

Algorithms reveal forecasting power of tweets

September 11 2014



Sang Won Yoon had a good Chinese meal recently—not always easy in America. It's on his mind.

Imagine, he says, that you and your co-workers plan via social media to head for lunch about 12:30 p.m. most Thursdays. Usually that Italian place downtown. Frequently tweet about traffic on the way.

Now imagine that at 10 a.m., you're tweeted a coupon from the Chinese place near the Italian joint—and directions around a traffic jam that will

start in about 90 minutes. Score one Sichuan hot pot.

Yoon can make that happen. He and fellow Binghamton University systems scientist Sarah Lam have been working with Binghamton alumnus Nathan Gnanasambandam, a senior researcher at the Palo Alto Research Center (PARC), a division of Xerox Research. They used 500 million tweets to develop algorithms that not only paint a picture of everyday human dynamics, but can predict an individual's behavior hours in advance. The team, which also included graduate students Keith Thompson and Bichen Zheng, recently published their findings in *Industrial Engineer*.

Think about what your typical social media post says about you: when you posted, where you were. Your networking relationships can be learned—and with context-based algorithms like those PARC and Binghamton University have developed—what you plan. They use what is called an artificial neural network.

How sure are they? Better than 90 percent for a typical social media user in a three-hour horizon. "If you look at the picture, it's very static. But the individuals are all over the place," Yoon says.

Some people are very careful about what data they give out, but the algorithms can work pretty well with anonymized data. Usable predictions can be made more than 60 percent of the time, if the right data are aggregated. And that data isn't just coming from [social media](#): Think about sources such as credit card transactions, monitored telephone calls, e-mail, GPS data.

Creepy, perhaps, but this type of analysis also has benefits. Xerox, which has funded and participated in the team's ongoing research, can apply the tools to traffic. (It helps run the New York State Thruway's EZ-Pass system and parking services in several cities across the country.) Imagine

getting directions during an emergency that not only get you out of harm's way, but get you to someplace personal where you're safe, reducing the burden on emergency shelters. Or imagine directions that prevent a traffic jam, rather than simply route you around one.

Now apply that research tool to call and contact centers, which Xerox also runs. These methods can fuse data from call centers, online chat and e-mail help desks. "We give it structure—not all feeds have structure," says Gnanasambandam, who is also a visiting professor in Binghamton's department of systems science and industrial engineering.

"What if you call a company..." Yoon says, and Lam completes: "... And they know why you're calling before you call?"

Help desk associates can be cross-trained in topics so they face less downtime, or calls could be routed faster to the best specialist. Data about problems can be analyzed in near-real time, perhaps allowing fixes to be made before the customer realizes there's a problem. "That's not too far away from what's happening," Gnanasambandam says.

Now direct this approach toward healthcare—which provides about \$2 billion of Xerox's annual business—and researchers can build tools to help patients, doctors, hospitals, insurers and pharmaceutical companies better understand the complexities of public health or ferret out prescription or Medicaid fraud.

"There's a lot of different directions you can go," Lam says.

Including to Yoon's next Chinese meal.

Provided by Binghamton University

Citation: Algorithms reveal forecasting power of tweets (2014, September 11) retrieved 23 April 2024 from <https://phys.org/news/2014-09-algorithms-reveal-power-tweets.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.