Comprehensive Performance Assessment of Passive Crowdsourcing for Counting Pedestrians and Bikes

Project 2025
January 2022

Wen Cheng, PhD
Yongping Zhang, PhD, PE
Edward Clay

Walking and bicycling provide several health, environmental, and economic benefits to those performing these actions for transportation, communities, and local traffic areas alike. Due to the numerous benefits active transportation offers, it is essential to understand pedestrian and bicyclist traffic volumes at various times and locations in order to better plan and design active transportation-related infrastructures and establish associated policies. Over history, there are many typical methods available to record these volumes, including manual counting on site, permanent bike and pedestrian detectors/counters, counting via videos, and so on. In recent years, crowdsourcing has seen a rise in popularity due to the ease of collecting data this way compared to traditional methods. Nevertheless, crowdsourced data have been applied in much fewer studies, and their reliability and performance relative to other conventional methods are rarely documented. To this end, this study examines the consistency between crowdsourced and count data collected from a more typical way, such as permanent counters. Additionally, the research developed the adjustment factor between the crowdsourced and permanent counter data and estimated the annual average daily traffic (AADT) data based on hourly volume and other predictor variables such as time, day, weather, land use, and facility type.

Study Methods

The research team collected data that originated from four separate sources, including crowdsourced StreetLight data, the national archive for bicycle and pedestrian count data maintained by Portland State University, the permanent counter data from one local city (San Jose), and weather data from Weather Underground.

For consistency checking between the StreetLight data and the permanent counter ones, the team used four different statistical tools, assuming both linear and nonlinear associations. For the development of the systematic adjustment factor and estimation of AADT, both fixed and random intercept models based on the Integrated Nested

Laplace Approximation (INLA) package in R approach, were developed with the performance being evaluated under a set of assessment criteria.

Findings

1. Based on the data collected from some cities in California, the StreetLight count data for pedestrians and bicyclists appear to be a viable alternative to the permanent counters in specific various circumstances where the data outliers were removed. The former demonstrates a notable consistency with the latter from different perspectives, including statistical difference, linear association, and non-linear association.
2. The discrepancy between StreetLight and permanent counter data is much smaller after the SL data are adjusted by applying the developed adjustment factors using the different count models.

This pioneering research explores the consistency between crowdsourced count data and those from the permanent counters, and develops the associated adjustment factors.

Policy/Practice Recommendations

1. For agencies using crowdsourced data, based on the findings of the research recommend adjusting such data before its use due to the systematic error associated with the crowdsource data, which are reliant on cellular devices.
2. For agencies wishing for the AADT of the non-motorist counts but subject to limited resources, they may refer to the models developed in the study that link AADT with an hourly volume of specific time periods or develop their preferred models for their needs. Such models can estimate the AADT based on the hourly volume, which is much easier and more economical to collect than AADT that requires continuous data availability for at least one year. The estimated AADT can also be adjusted based on different conditions such as hours, month, weekday, temperature, land use, facility type, etc.

About the Authors

Wen Cheng is a professor from civil engineering department at Cal Poly Pomona. His specialty areas include statistical modeling and traffic safety.

Yongping Zhang is an associate professor from civil engineering department at Cal Poly Pomona. His specialty areas include transportation planning models and policy development.

Edward Clay is a research assistant from civil engineering department at Cal Poly Pomona. His specialty areas include statistical modeling, data mining and computer vision.

To Learn More

For more details about the study, download the full report at **transweb.sjsu.edu/research/2025**

MTI is a University Transportation Center sponsored by the U.S. Department of Transportation's Office of the Assistant Secretary for Research and Technology and by Caltrans. The Institute is located within San José State University's Lucas Graduate School of Business.