Using Artificial Intelligence to Improve Traffic Flows with Consideration of Data Privacy Principles

Tyler Reeb, Ph.D., California State University, Long Beach

tyler.reeb@csulb.edu

Co-authors: Gwen Shaffer, Hossein Jula and Anastasios Chassiakos, California State University Long Beach

CALIFORNIA STATE UNIVERSITY LONG BEACH

Project Background

Reducing delays caused by traffic signals is an ongoing challenge for city planners and traffic engineers. In a survey conducted by the National Transportation Operations Coalition, traffic signals account for upwards of 10% of the 295 million vehicle-hours of delay on major roadways.

Artificial Neural Networks (ANN) are a class of Artificial Intelligence (AI) systems that are increasingly being used as tools that can be integrated into traffic signal technologies to enhance delay estimation functionality.

However, in order to “train” operate and modify AI-powered Intelligent Transportation Systems (ITS), the operators of those systems must gather a vast amount of mobility data that raises a range of privacy and civil liberties concerns for community members.

CITT CENTER FOR INTERNATIONAL TRADE AND TRANSPORTATION CSULB

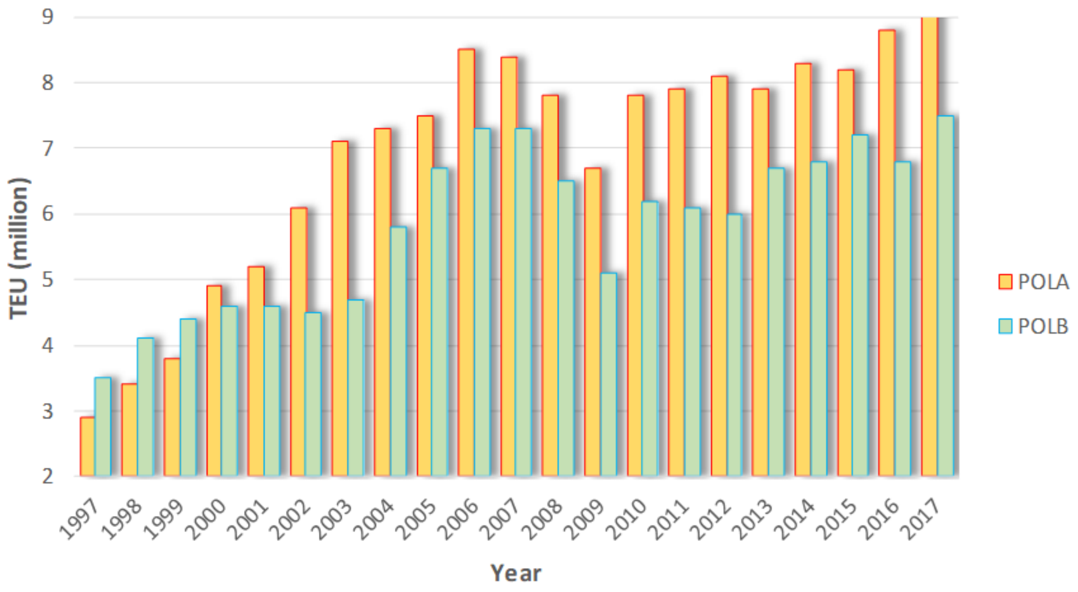
Study Objectives

The objectives for this study address two interrelated concerns: how to efficiently use artificial intelligence technologies to improve traffic flows without violating the data privacy of citizens.

- **AI Objective:** In this project, we explore ways that artificial-neural-network technologies can improve the efficacy of next-generation delay estimators. Our delay estimator will differ from previously observed delay estimators in that we include freight vehicle information as a part of our input parameters. Traffic delays at an intersection can be used as a measure of traffic efficiency at that intersection. The objective of this research is to develop methodologies to predict delays at intersections where increased heavy trucks are present.
- **Data Privacy Objective:** For this project, we led cohorts of study participants on a 1.5-mile community "datawalk" in West Long Beach—adjacent to the second busiest port in the United States—to gain insights into the level of trust residents feel when encountering smart technologies that track their movements and capture their image.

Why Choose Long Beach?

- The ports of Long Beach (POLB) and Los Angeles (POLA) form the largest container port complex in the U.S.
- Heavy trucks trips to and from the ports result in increased traffic delays at intersections around the ports.
- An intersection with a high rate of truck delays is inefficient and will cause increased congestion, noise, and greenhouse gas emissions for the local community.



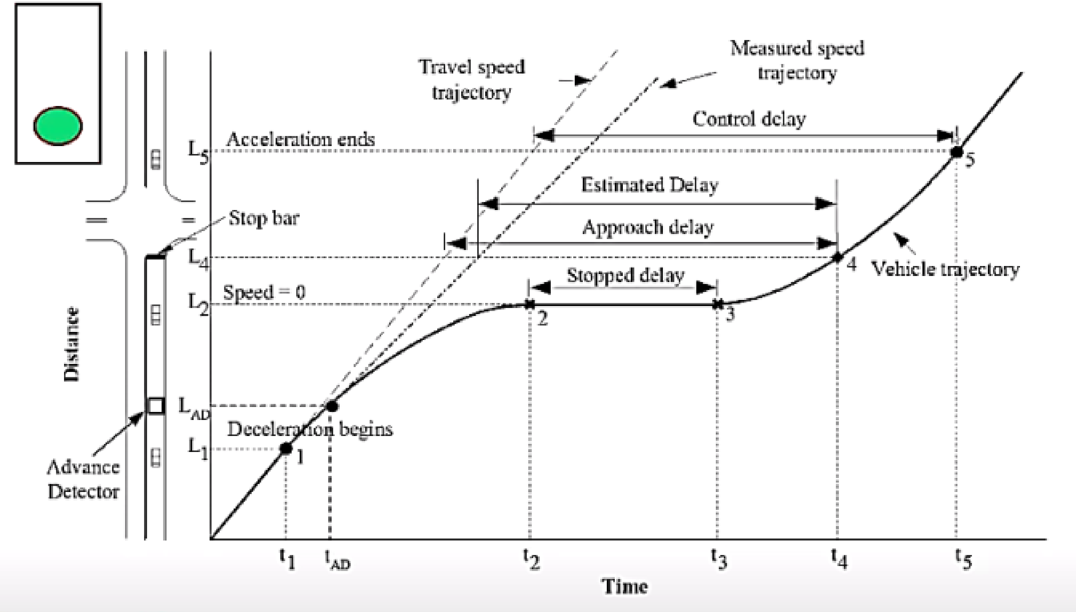
CALIFORNIA STATE UNIVERSITY LONG BEACH

What is a delay?

Types of delays:

- Stopped delay: delay due to complete stop.
- Approach delay: delay due to deceleration and stop time and acceleration.
- Control delay: delay due to deceleration, stop time, acceleration, and reaching the flow speed.

We adopt the control delay as the definition of delay here.



CITT CENTER FOR INTERNATIONAL TRADE AND TRANSPORTATION CSULB

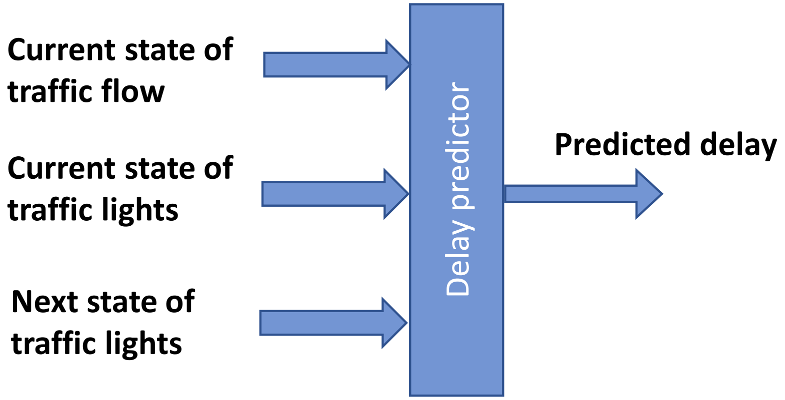
What is a delay predictor?

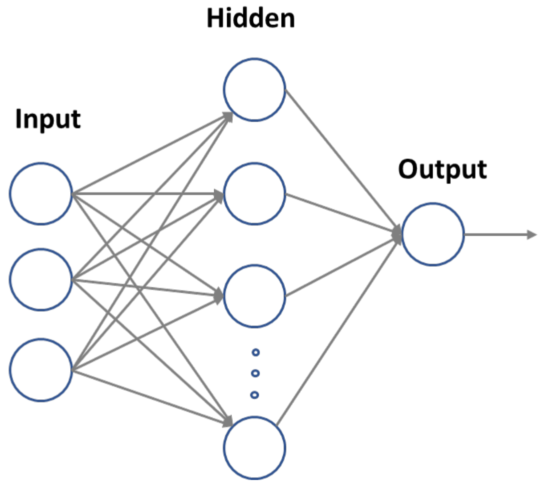
Predicts future delays based on the current state of traffic, current state of traffic lights, and the next state of traffic lights.

Current state of traffic:

- Number of passenger vehicles,
- Number of heavy trucks,
- Speed of passenger vehicles,
- Speed of heavy trucks, and
- Queue lengths.

We designed and developed a delay predictor model based on Artificial Neural Networks (NNs).





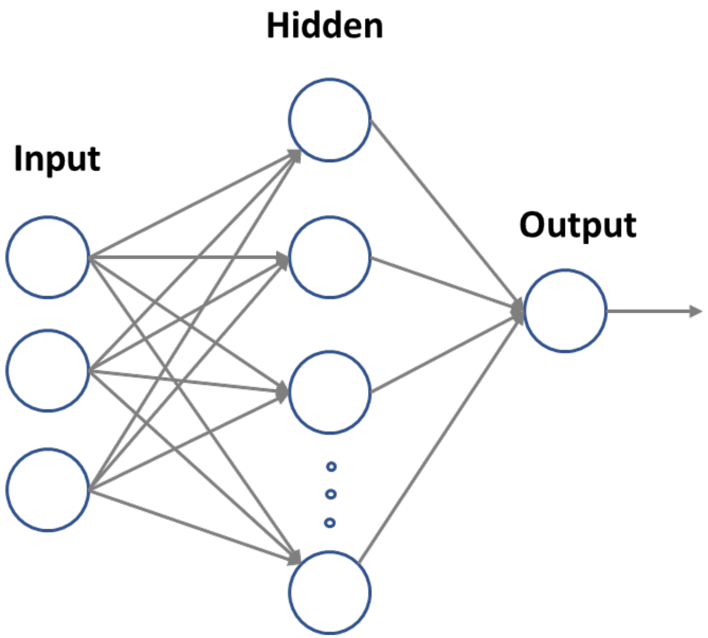
ANN Delay Predictor

After training, the designed ANN behaves like a non-linear time varying function that maps traffic flow and traffic light information to delay

$$D_{t+1} = f(S_t, L_t, L_{t+1}, t)$$

Where:

- D_{t+1} is delay prediction at time $t+1$ based on the information available at time t.
- $S_t = (N_P, V_P, N_T, V_P, Q)$ is the state of traffic at time t.
- L_t state of traffic lights at time t (G/Y/R).



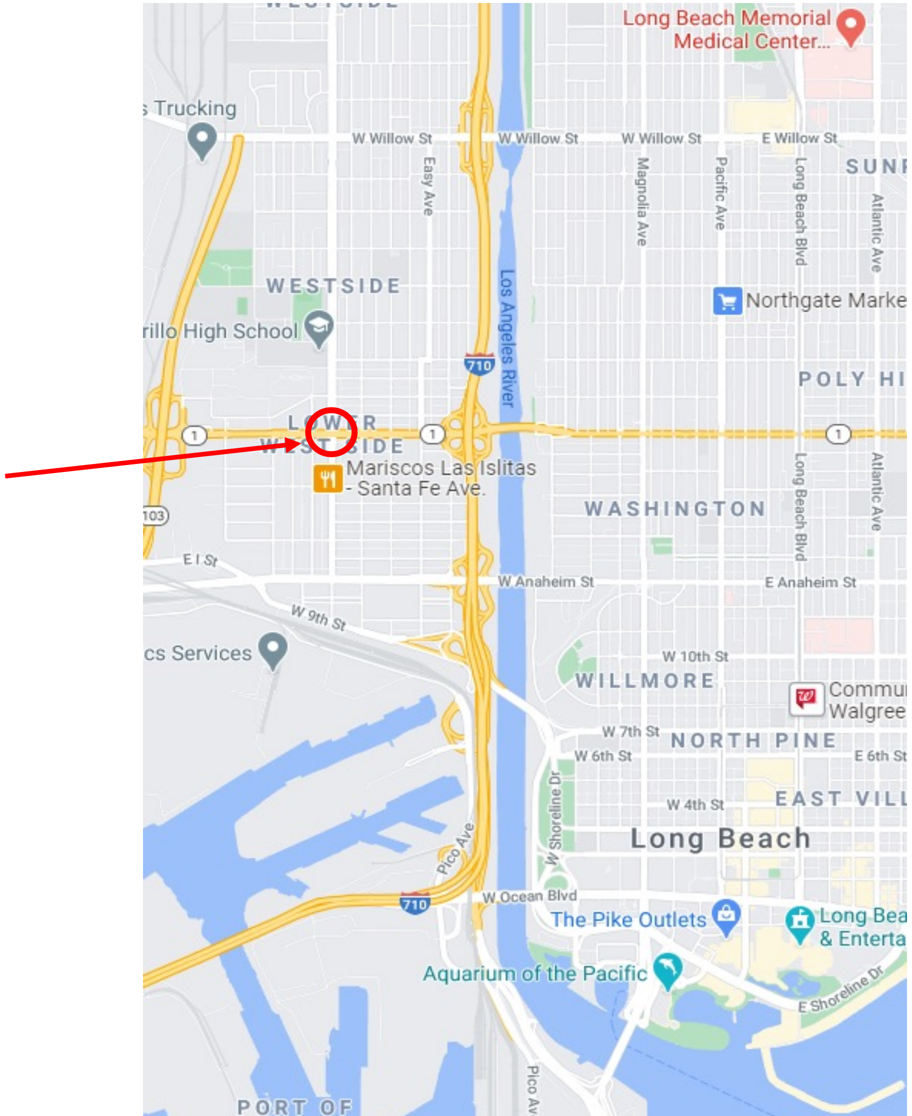
Our NN has one hidden layer with 10 – 12 neurons. We noticed that:

- A single layer is computationally fast.
- Increasing the number of hidden layers doesn’t substantially improve the quality of results.
- We used different training algorithms including Levenberg-Marquardt (LM) or Bayesian.

CALIFORNIA STATE UNIVERSITY
LONG BEACH

Study Objectives

- We have identified the PCH and Santa Fe intersection to collect data.
- The intersection is geographically close to the LB port and has elevated traffic of heavy trucks passing through.
- Our field measured data were used to train, calibrate, and evaluate our NN model.
- We have observed less than 20% difference, in average, between our predicted delays using our developed NN and the actual delays collected from the field.



CITT
CENTER FOR INTERNATIONAL TRADE AND TRANSPORTATION
CSULB

CALIFORNIA STATE UNIVERSITY
LONG BEACH

Data Privacy Objectives

- Most people are aware that social media platforms and corporate websites track their online behavior, analyze that data, store it and potentially share it.
- However, they often fail to realize that local governments also deploy devices and software that collect personally identifiable information.
- For this project, we led 32 study participants on a 1.5 mile community "datawalk" in West Long Beach—adjacent to the second busiest port in the United States—to gain insights into the level of trust residents feel when encountering smart technologies that track their movements and capture their image.

CITT
CENTER FOR INTERNATIONAL TRADE AND TRANSPORTATION
CSULB

Data Privacy Guidelines

- **In March 2021, Long Beach City Council adopted data privacy guidelines. Long Beach will:**
  - be publicly **transparent and accountable** in its collection and management practices of personal data.
  - work to **provide participatory, responsive feedback channels for residents** to inform the City's data collection and usage practices, exercise privacy complaints, and ensure the City is held accountable to these Guidelines.
  - **advance digital equity** and prioritize the needs of marginalized communities on matters pertaining to data and information management.
  - **use data in an ethical and non-discriminatory manner** to not reinforce existing racial biases and prejudiced decision-making.
  - **practice ethical data stewardship** throughout the data lifecycle to minimize misuse of personal data.

Community Data Walks

- Our methodology borrows from the "flashmob ethnography" framework conceived by design researcher Laura Forlano, who focused on identifying and highlighting the role of values in our built environment.
  - Each group member assumed a designated role: navigator/sketcher, note-taker/interviewer and photographer/videographer.
- We also borrow from "data walkshops" designed by Allison Powell, who characterizes her methodology as "a radically bottom-up process" of exploring and defining "data politics" from the perspectives of residents.

Study Objectives

- Datawalks are a powerful tool for connecting questions, concerns or investigations related to data with other dominant social challenges.
- Datawalk participants are exposed to issues ranging from racial inequality and environmental degradation to food deserts.
- Datawalks have the potential to spur changes in the study participant, in the technologies under study and, finally, in the relationship between the two.

CALIFORNIA STATE UNIVERSITY LONG BEACH

Cultures of Trust Framework

- Institutional trust is critical in order to maintain the stability of societies and cultures.
- Previous research suggests that smart city constituents will accept sharing their personal data only if officials establish trusted relationships among residents and visitors, and if those participants have the ability to control the use of their data.

CITT CENTER FOR INTERNATIONAL TRADE AND TRANSPORTATION CSULB

Research Questions

RQ1: What attitudes and trust levels do residents express regarding smart technologies that capture personally identifiable data? Do their attitudes and trust levels vary, depending on whether the technologies are deployed by the city or private entities?

RQ2: Based on the datawalk findings, what implications exist for smart cities deploying technologies that collect, store, analyze and sometimes share personal data about residents?

Study Objectives

Designed a route

The 1.5 mile route took study participants past multiple fast-food restaurants; a liquor store; through a residential neighborhood; into a city park; past a large public high school; and past a Long Beach police station.

Created a mobile app

The app enabled participants to "pin" smart technologies they encountered. Prompts asked walkers about their perceptions of technologies and their comfort levels with them.

Recruited participants

Promoted the July 1, 2021 datawalks on a dozen relevant social media pages. City staff also emailed info to "smart city initiative" stakeholders.

Facilitated 3 walks with 32 people

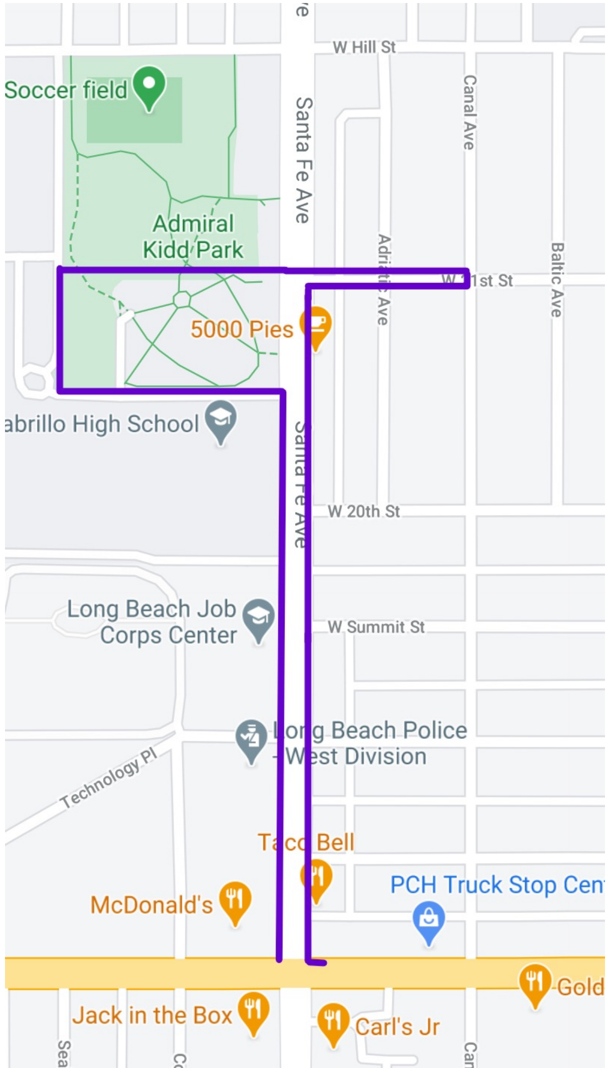
Walkers encountered commercial surveillance cameras, internet-connected bus kiosks/stops, residential surveillance cameras, license plate readers, public WiFi routers and city-owned surveillance cameras

Debriefed participants after walks

Asked if anything they observed during the walk surprised them; whether the city is transparent about data collection; whether the tech made them feel safe or violated; and whether the benefits outweigh privacy risks.

CALIFORNIA STATE UNIVERSITY
LONG BEACH

Community Data Walk Location



CITT
CENTER FOR INTERNATIONAL
TRADE AND TRANSPORTATION
CSULB

Findings

- Surveillance cameras:
- Comfort levels varied depending on whether participants encountered city-owned traffic cameras, or surveillance cameras installed by businesses or homeowners.

- Free wifi routers:
- Participants disliked being auto connected and questioned how private businesses and the city use their phone ID and analyze online behavior.

- Smart bus kiosk/apps:
- Participants appreciated tech that encourage/simplifies riding mass transit, but did not support LB Transit retaining or selling personally identifiable information..

- License plate readers:
- Concerns centered on how the LBPD stores and shares data collected through the tech, and the disproportionate impact on marginalized communities.

Discussions and Analysis

- Participants recognize the potential for law enforcement technologies to harm people of color and low-income communities, as well as for LBPD to "overreach."
- In many cases, participants expressed comfort with a particular device or platform, but articulated concerns about the retention and use of data collected by the technology. Comments included:

"Where do the data live?"

"I'm nervous about with whom the data are shared."

"I'm comfortable with the technology. However, I have concerns about the retention and use of that data."

"Not okay to store my data! I'm a law- abiding citizen!"

Discussion and Analysis

- Participants said the walk heightened their awareness of just how ubiquitous smart technologies are in the city.
- For instance, they realized that police vehicles are equipped with license plate readers, but not that this technology is also installed in vehicles belonging to the Parking Enforcement fleet. Or that residents encounter dozens of surveillance cameras when walking a few blocks.
- By encountering smart technologies within the context of an organized datawalk, study participants felt more aware of and more curious about technologies in their built environment:

"Is that a camera? Are we being recorded while we wait for the bus?"

"What is the weird box on the pole at the bus stop?"

"I've heard and read about this tech, and I find it unavoidable. It's interesting to identify them around the city, though."

Discussion and Analysis

- Smart cities must have a clear justification for why they collect specific types of data and articulate these reasons to the public (i.e. through annual privacy reports).
- Residents should play a role in determining policies for data collection, retention, access and use (i.e., through focus groups, public forums and similar in-person events).
- When possible, residents should be notified when the city is collecting personally identifiable data (i.e., signage posted near surveillance cameras, clear disclosure when installing apps).
- Local officials should not use personal data for secondary purposes without explicit buy-in from residents (i.e., recordings from a gunshot detection system installed on streetlights used to record street-level conversations).

Limitations of the Study

- About 63 percent of participants identified as white. Therefore, the study sample is not representative of Long Beach's population.
- Datawalks should be facilitated in communities citywide, beyond West Long Beach.
- More participants are necessary to confirm the findings and inform policy recommendations.

Big Takeaways

- Don’t underestimate the importance of community support for artificial intelligence and Intelligent Transportation Systems.
- Implementing real world mobility labs requires a great deal of coordination with local, state, and federal layers of government.
- It takes a multidisciplinary team to implement the mobility test labs of the future.

Thank you!

Tyler Reeb, Ph.D., California State University, Long Beach

tyler.reeb@csulb.edu

Co-authors: Gwen Shaffer, Hossein Jula and Anastasios Chassiakos, California State University Long Beach