National study offers new bike count models: Combining traditional counters and emerging GPS data

July 15 2022

Automated bicycle counter on Portland's Tilikum Crossing bridge. Credit: Cait McCusker

In order to make sure bicyclists' needs are considered when improving a transportation system, planners and engineers need to know how many people are biking, and where.

Traditional bicycle counters can provide data for limited sections of the bike network, often these counters are installed at important locations like trails or bridges. While limited in location, they count everyone who

bikes by. Meanwhile, GPS & mobile data cover the entire transportation network, but that data only represents those travelers who are using smartphones or GPS. Combining the traditional location-based data sources with this new, crowdsourced data could offer better accuracy than any could provide alone.

"Knowing how many people are bicycling on a street is really important for a number of reasons. As just a few examples, bicycle volumes give you a way to understand safety data and determine crash rates. They provide insight into where and how bicycle trips are taking place, which can help plan for new or improved facilities," said Nathan McNeil of Portland State University.

Supported by a pooled fund grant administered by the National Institute for Transportation and Communities (NITC), Dr. Sirisha Kothuri of Portland State University led a research project aimed at fusing traditional and emerging data sources together, to derive bicycle volumes for an entire transportation network. They developed three models and tested them in six cities: Dallas, Texas; Portland, Bend and Eugene, Oregon; Boulder, Colorado; and Charlotte, North Carolina.

Developing Three Bike Count Models

With Kothuri as principal investigator, the research team included Joe Broach and Nathan McNeil of PSU; Kate Hyun, Stephen Mattingly and Md. Mintu Miah of the University of Texas at Arlington; Krista Nordback of the University of North Carolina's Highway Safety Research Center, and Frank Proulx of Frank Proulx Consulting LLC.

First, the team conducted a literature review while cataloging and evaluating the available third-party data sources and existing applications. They chose the six study sites to represent a variety of urban and suburban contexts, with plenty of geographical diversity, and

existing bike data available. Of the six, Boulder, Charlotte and Dallas constituted basic sites, where one year of data (2019) was used for modeling. Portland, Bend, and Eugene in Oregon were considered enhanced sites, where three years of data (2017–2019) were used for model estimation.

The team chose three relatively new data sources: Strava, Streetlight Data, and GPS data from bike share systems in the case study cities. After collecting demographic, network, bike count and emerging data from the new sources for each of the cities, they developed three sets of models:

1. One with pooled data from all six cities,
2. another with just the pooled data from the three Oregon cities,
3. and finally a set of city-specific models.

The researchers then applied the results to each of the six study sites. The city-specific models generally performed the best, showing the most accuracy in predicting bicycle volumes. The scripts used to run the models will soon be published to GitHub, and a link will be posted on the project page for anyone interested in accessing the models.

In general, the various data sources appeared to be complementary to one another; that is, adding any two data sources together tended to outperform each data source on its own. Adding even more data should continue to refine accuracy. The findings from this study indicate that rather than replacing conventional bike data sources and count programs, big data sources like Strava and StreetLight actually make the old "small" data even more important.

"We will need more ground-truth counts for low-volume sites to capture the variety of locations, and that will make more robust models," said Kate Hyun of UTA.

Better Models Provide More Accurate Performance Measures For Transportation Agencies

Josh Roll, Research Analyst & Data Scientist at the Oregon Department of Transportation, served as the chair for the project's technical advisory committee. He believes the outcome of this research could help transportation agencies get a better handle on how many people are biking in their communities.

"At ODOT we just adopted "Bicycle Miles Traveled" as a new key performance measure, and we need a way to measure it, so this project very much helps to fill the gap on how we're going to do that. This research used cutting-edge data fusion techniques that could lay the groundwork for how transportation agencies like ODOT monitor bicycle activity across the system," Roll said.

For transportation agencies wishing to support active travel to meet various sustainability, public health, and climate-related goals, quickly having accurate data for the entire network would be a giant leap in the right direction.

Robust, organized, and accessible count programs will be essential to get the most out of emerging data sources. The more good, vetted data are available, the better models based on emerging sources will perform, so professionals managing bicycle count programs should focus on making data uniform and widely usable.

"In order to integrate all of these disparate data sources—automated and manual counts, opt-in apps like Strava, passively collected background data like Streetlight, and GPS-enabled bike sharing systems—into one coherent system, data professionals should organize their data to best take advantage of these new data fusion possibilities. This means making

sure nonmotorized data are accurate, consistent, and useful," said Sirisha Kothuri, lead researcher on the project.

More information: [ppms.trec.pdx.edu/media/projec …_Bicycle_Volumes.pdf](ppms.trec.pdx.edu/media/projec)

Provided by Portland State University