New England University Transportation Center
77 Massachusetts Avenue, E40-279
Cambridge, MA 02139
617.253.0753
utc.mit.edu

New England University Transportation Center

Year 24 Final Report

Grant Number: DTRT12-G-UTC01

Project Title:

Making More Value out of Transportation Data

Project Number:	Project End Date:	Submission Date:
UMAR24-23	8/31/15	12/31/15

Principal Investigator:	Daiheng Ni	**Co-Principal Investigator:**	Lixin Gao
Title:	Associate Professor	**Title:**	Professor
University:	University of Massachusetts/Amherst	**University:**	University of Massachusetts/Amherst
Email:	ni@engin.umass.edu	**Email:**	lgao@ecs.umass.edu
Phone:	413.545-5408	**Phone:**	413.545.4548

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or the use thereof.

The New England University Transportation Center is a consortium of 5 universities funded by the U.S. Department of Transportation, University Transportation Centers Program. Members of the consortium are MIT, the University of Connecticut, the University of Maine, the University of Massachusetts, and Harvard University. MIT is the lead university.

Over the past decades, tremendous investments have been made to advance technologies in computing, sensing, and communication to collect, process, and disseminate transportation related data. As such, transportation system operators and users have timely information to make educated decisions to improve safety and mitigate congestion. These investments can become more productive if the following problems can be properly addressed: (1) effective algorithms to process huge volume of transportation data, (2) creative approaches to merge a variety of transportation data, (3) advanced tools to increase the velocity of information retrieving, processing, and delivering. This research aspires to explore solutions to these problems and the effort will be facilitated by the Regional Traveler Information Center (RTIC) at UMass Amherst. As a starting point, this research will focus on two basic problems: (1) variety: developing effective algorithms to convert traffic data in different forms so that these data can be merged to provide complementary perspectives on the real system, and (2) velocity: developing applications to disseminate transportation information in a timely fashion to enable better decision-making by road users, system operators, and emergency responders..

Road traffic conditions are typically affected by past traffic condition and events like extreme weather and sport games. With the advance of Web, events and weather conditions can be readily retrieved in real-time. Large amount of data incorporating both traffic and contextual information can be gathered.

This research used field data collected from RTIC to implement the research outcome. This set of data consisted of 3.5 years' worth of information. In addition to traffic data and their associated timing information, three kinds of online information are collected. The contextual online information includes weather, local events, such as big sports events, musical concerts and graduation ceremonies, and local special days like storm days, legal holidays, university closing days and so on. Traffic information using E-ZPass RFID data in western Massachusetts is collected from January 1, 2010 to June 30, 2013. The E-ZPass system is a non-contact automatically identification system which consists of mobile tags and fixed readers. When a vehicle equipped with a E-ZPass tag passes a reader, current time, reader ID and tag ID are recorded. As shown in Figure 1, four sensors shown with red numbers monitor two unconnected rural highways. These two roads are the most important connections among 5 higher education institutions including UMass Amherst. According to UMass Transit Information Center the flow rate is about 960 veh/hr in rush hour and 480 veh/hr for average, at the same time about 8% vehicles passed by those readers are equipped with EZPass tags.

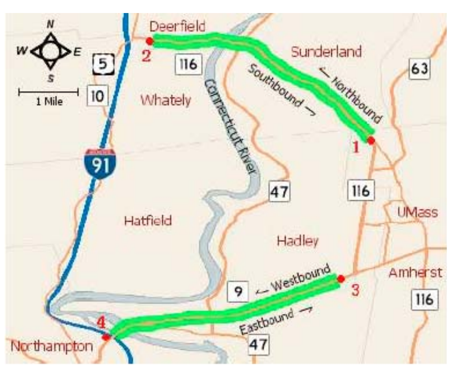

Figure 1. Test site

In this research, we propose a traffic condition prediction system taking advantage of big data processing technology. RFID-based system has been deployed for monitoring road traffic. By incorporating data from both road traffic monitoring system and online contextual information, we propose a hierarchical Bayesian network to predict road traffic condition.

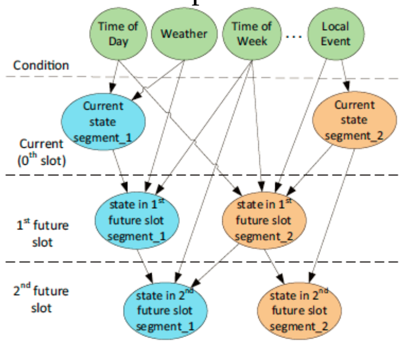

Figure 2. An example Bayesian network

The principle of this approach is the following. Theoretically after incorporating the current traffic information, the traffic system can be modeled with causal graphical models. In this model, traffic conditions on each road link in each time window and other contextual factors are modeled as nodes and the causal relationships among different factors are modeled as direct edges from reason nodes to result nodes. Different states of each factor are represented by the different values of corresponding node. There is a hierarchical property on this predicting system. According to law of causality, at any time point a current event is impossible to be affect by future events and other events happening simultaneously. So nodes representing traffic condition of different time can be separated into several groups. As a simple example shown in Figure 2, when arranged with temporal ordering as Hierarchical Bayesian Network (HBN). Using this model structure, we predict the traffic conditions in next 2 time units.

The result of implementation of the proposed model is presented in comparison with a conventional one called naive model, see Figure 3. The major difference of the two models is what kind of information the prediction is based on. So in the figure the lines of the naive model are denoted by "Context" and the lines of the HBN model are denoted by "Context+Traffic". Using the naive model, the accuracies on normal and congested conditions are 92% and 19%. Incorporated with current traffic condition and applied with the HBN model, the accuracies rise to 94% on normal cases and 72% on congested cases. The accuracy of predicting congested cases is improved about 3.8 times from 19% to 72% after taking the current traffic into consideration while the accuracy of predicting normal cases improves a little. Analysis on the trained HBN model shows that bad traffic conditions have significant influence on the other bad conditions, especially that of the same road segment.

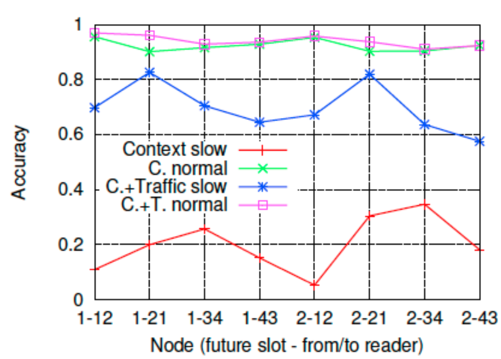

Figure 3. Implementation result