

Transportation
Safety and Mobility

MAY 2019

Project Title:

Combining California Household Travel Survey Data with Harvested Social Media Information

Task Number: 2874

Start Date: April 15, 2015

Completion Date: March 31, 2016

Task Manager:

Hassan Aboukhadijeh
Transportation Engineer
haboukha@dot.ca.gov

Combining California Household Travel Survey Data with Harvested Social Media Information

Develop a program to identify and address pedestrian safety problems in California, with the goal of reducing pedestrian fatalities and injuries.

WHAT IS THE NEED?

When something is changing in the infrastructure surrounding residences and businesses, we also expect to see travel demand to change. For example, providing a new travel option may motivate people to abandon an older option and improving the flow of traffic may encourage people to travel more. To detect changes in travel behavior we first need an inventory of current conditions and behavior, or baseline travel demand. This inventory should include information about usual fluctuations in travel demand by day of week and by season. Though a comprehensive travel demand inventory has never been created before, information from traffic counts and surveys can help us develop this baseline for households living in California. This baseline description will allow us to determine whether changes in behavior at specific locations are the result of “natural” fluctuations of travel behavior or the result of new infrastructure. Since a complete baseline and procedures to detect deviations from the baseline do not currently exist for the entire State of California, in this project we will perform a pilot test with existing data and enhancements from social media data.

WHAT WAS OUR GOAL?

The goal of the project is to produce an automated procedure of baseline travel demand creation and a method to detect changes in travel demand at specific locations.



Caltrans provides a safe, sustainable, integrated and efficient transportation system to enhance California's economy and livability.

WHAT DID WE DO?

In this research project we used data from multiple sources to produce statewide travel patterns and large-scale estimates of induced travel demand. In this way we developed a baseline short- and long- distance travel inventory that includes statewide vehicle miles traveled (VMT). We created a procedure to monitor the evolution of travel in California using data from social media adjusted by region and correlated with land uses at fine geographic areas. We estimated induced demand by monitoring regions of California, which have experienced major changes in land use or other policy, for a concomitant travel demand change. We first developed a synthetic inventory of travel in California using data in the California Household Travel Survey (CHTS), Origin-Destination data from the Statewide Travel Demand Model, and harvested social media data. Second, we developed a conversion procedure to transform social media data into origin-destination travel and trip lengths to produce estimates of VMT. Third, we created an automated procedure to harvest data and convert them into travel predictions statewide and then used to derive estimates of induced demand. This project also allowed us to plan the creation of an observatory for ongoing data collection of this type

WHAT WAS THE OUTCOME?

In this research we produced statewide travel framework and estimates of travel demand by using data from multiple existing sources such as California Household Travel Survey, Origin-Destination data from the Statewide Travel Demand Model, and social media data. We developed a short and long-distance travel inventory that includes statewide vehicle miles traveled (VMT). Furthermore, we developed an automated conversion procedure to transform

social media data into origin-destination travel and trip lengths to produce estimates of VMT.

A variety of research directions have emerged from lessons learned in this project. First, with the comparison of Twitter trips and synthetic population, we found walking trips are strongly related to twitter trips, so our immediate next step is to perform in-depth analysis of twitter trip and its relationship with walking trips. It is possible that this relationship is due to different land uses and resident characteristics not captured in the analysis of this project and can explain the relationship. In addition, although 6-month observation was a long period of data collection, it would be better to collect data for more than a year like the California Household Travel Survey. In this way, we can observe the year-long dynamics of travel behavior. Moreover, we envision the creation of an observatory project in which social media data are collected for more than a year. This could provide valuable information for not only Caltrans but also the MPOs. Twitter is used heavily by a segment of the population for which we have limited travel behavior data. This segment includes students residing in group quarters, and social media may be the only currently available source to understand their behavior. Developing a small scale survey that is also informed by social media will provide invaluable information for modeling and simulation of travel behavior for this group. In addition, as a first step we could create a hot spot analysis and identify if many of the trips we estimated in this project have their origins at colleges and universities and then design surveys that target the locations with the highest number of tweets.

WHAT IS THE BENEFIT?

The baseline inventory will provide a quantified reference of travel demand statewide that will include day of the week and month-to-month travel demand variation. This will allow us to identify when an anomaly in travel demand happens and distinguish between induced demand from events that attract people at a specific location for a short amount of time such as a concert and other unpredictable occurrences.

IMAGES

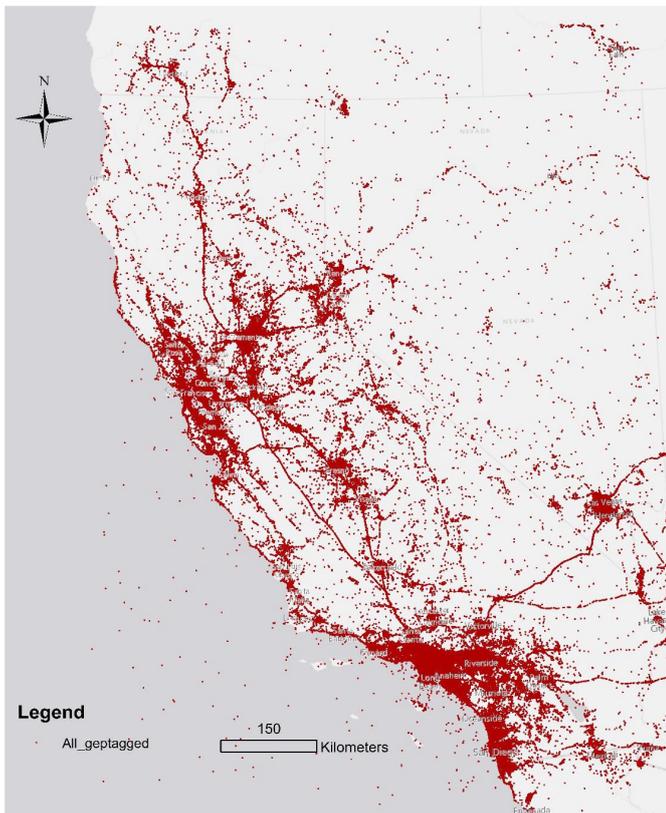


FIGURE 1: All Geo-tagged Tweets

The contents of this document reflect the views of the authors, who are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the California Department of Transportation, the State of California, or the Federal Highway Administration. This document does not constitute a standard, specification, or regulation. No part of this publication should be construed as an endorsement for a commercial product, manufacturer, contractor, or consultant. Any trade names or photos of commercial products appearing in this document are for clarity only.