

Data on taxi routes and points of interest may improve crime predictions

August 16 2016

Data on how taxis travel through communities and on how people label points of interest on social media could help analysts and criminologists better understand neighborhood crime rates in a city, according to Penn State researchers.

Analysis of data from points of interest in Chicago—including restaurants, shops, nightclubs and transit stations—designated by members of FourSquare, a [social media](#) site, along with the city's taxi flow information, offered significantly more accurate estimates of crime rates compared to traditional means. Crime analysts currently mainly rely on demographic and geographic data to study crime and predict trends.

Big data projects could improve understanding of crime and help planners make better decisions, as well as allow communities and police to use their resources to more efficiently fight crime, said Jessie Li, assistant professor of information sciences and technology. Taxi routes are like hyperlinks, connecting different communities with each other, added Li, who worked with Hongjian Wang, doctoral student in [information sciences](#) and technology; Daniel Kifer, associate professor in computer science and engineering and Corina Graif, assistant professor of sociology and criminology, all at Penn State.

"We had this idea that taxis serve as hyperlinks because people are not only influenced by the nearby location, but they are also frequently influenced by the places they go to," said Li. "For example, your home

may be a half hour drive from your work; they are not spatially close. But you spend a lot of time there and you end up being influenced by people, such as your colleagues, there."

Points-of-interest information may improve crime statistic analysis because it shows how certain areas are used and why people want to be there, according to the researchers, who present their findings today (Aug. 15) at the conference on Knowledge Discovery and Data Mining in San Francisco, Calif.

"According to the data, areas with nightclubs tend to be low crime areas, at least in Chicago, which may be a surprise to many," said Li.

"However, it may reflect the people's choices to be there—they want to go to a nightclub that is safe, not one that's dangerous."

Li said that this study also points to how the field of [big data](#) is providing both new sources of data and new ways to explore the implications of that data.

Big data can often show correlations between the sources of data and certain effects, such as [crime](#), which is helpful for making predictions. However, Li pointed out that the sources of data are not necessarily causing the effect.

"What we see here is a correlation between the taxi and points-of-interest data and [crime rates](#)," said Li. "The data show us the correlation, but, scientifically, as far as a cause, we don't know."

The researchers used data on taxi trip records in Chicago, which included pickup and drop off times and locations, operation time and total fare amount, from October to December 2013. They also gathered 112,000 points-of-interest from FourSquare for the study. Statistics on crimes in Chicago were gathered from the city's data portal and

demographic details included information on population, poverty, disadvantage index and ethnic diversity.

Provided by Pennsylvania State University

Citation: Data on taxi routes and points of interest may improve crime predictions (2016, August 16) retrieved 30 April 2024 from <https://phys.org/news/2016-08-taxi-routes-crime.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.