COVID Information Commons (CIC) Research Lightning Talk

Transcript of a Presentation by Peter Pirolli (Florida Institute for Human and Machine Cognition), February 10, 2021

Title: Improving Computational Epidemiology with Higher Fidelity Models of Human Behavior

Peter Pirolli CIC Database Profile

NSF Award #: 2033390

Youtube Recording with Slides

May 2021 CIC Webinar Information

Transcript Editor: Saanya Subasinghe

Transcript

Peter Pirolli:

Slide 1

This is a project about improving computational epidemiology with higher fidelity models of human behavior. This is a project that I'm doing with my Co-PIs, Christian Lebiere and Mark Orr,]. Christian's at CMU [Carnegie Mellon University] and Mark is at University of Virginia. And we have a larger cast of contributor - contributors that are working on all varieties of things ranging from natural language processing to epidemiology.

Slide 2

This project was motivated by the realization that last year, we were in the midst of, historically, the most massive attempt ever to change human behavior. And I'm talking about specifically non-pharmaceutical interventions, such as social distancing, hand washing, mask wearing and now, vaccination.

Slide 3

And decision makers and people who are trying to manage public health, rely on epidemiological models to forecast rates of infections and deaths, and to try to understand what the possible effects are of these NPIs - non-pharmaceutical interventions. Unfortunately, a lot of these models are not very detailed and have a huge abundance of uncertainty. And to some degree, we believe that that's partly because they do not really have accurate models of how people respond, psychologically, and behaviorally to NPIs into what is going on in the environment around them. And just to give you a concrete example, this is one of the many models that we've seen on the news, or on the web over the past year. It's a snapshot that I took in October. And on the right hand side of that graph is a projection for the fall - the month following October, when this was presented. And what you can see is this huge pink error bar around the prediction and that the difference between the top and the bottom of that confidence interval is an order of magnitude. And within that pink area, you can reasonably say: things might go up, they might stay the same or they might go down. So these models have a large degree of uncertainty.

Slide 4

And we are making the bet that by understanding and modeling more specifically an individual's psychology, that we will be able to do better. And this is partly because I think we all believe that people's beliefs and attitudes and intentions and self-efficacy all have an impact on how they respond. And there's certainly evidence out there that that is the case. And it is also the case that these responses seem to change over time. So we've heard a lot about COVID fatigue, and how people's attitudes change over time. And these things also seem to vary across regions. So some regions seem to respond differently than others.

Slide 5

So our aim was to build computational predictive models that are based on a variety of things that we had already been working on. So one set of things was around theories of individual health psychology that some of us had worked on. Another is a large amount of experience with a particular theory and computational modeling system called ACT-R, which allows us to build computational models of behavior change and to develop agent-based simulations. And so out of this, our goal was to develop what we call psychologically valid agents that we can build into agent-based models that will allow us to accurately predict the dynamics of changing behavior over time and how those dynamics are impacted by these NPIs by government messaging, mass media, social media, disinformation, campaigns, etc.

Slide 6

The theory itself, that's the core of our work ACT-R is a constrained, very principled framework for modeling human behavior. It's a theory of the structure of the brain and the functioning of the mind. It's also a simulation environment. And it essentially says, how the modules of the brain that carry out goals and memory and perception - how they operate together dynamically over time to produce behavior, that allows us to model both the symbolic knowledge that people have as well as their statistical adaptivity to the things that are going on around them. And it includes about 45 years of research, based in the laboratory as well as real world applications, a lot of fMRI and EEG imaging data.

Slide 7

And so one way to think of it is - we're trying to build these individual level agents that can simulate the, what we call the response profiles of people. That is, you know, whether they will, in fact wash their hands or wear masks or social distance, or are they going to go out and party or go to - go out to restaurants. And these agents are going to be seated with representations of individual level attitudes and beliefs and intentions. And then those agents will be embedded in an agent-based simulation of given regions and periods. And from that, we want to be able to predict actual behavior that we will compare against some proxy measures that we have of behavior, including mobility data from Unicast, and mass scoring data that's collected daily from COVIDcast. And the way that we are seating these models is using a variety of data that are out there already, including these daily polling data sites. We're also doing a lot of analysis of mass media and online information. And using that to, to get representations that we think characterize individuals in these different regions over time.

Slide 8

So just to give you some examples, we're, we're ingesting a dataset called Third Eye Chyron dataset, which is basically a texturized version of CNN, MSNBC, and all the other major news networks.

Slide 9

We've got access to a variety of datasets of Twitter, including geo-tagged COVID dataset for the world. And our CMU partners have a system called CASOS, which analyzes data from the United States in great detail and great volumes. And so here are some plots of pro versus con tweet volumes in a variety of cities in California that we've collected. And we use GPS mobility tracking provided by Unicast, as well as a day by day polling data about behaviors from the CMU COVIDcast folks at Delphi.

Slide 10

So just to give you one thread of analysis that we're doing, a bunch of folks who are doing natural language processing, and machine learning over Twitter are inducing what we call stances, which are representations of attitudes, beliefs, and intentions from individual level tweets, which are then aggregated up to the users who are making those tweets. And these are stances or attitudes, beliefs towards certain things like mask wearing, or social distancing. We do that at large scale, and then use the representations that come out of that natural language processing to see the representations inside of this psychological valid agent, computational agent, that we're using in our sim[ulations].

Slide 11

Just to give you some ideas of the kinds of behaviors or phenomena that we're trying to model - using our psychologically valid agents, we are modeling a variety of phenomena, these are just a couple on the left here. One phenomena that one sees over and over again, across the world is that as the pandemic hit, there was a great decrease in the effective transmission rates down to around one and then this kind of dampened oscillation around, around a transmission rate of one that seemed to indicate that people were adjusting their behavior to modulate that transmission rate. And our psychological models can in fact, model mask wearing in relation to what's going on in the ambient environment in that kind of dampened oscillating pattern, which is in that lower quadrant there. On the right here are just showing how, how, at the aggregate level, this is for four states of polling data about mask wearing, we can - our models can predict pretty well what the actual mask learning probabilities will be in those four states and we can get down into finer grained regional areas.

Slide 12

So if you want to find out more, please contact me [ppirolli@ihmc.org]. And I just want to mention that our research is funded by NSF and IARPA, and I want to thank the various folks at the bottom here for providing us with data. Thank you!