Mapping the Social Network of Coronavirus

To slow the virus, Alessandro Vespignani and other analysts are racing to model the behavior of its human host.

By **Benedict Carey**

March 13, 2020

BOSTON — The offices of the Network Science Institute at Northeastern University sit 10 floors above Boston's Back Bay. Wraparound windows offer a floating panorama of the city, from Boston Common to Fenway Park, as a half-dozen young analysts toil quietly at computers.

At 10 a.m. on a recent morning, with the early calls to the World Health Organization and European doctors complete and the check-in with the Centers for Disease Control and Prevention scheduled for later, Alessandro Vespignani, the institute's director, had some time to work the room. In a black blazer and jeans, he moved from cubicle to cubicle, giving each member of his team the latest updates on the coronavirus pandemic.

"We call this 'wartime,'" Dr. Vespignani said later in his office; he was seated, but his hands hadn't stopped moving. "Before this, we were working on Ebola, and Zika, and when these things are spreading, you are working on the fly, you don't stop. You are continually modeling networks."

Historically, scientists trying to anticipate the trajectory of infectious diseases focused on properties of the agent itself, like its level of contagion and lethality. But infectious diseases need help to spread their misery: humans meeting humans, in person. In the past decade or so, leading investigators have begun to incorporate social networks into their models, trying to identify and analyze patterns of individual behavior that amplify or mute potential pandemics.

Those findings, in turn, inform policy recommendations. When does it make sense to shut down schools or workplaces? When will closing a border make a difference, and when won't it? World health officials consult with social network modelers on a near daily basis, and Dr. Vespignani's lab is part of one of several consortiums being consulted in the crucial and perhaps disruptive decisions coming in the next few weeks. On Friday, in an analysis posted by the journal Science, the group estimated that China's travel ban on Wuhan delayed the growth of the epidemic by only a few days in mainland China and by two to three weeks elsewhere. "Moving forward we expect that travel restrictions to COVID-19 affected areas will have modest effects," the team concluded.

"Today, with the enormous computing power available on the cloud, Dr. Vespignani and other colleagues can model the entire world using" publicly available data, said Dr. Elizabeth Halloran, a professor of biostatistics at the University of Washington and a senior researcher at the Fred Hutchinson Cancer Research Center. "On the one hand, there is the rise of network science, and on the other, there is the enormous rise in computing power."

Dr. Vespignani came to network analysis through physics. After completing a Ph.D. in his native Italy, he took up postdoctoral studies at Yale, where he began to focus on applying computational techniques to epidemiology and geographical data.

"Look, I am Roman, and I am a fan of Lazio," the soccer team, he said. "We were in first place — finally, after how many years? — and some fans think the coronavirus is a conspiracy against Lazio. I don't say this to be funny, but to say: Each social network functions in its own way."

He was on his feet again and roaming past a row of glass-walled offices. At one point he stuck his head into an office where Ana Pastore y Piontti, a physicist and research associate, was working on one of the problems du jour: school closings, analyzed state by state and region by region. Health officials across the country are grappling with whether to close local schools — which ones, how soon, and for how long.

"Ana's working on this right now, we want to be able to estimate the effects," Dr. Vespignani said.

Her project, like many others at the institute, uses census data, which reveals the composition of nearly every American household: the number of adults and children, and their ages. From a single household, a large map can be constructed. First, the connections between mom, dad, son and daughter. Added next are dad's connections at the shop, mom's at an office, and the children's at their respective schools. The analysis might determine that, say, a 12-year-old boy living in central Redmond, Wash., near Seattle, will come into regular contact with his parents, his sister, and an average of 20.5 fellow students at his local middle school.

Repeating the process with nearby households generates a dense digital map of interconnections over an entire community. On Dr. Pastore y Piontti's computer monitor, it resembles a complex electrical circuit, with multicolored wires and cables to and from packed hubs of interaction.

"Think of it like tracing all regular interactions in the video game SimCity," she said.

To this map, she adds still more connections, incorporating data on travel in and out of that community — by air, train or bus (if such information is available). The final result, which she calls a "contact matrix," looks like a rough heat map — a colored slide showing who is most likely to interact with whom, by age. From this she subtracts out all the school interactions, revealing an estimate of how many fewer interactions — and potential new infections — would occur by closing certain schools.

"Each country, each state, can be very different, depending on the patterns of interaction and compositions of households," Dr. Pastore y Piontti said. "And then there is the question of what is most effective: a week of closing, or two weeks, or closed until next school year."

Dr. Vespignani had disappeared back into his own office with a pair of senior analysts. They were huddled around a speakerphone, running through the latest modeling changes with an outside researcher. The lab is part of a consortium that advises the C.D.C., and fields continual calls from infectious-disease mapping operations around the world.

The conversation and consulting are nonstop, because the institute must navigate the limitations inherent to all predictive modeling. One challenge is that important venues of disease progression cannot all be anticipated: cruise ships, for example. Another is factoring in random events — say, an infected person who suddenly decides that now is the moment to take a dream trip to Spain.

"It may seem like a small thing at the time, but after the fact you say, 'Oh yeah, that was hugely important,'" said Duncan Watts, a computer and information scientist at the University of Pennsylvania. "How do you handle these unexpected factors?"

Finally, as people become more informed about the coronavirus, their behavior will change, sometimes drastically and en masse.

"A good analogy is a storm," said Dr. Steven Riley, a professor of infectious-disease dynamics at Imperial College London, which has done modeling for decades. "You can forecast a bad storm in a particular place, and people will take out an umbrella and put on a coat. Well, the impact is less for those people but has no effect on the storm. With infectious diseases,

people's precautions — like social distancing — do change the trajectory of the disease, and it's very hard to predict or model that."

By now Dr. Vespignani, in motion again, had cornered a visiting colleague, Mauricio Santillana, director of the Machine Intelligence Research Lab at Harvard Medical School.

Dr. Santillana works to understand how the behavior of individuals changes from day to day in the midst of a pandemic. For insight, he draws on a vast array of variables, including mentions of certain words — "fever," "pneumonia," "coronavirus" — in online searches and social-media comments. Together, he and Dr. Vespignani are trying to work out how to best incorporate this continuously updated analysis into the travel and geographical models used at the institute.

"We can look, for example, at when X number of people are searching for 'fever' online, there were Y number of people who ended up in the hospital," Dr. Santillana said. "We can then use that kind of day-to-day data to continually update these social-network models."

All of this, in raw computational terms, is just the beginning of the campaign. No single predictive model is enough; the Vespignani lab, and their colleagues around the world, run millions of simulations regularly, to help gauge which outcomes are the most likely in a world that changes daily. Google has granted him free space in the cloud to do so, because the in-house computing power is not nearly fast enough.

How well this modeling works, and whether it will help contain the virus, is hard to know while the battle is being waged, Dr. Vespignani said: "It's in peacetime, between outbreaks, that we can do the real science, and improve the models. Let us hope that comes soon."