Efficient Transport Simulation With Restricted Batch-Mode Active Learning

Simulation modeling is a well-known and recurrent approach to study the performance of urban systems. Taking into account the recent and continuous transformations within increasingly complex and multidimensional cities, the use of simulation tools is, in many cases, the only feasible and reliable approach to analyze such dynamic systems. However, simulation models can become very time consuming when detailed input-space exploration is needed. To tackle this problem, simulation metamodels are often used to approximate the simulators’ results. In this paper, we propose an active learning algorithm based on the Gaussian process (GP) framework that gathers the most informative simulation data points in batches, according to both their predictive variances and to the relative distance between them. This allows us to explore the simulators’ input space with fewer data points and in parallel, and thus in a more efficient way, while avoiding computationally expensive simulation runs in the process. We take advantage of the closeness notion encoded into the GP to select batches of points in such a way that they do not belong to the same high-variance neighborhoods. In addition, we also suggest two simple and practical user-defined stopping criteria so that the iterative learning procedure can be fully automated. We illustrate this methodology using three experimental settings. The results show that the proposed methodology is able to improve the exploration efficiency of the simulation input space in comparison with non-restricted batch-mode active learning procedures.
Multi-Output Gaussian Processes for Crowdsourced Traffic Data Imputation

Traffic speed data imputation is a fundamental challenge for data-driven transport analysis. In recent years, with the ubiquity of GPS-enabled devices and the widespread use of crowdsourcing alternatives for the collection of traffic data, transportation professionals increasingly look to such user-generated data for a good deal of analysis, planning, and decision support applications. However, due to the mechanics of the data collection process, crowdsourced traffic data such as probe-vehicle data is highly prone to missing observations, making accurate imputation crucial for the success of any application that makes use of that type of data. In this paper, we propose the use of multi-output Gaussian processes (GPs) to model the complex spatial and temporal patterns in crowdsourced traffic data. While the Bayesian nonparametric formalism of GPs allows us to model observation uncertainty, the multi-output extension based on convolution processes effectively enables us to capture complex spatial dependencies between nearby road segments. Using six months of crowdsourced traffic speed data or "probe vehicle data" for several locations in Copenhagen, the proposed approach is empirically shown to significantly outperform popular state-of-the-art imputation methods.

General information
State: Accepted/In press
Organisations: Technical University of Denmark, Department of Management Engineering, Transport Modelling, University of Washington
Authors: Rodrigues, F. (Intern), Henrickson, K. (Ekstern), Pereira, F. C. (Intern)
Number of pages: 10
Publication date: 28 Apr 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: IEEE Transactions on Intelligent Transportation Systems
ISSN (Print): 1524-9050
Ratings:
Web of Science (2018): Indexed yes
Scopus rating (2017): CiteScore 5.14
Web of Science (2017): Impact factor 4.051
Web of Science (2017): Indexed yes
Scopus rating (2016): CiteScore 4.44
Web of Science (2016): Impact factor 3.724
Web of Science (2016): Indexed yes
Scopus rating (2015): CiteScore 4.64
Web of Science (2015): Impact factor 2.534
Scopus rating (2014): CiteScore 4.16
Web of Science (2014): Impact factor 2.377
Scopus rating (2013): CiteScore 4.99
Web of Science (2013): Impact factor 2.472

10.1109/TITS.2018.2842695
Combining time-series and textual data for taxi demand prediction in event areas: a deep learning approach

Accurate time-series forecasting is vital for numerous areas of application such as transportation, energy, finance, economics, etc. However, while modern techniques are able to explore large sets of temporal data to build forecasting models, they typically neglect valuable information that is often available under the form of unstructured text. Although this data is in a radically different format, it often contains contextual explanations for many of the patterns that are observed in the temporal data. In this paper, we propose two deep learning architectures that leverage word embeddings, convolutional layers and attention mechanisms for combining text information with time-series data. We apply these approaches for the problem of taxi demand forecasting in event areas. Using publicly available taxi data from New York, we empirically show that by fusing these two complementary cross-modal sources of information, the proposed models are able to significantly reduce the error in the forecasts.

General information
State: Accepted/In press
Organisations: Department of Management Engineering, Transport Modelling
Authors: Rodrigues, F. (Intern), Markou, I. (Intern), Pereira, F. C. (Intern)
Number of pages: 20
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: Information Fusion
ISSN (Print): 1566-2535
Ratings:
BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 7.23 SNIP 3.382 SJR 1.832
Web of Science (2017): Impact factor 6.639
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 5.99 SNIP 3.369 SJR 1.679
Web of Science (2016): Impact factor 5.667
BFI (2015): BFI-level 1
Scopus rating (2015): CiteScore 4.82 SNIP 3.607 SJR 1.586
BFI (2014): BFI-level 1
Scopus rating (2014): CiteScore 4.6 SNIP 4.101 SJR 1.27
Web of Science (2014): Impact factor 3.681
BFI (2013): BFI-level 1
Scopus rating (2013): CiteScore 3.74 SNIP 4.052 SJR 0.647
Web of Science (2013): Impact factor 3.472
BFI (2012): BFI-level 1
Scopus rating (2012): CiteScore 3.06 SNIP 4.048 SJR 1.151
Web of Science (2012): Impact factor 2.262
Deep Learning from Crowds

Over the last few years, deep learning has revolutionized the field of machine learning by dramatically improving the state-of-the-art in various domains. However, as the size of supervised artificial neural networks grows, typically so does the need for larger labeled datasets. Recently, crowdsourcing has established itself as an efficient and cost-effective solution for labeling large sets of data in a scalable manner, but it often requires aggregating labels from multiple noisy contributors with different levels of expertise. In this paper, we address the problem of learning deep neural networks from crowds. We begin by describing an EM algorithm for jointly learning the parameters of the network and the reliabilities of the annotators. Then, a novel general-purpose crowd layer is proposed, which allows us to train deep neural networks end-to-end, directly from the noisy labels of multiple annotators, using only backpropagation. We empirically show that the proposed approach is able to internally capture the reliability and biases of different annotators and achieve new state-of-the-art results for various crowdsourced datasets across different settings, namely classification, regression and sequence labeling.

Event characteristics that disrupt transport system’s balance

The life of the city is often reflected in traffic patterns: popular sporting events draw crowds, holidays create disruptions, protests may result in road closures, etc. Decades of research on travel demand and network modelling already provide satisfying predictive tools. However, the main research focus has been on regular behaviour, such as peak/off-peak cycles, regular functioning of the infrastructure, and normal weather conditions. Consequently, non-recurrent events severely challenge such models. Under non-recurrent circumstances, the typically expected correlation structures (e.g. between demand flows in neighbor areas; between current and recent values of traffic speeds or travel times) are
drastically affected, severely affecting predictions. It is therefore necessary to take into consideration data from different sources. The objective of this research is the development of a methodology that correlates high taxi demand observations with popular events retrieved from Social Media platforms. Using NYC taxi trips public dataset, the average demand of the day was determined using kernel density analysis. Days that showed significant outliers compared to the average day were further studied using a dataset of around 116000 events. The second dataset was retrieved from the Web for the same 6 months period through the direct use of APIs. The correlation step includes the comparison of spatial and temporal kernel density depiction of taxi pick-up locations and events retrieved details. Through the correlation evaluation of traffic data and semantic information, conclusions were made on how the demand of taxi pick-ups changes based on certain event characteristics.

**General information**
State: Submitted
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling
Authors: Markou, I. (Intern), Rodrigues, F. (Intern), Pereira, F. C. (Intern)
Number of pages: 10
Publication date: 2018

**Host publication information**
Title of host publication: Transportation Research Procedia
Main Research Area: Technical/natural sciences
Conference: International Conference on Intelligent Transport Systems in Theory and Practice, mobil. TUM 2017, Munich, Germany, 04/07/2017 - 04/07/2017
Source-ID: 142689102
Publication: Research - peer-review › Article in proceedings – Annual report year: 2018

**Heteroscedastic Gaussian processes for uncertainty modeling in large-scale crowdsourced traffic data**
Accurately modeling traffic speeds is a fundamental part of efficient intelligent transportation systems. Nowadays, with the widespread deployment of GPS-enabled devices, it has become possible to crowdsource the collection of speed information to road users (e.g. through mobile applications or dedicated in-vehicle devices). Despite its rather wide spatial coverage, crowdsourced speed data also brings very important challenges, such as the highly variable measurement noise in the data due to a variety of driving behaviors and sample sizes. When not properly accounted for, this noise can severely compromise any application that relies on accurate traffic data. In this article, we propose the use of heteroscedastic Gaussian processes (HGP) to model the time-varying uncertainty in large-scale crowdsourced traffic data. Furthermore, we develop a HGP conditioned on sample size and traffic regime (SSRC-HGP), which makes use of sample size information (probe vehicles per minute) as well as previous observed speeds, in order to more accurately model the uncertainty in observed speeds. Using 6months of crowdsourced traffic data from Copenhagen, we empirically show that the proposed heteroscedastic models produce significantly better predictive distributions when compared to current state-of-the-art methods for both speed imputation and short-term forecasting tasks.

**General information**
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling
Authors: Rodrigues, F. (Intern), Pereira, F. C. (Intern)
Pages: 636-651
Publication date: 2018
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Transportation Research. Part C: Emerging Technologies
Volume: 95
ISSN (Print): 0968-090X
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 5.17 SJR 2.293 SNIP 2.907
Web of Science (2017): Impact factor 3.968
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 4.43 SJR 1.998 SNIP 2.638
Web of Science (2016): Impact factor 3.805
Overview of traffic incident duration analysis and prediction
Non-recurrent congestion caused by traffic incident is difficult to predict but should be dealt with in a timely and effective manner to reduce its influence on road capacity reduction and enormous travel time loss. Influence factor analysis and reasonable prediction of traffic incident duration are important in traffic incident management to predict incident impacts and aid in the implementation of appropriate traffic operation strategies. The objective of this study is to conduct a thorough review and discusses the research evolution, mainly including the different phases of incident duration, data resources, and the various methods that are applied in the traffic incident duration influence factor analysis and duration time prediction. In order to achieve the goal of this study, we presented a systematic review of traffic incident duration time estimation and prediction methods developed based on various data resources, methodologies etc. based on the previous
studies, we analyse (i) Data resources and characteristics: different traffic incident time phases, data set size, incident
types, duration time distribution, available data resources, significant influence factors and unobserved heterogeneity and
randomness, (ii) traffic incident duration analysis methods, mainly including hazard-based duration model and regression
and statistical tests, (iii) traffic incident duration prediction methods and evaluation of prediction accuracy. After a
comprehensive review of literature, this study identifies and analyses future challenges and what can be achieved in the
future to estimate and predict the traffic incident duration time.

General information
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, Tsinghua University,
Massachusetts Institute of Technology
Authors: Li, R. (Ekstern), Pereira, F. C. (Intern), Ben-Akiva, M. E. (Ekstern)
Number of pages: 13
Publication date: 2018
Main Research Area: Technical/natural sciences

Publication information
Journal: European Transport Research Review
Volume: 10
Issue number: 2
Article number: 22
ISSN (Print): 1867-0717
Ratings:
Web of Science (2018): Indexed yes
Scopus rating (2017): CiteScore 1.67 SJR 0.51 SNIP 1.185
Web of Science (2017): Impact factor 1.758
Web of Science (2017): Indexed yes
Scopus rating (2016): SJR 0.419 SNIP 1.002 CiteScore 1.15
Web of Science (2016): Impact factor 0.962
Scopus rating (2015): SJR 0.394 SNIP 0.651 CiteScore 0.95
Web of Science (2015): Impact factor 0.672
Scopus rating (2014): SJR 0.425 SNIP 0.91 CiteScore 0.74
Scopus rating (2013): SJR 0.421 SNIP 0.781 CiteScore 0.75
ISI indexed (2013): ISI indexed no
Scopus rating (2012): SJR 0.26 SNIP 0.653 CiteScore 0.62
ISI indexed (2012): ISI indexed no
Scopus rating (2011): SJR 0.113 SNIP 0.665 CiteScore 0.42
ISI indexed (2011): ISI indexed no
Scopus rating (2010): SJR 0.164 SNIP 0.886
Original language: English
Incident duration analysis, Traffic incident duration prediction, Hazard-based duration model, Data mining, Influence
factors
Electronic versions:
filestore_2_.pdf
DOIs:
10.1186/s12544-018-0300-1

Bibliographical note
© The Author(s). 2018 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0
International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and
reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to
the Creative Commons license, and indicate if changes were made.
Source: FindIt
Source-ID: 2435074298
Publication: Research - peer-review › Journal article – Annual report year: 2018

Social network analysis in future transportation systems: Contributions on observability, behaviour and structure

General information
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, Technical University of
Munich, University of Cyprus
A Bayesian Additive Model for Understanding Public Transport Usage in Special Events

Public special events, like sports games, concerts and festivals are well known to create disruptions in transportation systems, often catching the operators by surprise. Although these are usually planned well in advance, their impact is difficult to predict, even when organisers and transportation operators coordinate. The problem highly increases when several events happen concurrently. To solve these problems, costly processes, heavily reliant on manual search and personal experience, are usual practice in large cities like Singapore, London or Tokyo. This paper presents a Bayesian additive model with Gaussian process components that combines smart card records from public transport with context information about events that is continuously mined from the Web. We develop an efficient approximate inference algorithm using expectation propagation, which allows us to predict the total number of public transportation trips to the special event areas, thereby contributing to a more adaptive transportation system. Furthermore, for multiple concurrent event scenarios, the proposed algorithm is able to disaggregate gross trip counts into their most likely components related to specific events and routine behavior. Using real data from Singapore, we show that the presented model outperforms the best baseline model by up to 26 percent in R-2 and also has explanatory power for its individual components.

General information
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, Singapore-MIT Alliance for Research and Technology, University of Coimbra
Authors: Rodrigues, F. (Intern), Borysov, S. S. (Ekstern), Ribeiro, B. (Ekstern), Pereira, F. C. (Intern)
Pages: 2113-2126
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: I E E E Transactions on Pattern Analysis and Machine Intelligence
Volume: 39
Issue number: 11
ISSN (Print): 0162-8828
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 12.75 SJR 2.367 SNIP 6.357
Web of Science (2017): Impact factor 9.455
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 13.29 SJR 5.388 SNIP 6.403
Web of Science (2016): Impact factor 8.329
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 3.475 SNIP 7.634 CiteScore 11.05
Web of Science (2014): Impact factor 5.781
BFI (2013): BFI-level 2
Due to environmental and economic stress, strong investment exists now towards adaptive transport systems that can efficiently utilize capacity, minimizing costs and environmental impacts. The common vision is a system that dynamically changes itself (the supply) to anticipate traveler needs (the demand). In some occasions, unexpected and unwanted demand patterns are noticed in the traffic network that lead to system failures and cost implications. Significantly low speeds or excessively low flows at an unforeseeable time are only some of the phenomena that are often noticed and need to be explained for transport system’s better future response. The objective of this research is the formulation of a proper methodology that identifies anomalies on traffic networks and correlates them with special events using internet data. Our main subject of interest is the investigation of why traffic congestion is happening as well as why there are demand fluctuations in days were there are no apparent reasons for the occurrence of such phenomena. We evaluated our system using Google’s NYC taxi trips public dataset. We defined initially the “normality” baseline and thereunder we studied individual days’ demand patterns for outliers’ detection. Our approach enabled us to detect demand fluctuations, analyze and correlate them with disruptive events scenarios like extreme weather conditions, public holidays, religious festivities and parades. Using kernel density analysis, the affected areas as well as the significance of the observed differences compared to the average day are depicted.
Enabling Bus Transit Service Quality Co-Monitoring Through Smartphone-Based Platform

The growing ubiquity of smartphones offers public transit agencies an opportunity to transform ways to measure, monitor, and manage service performance. The potential of a new tool is demonstrated for engaging customers in measuring satisfaction and co-monitoring [Editor’s note: This is the authors’ word, meaning “agencies using public feedback to supplement official monitoring and regulation.”] bus service quality. The pilot project adapted a smartphone-based travel survey system, Future Mobility Sensing, to collect real-time customer feedback and objective operational measurements on specific bus trips. The system used a combination of GPS, Wi-Fi, Bluetooth, and accelerometer data to track transit trips while soliciting users’ feedback on trip experience. Though not necessarily intended to replace traditional monitoring channels and processes, these data can complement official performance monitoring through a more real-time, customer-centric perspective. The pilot project operated publicly for 3 months on the Silver Line bus rapid transit in Boston, Massachusetts. Seventy-six participants completed the entrance survey; half of them actively participated and completed more than 500 questionnaires while on board either at the end of a trip, at the end of a day, or both. Participation was biased toward frequent Silver Line users, the majority of whom were white and of higher income. Indicative models of user-reported satisfaction reveal some interesting relationships, but the models can be improved by fusing the app-collected data with actual performance characteristics. Broader and more sustained user engagement remains a critical future challenge.
Enhancing Resilience to Disasters using Social Media

During the last decade, Social Media (SM) have emerged as a prominent trend in social communication, with online platforms such as Facebook and Twitter to conquer the internet space with millions of visitors per day. SM usage generates an astonishing amount of information, which could be used for scarcely experienced situations, such as mass convergence and emergency events. This study presents a preliminary exploratory analysis on examining the capacity of Social Media to extract information on individuals choices during evacuation. We collect tweets from the evacuation in Oroville, California USA due to danger of flood and the evacuation. The data is used for the creation of a user sample which allows the collection of historical data. The historical data is compared with the data collected during and after the evacuation. The goal of this comparison is the extraction of potential information related to the evacuation.
Learning Supervised Topic Models for Classification and Regression from Crowds

The growing need to analyze large collections of documents has led to great developments in topic modeling. Since documents are frequently associated with other related variables, such as labels or ratings, much interest has been placed on supervised topic models. However, the nature of most annotation tasks, prone to ambiguity and noise, often with high volumes of documents, deem learning under a single-annotator assumption unrealistic or unpractical for most real-world applications. In this article, we propose two supervised topic models, one for classification and another for regression problems, which account for the heterogeneity and biases among different annotators that are encountered in practice when learning from crowds. We develop an efficient stochastic variational inference algorithm that is able to scale to very large datasets, and we empirically demonstrate the advantages of the proposed model over state-of-the-art approaches.
Probabilistic Modeling and Visualization for Bankruptcy Prediction

In accounting and finance domains, bankruptcy prediction is of great utility for all of the economic stakeholders. The challenge of accurate assessment of business failure prediction, specially under scenarios of financial crisis, is known to be complicated. Although there have been many successful studies on bankruptcy detection, seldom probabilistic approaches were carried out. In this paper we assume a probabilistic point-of-view by applying Gaussian Processes (GP) in the context of bankruptcy prediction, comparing it against the Support Vector Machines (SVM) and the Logistic Regression (LR). Using real-world bankruptcy data, an in-depth analysis is conducted showing that, in addition to a probabilistic interpretation, the GP can effectively improve the bankruptcy prediction performance with high accuracy when compared to the other approaches. We additionally generate a complete graphical visualization to improve our understanding of the different attained performances, effectively compiling all the conducted experiments in a meaningful way. We complete our study with an entropy-based analysis that highlights the uncertainty handling properties provided by the GP, crucial for prediction tasks under extremely competitive and volatile business environments.

General information
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, University of Coimbra
Authors: Antunes, F. (Ekstern), Ribeiro, B. (Ekstern), Pereira, F. C. (Intern)
Pages: 831-843
Use of Taxi-Trip Data in Analysis of Demand Patterns for Detection and Explanation of Anomalies

Because of environmental and economic stress, current strong investment in adaptive transport systems can efficiently use capacity, minimizing costs and environmental impacts. The common vision is of a system that dynamically changes itself (the supply) to anticipate the needs of travelers (the demand). In some occasions, unexpected and unwanted demand patterns are noticed in the traffic network; these patterns lead to system failures and cost implications. Significantly, low speeds or excessively low flows at an unforeseeable time are only some of the phenomena that are often noticed and need to be explained for a transport system to develop a better future response. The objective of this research was the formulation of a methodology that could identify anomalies on traffic networks and correlate them with special events by using Internet data. The main subject of interest in this study was the investigation of why traffic congestion was occurring as well as why demand fluctuated on days when there were no apparent reasons for such phenomena. The system was evaluated by using Google’s public data set for taxi trips in New York City. A “normality” baseline was defined at the outset and then used in the subsequent study of the demand patterns of individual days to detect outliers. With the use of this approach it was possible to detect fluctuations in demand and to analyze and correlate them with disruptive event scenarios such as extreme weather conditions, public holidays, religious festivities, and parades. Kernel density analysis was used so that the affected areas, as well as the significance of the observed differences compared with the average day, could be depicted.

General information
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling
Authors: Markou, I. (Intern), Rodrigues, F. (Intern), Pereira, F. C. (Intern)
Pages: 129-138
Publication date: 2017
Main Research Area: Technical/natural sciences

Publication information
Journal: Transportation Research Record
Volume: 2643
ISSN (Print): 0361-1981
Ratings:
BFI (2018): BFI-level 1
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 1
Scopus rating (2017): CiteScore 0.83 SJR 0.589 SNIP 0.708
Web of Science (2017): Impact factor 0.695
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 0.75 SJR 0.557 SNIP 0.81
Web of Science (2016): Impact factor 0.592
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.612 SNIP 0.821 CiteScore 0.6
Web of Science (2015): Impact factor 0.522
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 0.562 SNIP 0.876 CiteScore 0.58
Web of Science (2014): Impact factor 0.544
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 0.635 SNIP 0.958 CiteScore 0.76
Web of Science (2013): Impact factor 0.556
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 0.573 SNIP 1.062 CiteScore 0.6
Web of Science (2012): Impact factor 0.442
I E E E Intelligent Systems

Mapping Social Media for Transportation Studies

General information
State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, Technical University of Munich
Authors: Chaniotakis, E. (Ekstern), Antoniou, C. (Ekstern), Pereira, F. C. (Intern)
Number of pages: 7
Pages: 64-70
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Journal: I E E E Intelligent Systems
Volume: 31
Issue number: 6
ISSN (Print): 1541-1672
Ratings:
BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
Arrival time predictions for the next available bus or train are a key component of modern traveler information systems (TISs). A great deal of research has been conducted within the intelligent transportation system community in developing an assortment of different algorithms that seek to increase the accuracy of these predictions. However, the inherent stochastic and nonlinear nature of these systems, particularly in the case of bus transport, means that these predictions suffer from variable sources of error, stemming from variations in weather conditions, bus bunching, and numerous other...
sources. In this paper, we tackle the issue of uncertainty in bus arrival time predictions using an alternative approach. Rather than endeavor to develop a superior method for prediction, we take existing predictions from a TIS and treat the algorithm generating them as a black box. The presence of heteroscedasticity in the predictions is demonstrated and then a metamodel approach is deployed, which augments existing predictive systems using quantile regression to place bounds on the associated error. As a case study, this approach is applied to data from a real-world TIS in Boston. This method allows bounds on the predicted arrival time to be estimated, which give a measure of the uncertainty associated with the individual predictions. This represents to the best of our knowledge the first application of methods to handle the uncertainty in bus arrival times that explicitly takes into account the inherent heteroscedasticity. The metamodel approach is agnostic to the process generating the predictions, which ensures the methodology is implementable in any system.

**General information**

State: Published
Organisations: Department of Management Engineering, Transport DTU, Transport Modelling, Singapore-MIT Alliance, Massachusetts Institute of Technology, Northeastern University
Authors: O'Sullivan, A. (Ekstern), Pereira, F. C. (Intern), Zhao, J. (Ekstern), Koutsopoulos, H. N. (Ekstern)
Number of pages: 11
Pages: 3286-3296
Publication date: 2016
Main Research Area: Technical/natural sciences

**Publication information**

Journal: IEEE Transactions on Intelligent Transportation Systems
Volume: 17
Issue number: 11
ISSN (Print): 1524-9050
Ratings:

BFI (2018): BFI-level 2
Web of Science (2018): Indexed yes
BFI (2017): BFI-level 2
Scopus rating (2017): CiteScore 5.14
Web of Science (2017): Impact factor 4.051
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 4.44
Web of Science (2016): Impact factor 3.724
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 4.64
Web of Science (2015): Impact factor 2.534
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 4.16
Web of Science (2014): Impact factor 2.377
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 4.99
Web of Science (2013): Impact factor 2.472
BFI (2012): BFI-level 2
Scopus rating (2012): CiteScore 4.4
Web of Science (2012): Impact factor 3.064
BFI (2011): BFI-level 2
Scopus rating (2011): CiteScore 5.11
Web of Science (2011): Impact factor 3.452
BFI (2010): BFI-level 2
Web of Science (2010): Impact factor 2.258
BFI (2009): BFI-level 2
BFI (2008): BFI-level 1
Original language: English
Electronic versions:
QRbus.pdf
DOIs:
Using internet search queries to predict human mobility in social events

While our transport systems are generally designed for habitual behavior, the dynamics of large and mega cities systematically push it to its limits. Particularly, transport planning and operations in large events are well known to be a challenge. Not only do they imply stress to the system on an irregular basis, their associated mobility behavior is also difficult to predict. Previous studies have shown a strong correlation between number of public transport arrivals with the semi-structured data mined from online announcement websites. However, these models tend to be complex in form and demand substantial information retrieval, extraction and data cleaning work, and so they are difficult to generalize from city to city. In contrast, this paper focuses on enriching previously mined information about special events using automated web search queries. Since this context data comes in unstructured natural language form, we employ supervised topic model to correlate it with real measurements of transport usage. In this way, the proposed approach is more generic and a transit agency can start planning ahead as early as the event is announced on the web. The results show that using information mined from the web search not only shows high prediction accuracy of public transport demand, but also potentially provides interesting insights about popular event categories based on extracted topics.

General information
State: Published
Organisations: Department of Management Engineering, Transport Modelling, Transport DTU, KTH - Royal Institute of Technology, MIT-Singapore Program, Los Alamos National Laboratory
Authors: Borysov, S. (Ekstern), Lourenco, M. (Ekstern), Rodrigues, F. (Intern), Balatsky, A. (Ekstern), Pereira, F. C. (Intern)
Number of pages: 6
Pages: 1342-1347
Publication date: 2016

Host publication information
Title of host publication: Proceedings of 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)
Publisher: IEEE
Main Research Area: Technical/natural sciences
Data models, Predictive models, Internet, Urban areas, Web search, Planning
DOIs:
10.1109/ITSC.2016.7795731
Source: Findit
Source-ID: 2350291944
Publication: Research - peer-review › Article in proceedings – Annual report year: 2016

DynaMIT2.0: architecture design and preliminary results on real-time data fusion for traffic prediction and crisis management
The ability to monitor and predict in real-time the state of the transportation network is a valuable tool for both transportation administrators and travellers. While many solutions exist for this task, they are generally much more successful in recurrent scenarios than in non-recurrent ones. Paradoxically, it is in the latter case that such tools can make the difference. Therefore, the dynamic traffic assignment and simulation based prediction system such as DynaMIT (1) demonstrates high effectiveness in the context of sudden network disturbance or demand pattern changes. This paper presents the design, development and implementation of new components and modules of DynaMIT 2.0 which is an extension of its predecessor with recent enhancements on online calibration, context mining, scenario analyser and strategy simulation capability. Also, some preliminary results are presented using Singapore expressway to show the actual benefit of the system.

General information
State: Published
Organisations: Department of Transport, Singapore-MIT Alliance for Research and Technology, National Technical University of Athens, Massachusetts Institute of Technology
Authors: Lu, Y. (Ekstern), Pereira, F. C. (Intern), Seshadri, R. (Ekstern), O'Sullivan, A. (Ekstern), Antoniou, C. (Ekstern), Ben-Akiva, M. (Ekstern)
Number of pages: 6
Pages: 2250-2255
Host publication Information
Title of host publication: Proceedings of the 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC)
Publisher: IEEE
ISBN (Print): 978-1-4673-6595-6
Main Research Area: Technical/natural sciences
data handling, data mining, emergency management, road traffic, sensor fusion, traffic information systems, Transportation, architecture design, Calibration, Computational modeling, context mining, crisis management, Data models, demand pattern changes, dynamic traffic assignment, dynamic traffic simulation, DynaMIT2.0, Prediction algorithms, Predictive models, real-time data fusion, Real-time systems, scenario analyser, Sensors, Singapore expressway, strategy simulation capability, sudden network disturbance, traffic prediction, transportation administration, transportation network state monitoring, transportation network state prediction, travellers
DOIs:
10.1109/ITSC.2015.363
Source: FindIt
Source-ID: 276541181
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015

Intelligent Transport Systems in the Smart City
General information
State: Published
Organisations: Department of Transport
Authors: Pereira, F. C. (Intern)
Number of pages: 1
Publication date: 2015

Host publication Information
Title of host publication: Book of Abstracts. DTU's Sustain Conference 2015
Place of publication: Lyngby
Publisher: Technical University of Denmark (DTU)
Article number: L-9
Main Research Area: Technical/natural sciences
Conference: DTU Sustain Conference 2015, Lyngby, Denmark, 17/12/2015 - 17/12/2015
Electronic versions:
L9_DTU_Sustain_2015.pdf
Publication: Research - peer-review › Conference abstract in proceedings – Annual report year: 2015

Projects:

Autonomous Bus Demand Modeling from Big Data
Department of Management Engineering
Period: 01/01/2018 → 31/12/2020
Number of participants: 3
Phd Student:
Peled, Inon (Intern)
Supervisor:
Dauwels, Justin (Ekstern)
Main Supervisor:
Pereira, Francisco Camara (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD
Traffic controlled by Air Quality - Organic City

Increased urbanisation puts pressure on city infrastructures. In traffic this infers increased congestion detriment to health and to the more general liveability of the city. As a City, Copenhagen has a strong will to find new ways to management the increased urbanisation in order to provide citizens with a high quality of life. Smart city solutions hold potentials to restructure the management of the city with digital solutions rather than cumbersome and expensive physical investments. However, many of these solutions have no or little track record, and needs testing – by the civil services operating the new solutions and by the citizens living in them. Our motivation to apply to this call is a genuine interest in building new solutions to our citizens that provide them higher quality of life. The ability to try and test solutions with agility and citizens’ feedback holds a promise to shortcut and to accelerate the transition of the city to the future.

The aim of this project is therefore twofold: 1) to build capacity to swiftly conduct experiments in the settings where solutions have to be implemented, 2) to generate new knowledge about the relation of traffic management to air quality. Combined the project aims to establish proofs of value of smart city solutions in traffic to health and related public expenses and to the ability to implement new city workflows of prototyping and testing. Previous Organicity participants have already focused on the topics of air quality and traffic:

AirPublic through high-resolution, real-time and low cost spatial maps of air pollution in London through small sensors in an effort to produce data that can enable informed decision making related to the problem.

While the use of real-time data in both instances is intriguing we are of the opinion that such data is only needed for a limited time and not forever. We believe that the gathering of real-time data through a reasonable time period can provide enough information to still make changes possible if that information is integrated into the planning and management aspect of a city. Thus, we propose a project for improving the air pollution of a city through changes in traffic management through a limited time period of real-time data gathering. The project design monitor two stretches of roads of similar design and with consistent traffic flows. Sensors placed at intersections will monitor air quality (NOx, CO, CO2, PM10 and PM2.5). Traffic lights would then be optimized for better health possibilities by altering the traffic flow and registering differences.

Department of Environmental Engineering
Air, Land & Water Resources
Department of Management Engineering
Transport DTU
Transport Modelling
Copenhagen Solutions Lab
Period: 01/10/2017 → 31/03/2018
Number of participants: 2
Project participant:
Mikkelsen, Teis Nørgaard (Intern)
Pereira, Francisco Camara (Intern)

Multi-model bus Arrival Predictinwith Intelligent Handling of Uncertainties
Department of Management Engineering
Period: 01/02/2017 → 30/01/2021
Number of participants: 3
Phd Student:
Petersen, Niklas Christoffer (Intern)
Supervisor:
Heckscher, Annette (Ekstern)
Main Supervisor:
Pereira, Francisco Camara (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Ansat eksternt
Project: PhD

ATIS and modal shift: the role and the effectiveness of information provision and perception
Department of Management Engineering
Period: 15/12/2015 → 30/06/2019
Number of participants: 3
Phd Student:
Mehdizadeh Dastjerdi, Aliasghar (Intern)
Supervisor:
Kaplan, Sigal (Intern)
Main Supervisor:
Pereira, Francisco Camara (Intern)

**Financing sources**
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

**Machine Learning and Mobility**
Department of Management Engineering
Period: 15/12/2015 → 14/12/2018
Number of participants: 3
Phd Student:
Markou, Ioulia (Intern)
Supervisor:
Larsen, Allan (Intern)
Main Supervisor:
Pereira, Francisco Camara (Intern)

**Financing sources**
Source: Internal funding (public)
Name of research programme: Institut stipendie (DTU)
Project: PhD

**Intelligent Composition of Buffer Times in Railway Scheduling**
Department of Management Engineering
Period: 15/12/2014 → 04/04/2018
Number of participants: 6
Phd Student:
Cerreto, Fabrizio (Intern)
Supervisor:
Harrod, Steven (Intern)
Main Supervisor:
Nielsen, Otto Anker (Intern)
Examiner:
Pereira, Francisco Camara (Intern)
Krasemann, Johanna Törnquist (Ekstern)
Preston, John (Ekstern)

**Financing sources**
Source: Internal funding (public)
Name of research programme: Institut stipendie (DTU)
Project: PhD

**Activities:**

**Projects for Mobility in the Smart City**
Period: 1 Jun 2018 → 30 Jun 2018
Inon Peled (Organizer)
Francisco Camara Pereira (Organizer)
Department of Management Engineering
Transport DTU
Transport Modelling
Degree of recognition: Local

Related event

Projects for Mobility in the Smart City
01/06/2018 → 30/06/2018
Activity: Attending an event › Participating in or organising workshops, courses, seminars etc.