebook friends).

Using the lens of a social network, traditional epidemic models assume that any individual can infect any other individual. Simplified, the models' social network looks something like this:





However, networks do not need to be completely connected. There is another network model called “small world networks.” Here, people organize into various clusters or cliques (e.g., family, friends, colleagues). This is how such a small world network night look:



The example above shows connections between real people for the purpose of tracking an outbreak of sexually transmitted diseases. The key point is that in real life people tend to cluster together, with some “nodes” or people connecting the various groups. People with a lot of inter-group connections can be viewed as super influencers — or super spreaders. When companies want to reach the most people with a marketing message, they target these super influencers. The messages then cascade across the network “virally.”

A research report from the University of Michigan argues that such small world networks apply to the coronavirus. Accordingly, the virus spreads mainly through cliques. At first, the virus spreads exponentially, but this levels off and then drops sharply as the supply of nearby uninfected targets (people or nodes) is exhausted.

Which brings us to subways.

Here is a striking slide from the White House's April 18 update on the coronavirus, showing cumulative coronavirus cases in the U.S.'s top 25 metro areas:

Clearly, something different is going on in the New York City metro area. And the answer might be the subways, according to researchers at MIT, who argue: "New York City's multitentacled subway system was a major disseminator — if not the principal transmission vehicle — of coronavirus infection during the initial takeoff of the massive epidemic that became evident throughout the city during March 2020."

In a series of tweets, political commentator Buck Sexton summarized the MIT findings this way:

"New Yorkers have thought all along that the subway was the covid-19 super spread[.] But with the panic driven lockdown, authorities took a virus that had already spread widely and forced people into more crowded subway cars and buses with limited service. This made it worse….

"So once the subway lines had already help spread this disease to 10,000s of NYC residents, the city authorities told everyone to lockdown at home[.] Queens and Brooklyn have a huge number of multigenerational families living under one roof[.] Such [i]ntra-familial spread can be lethal."

And so this is the alternative to "perfect mixing": The spread is largely through clusters of families, friends and colleagues — small world networks — with the subway connecting otherwise disparate nodes. The subway itself acts as a super spreader.

Viewing the virus through the lens of a small world network also impacts thinking about herd immunity, or how much of the population has to have had the disease and developed immunity to curtail further spreading. Traditional models peg that number at 60 or 70 percent. But, again, traditional models treat all people (or nodes) as equals — they are equally likely to catch or spread the disease. But a number of new papers (here, here and here) argue that if enough super spreaders develop immunity, then the 60 or 70 percent level can drop into the 20 percent or so range. Meaning only 20 percent of the population (if made up mostly of super spreaders) needs to have developed immunity to check the growth of the disease.

Final note: some have argued that the traditional models did their job of communicating danger and prompting action. But like the models' underlying premise of random mixing, policy interventions have been broad and sweeping. The small world model suggests more targeted actions aimed at specific clusters (e.g., nursing homes) and key nodes (super spreaders).

Coronavirus Social Network Analysis Graph Analytics