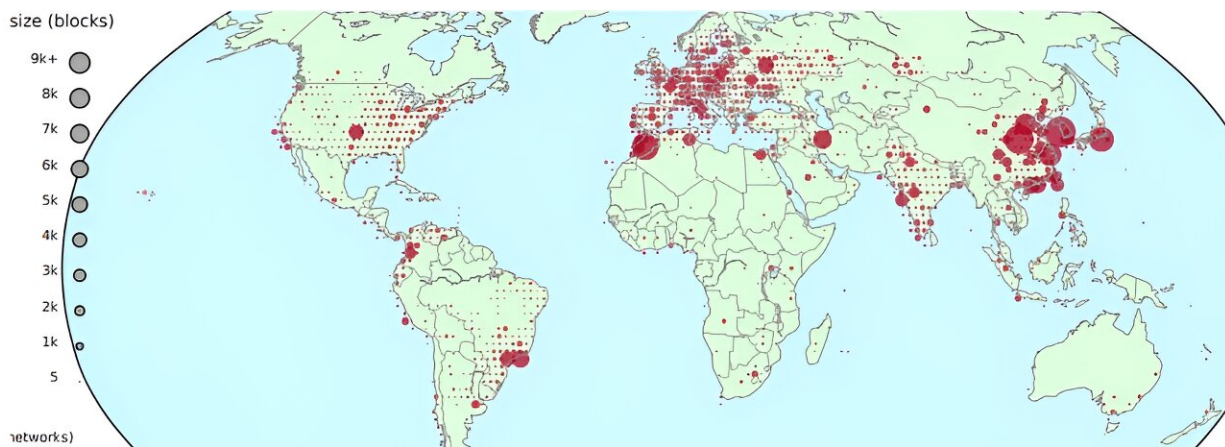# Spotting human activity in internet usage data

November 1 2023, by Julia Cohen

The number of change-sensitive blocks (circle area) by geolocation (in a 2 × 2 ∘ gridcell). Dataset: 2020m1. Credit: *Proceedings of the 2023 ACM on Internet Measurement Conference* (2023). DOI: 10.1145/3618257.3624796

Most people call an internet outage an annoyance. The researchers at USC's ANT Lab (Analysis of Network Traffic) call it a data point.

The ANT Lab, based out of USC Viterbi's Information Sciences Institute (ISI), has been scanning the public internet looking for outages since 2014. USC Ph.D. student Xiao Song and John Heidemann, principal scientist at ISI and research professor of Computer Science at USC Viterbi School of Engineering, were considering this data when the COVID-19 epidemic started in 2020.

Heidemann said, "We were inspired by the network at ISI, where we could see four or five bumps every week, corresponding with the work week. After the Martin Luther King Holiday we saw four bumps in that work week because nobody came in on that Monday. And then we saw COVID hit, and all of a sudden there were no bumps."

The "bumps" they were seeing were ISI employees' laptops and IP addresses connecting to the ISI network when they were at work. Heidemann and Song thought perhaps this could be generalized and applied across the entire internet, to see if they could pick up signals of human activity from internet usage data.

Their resulting paper, "Inferring Changes in Daily Human Activity from Internet Response," is the first demonstration of inferring changes in human activity, such as the transition to work-from-home, from IP responsiveness, and an important example of using the internet to understand our world.

Heidemann and Song presented their paper at the 2023 Internet Measurement Conference, held in Montreal, Canada, from October 24–26, 2023.

The internet isn't out, but employees are

Since 2013, the ANT Lab has had an ongoing project that actively probes the internet to detect outages worldwide (currently 5 million networks measured every 11 minutes).

Heidemann and Song used this existing data to look for and analyze changes in internet usage worldwide that could indicate something about human behavior. They developed algorithms to clean the data, extract underlying trends, and detect changes in activity.

They found that, by using their algorithms, they could identify work-from-home orders that were put in place due to the emergence of COVID-19 in 2020. They could also identify other changes in human activity, such as national holidays and government-mandated curfews.

Song explained, "We looked for significant changes in our human behavior change maps and compared those change event dates to news reports for the same location. For example, around late March 2020, network usage plummeted in Manila, Philippines. The news timeline confirmed that the change we saw correlated with Manila's COVID lockdown which began on March 15, 2020."

Real world events spotted via internet activity: Two case studies

China in January 2020

Using their technique, the team detected activity changes in China in late January 2020. They correlated these changes with two concurrent events: the Wuhan COVID lockdown and the week-long Spring Festival, a national holiday where people typically do not go into their offices. Since the Wuhan lockdown and Spring Festival were concurrent events, they cannot attribute network changes specifically to either one.

India in February and March 2020

The team detected network changes for several days in India in both February and March 2020. When they looked at news in the area, they found that the February network activity correlated with riots in India associated with immigration law protests. While the March network activity corresponded with the first COVID-related curfew in India and subsequent lockdown order.

These [case studies](#) suggest that changes in [human behavior](#) that lead them to work from home can have multiple causes, but their outcome on the internet is similar.

What's next?

Heidemann said, "Our first goal was just curiosity—can we see human activity in the internet?" Now that the team has shown they can, what can they do with and learn from this information. He continued, "In the context of COVID, we could explore questions like: what countries have lockdowns or stay-at-home orders? When do they have them? If you have a stay-at-home order, do people actually follow it or do they not comply and go into work anyway?"

This ability to detect trends in [human activity](#) from the [internet](#) data provides a new ability to understand our world, complementing other sources of public information such as news reports and wastewater virus observation.

Heidemann explained, "People use wastewater detection systems to understand the background level of COVID in a city. And that's great, because it's an anonymous and reliable method to judge what's happening in a city as a whole. I would hope that our technique could provide a similar kind of anonymous, independent, third party observation about what's going on, and that might feed into public health decisions."