Journal of Visual Languages and Computing 25 (2014) 973–980

Contents lists available at ScienceDirect

Journal of Visual Languages and Computing

journal homepage: www.elsevier.com/locate/jvlc

Short Paper

Predicting and visualizing traffic congestion in the presence of planned special events

Simon Kwoczek [a,b,*], Sergio Di Martino [a], Wolfgang Nejdl [b]

[a] Group Research, Volkswagen AG, Germany
[b] L3S Research Center, University of Hanover, Germany

ARTICLE INFO

Article history:
Received 28 September 2014
Received in revised form
24 October 2014
Accepted 28 October 2014
This paper has been recommended for acceptance by S.-K. Chang
Available online 7 November 2014

Keywords:
Mobility
Traffic prediction
Visualization
Machine Learning
Geographic information systems

ABSTRACT

The recent availability of datasets on transportation networks with higher spatial and temporal resolution is enabling new research activities in the fields of Territorial Intelligence and Smart Cities. Among these, many research efforts are aimed at predicting traffic congestions to alleviate their negative effects on society, mainly by learning recurring mobility patterns. Within this field, in this paper we propose an integrated solution to predict and visualize non-recurring traffic congestion in urban environments caused by Planned Special Events (PSE), such as a soccer game or a concert. Predictions are done by means of two Machine Learning-based techniques. These have been proven to successfully outperform current state of the art predictions by 35% in an empirical assessment we conducted over a time frame of 7 months within the inner city of Cologne, Germany. The predicted congestions are fed into a specifically conceived visualization tool we designed to allow Decision Makers to evaluate the situation and take actions to improve mobility.

1. Introduction

In times of ongoing urbanization and steady growth of mobility demands, traffic congestions cost billions of dollars to the society every year [13]. Within the last years, thanks to the advances in sensor technologies (like Smartphones and GPS handhelds) and storage capabilities, mobility datasets with high spatial and temporal resolution have become available. This has led to advanced investigations on the impact of different influencing factors on traffic congestions, as daily concerts or rush hour, traffic lights, construction zones, etc. (i.e. [2,10,8]). Most of these factors are recurring on a regular base (i.e. rush hour), exist only once for limited time (i.e. construction zones) or their occurrences are unpredictable (i.e. accidents). Current state of the art commercial solutions are able to predict well recurring traffic situations, as this behavior can be easily learned from historical data [19]. On the other hand, also non-periodic events with an expected large attendance (known also as Planned Special Events, or PSE as introduced in [7]), such as, for instance, concerts or soccer games play a major role for delays in everyday transportation [10]. As example, the concert of Rihanna in Johannesburg (South Africa) in October 2013 caused people to sit in traffic for as long as five hours, trying to reach the stadium. Similarly, the concert of Robbie Williams in London, in 2003, caused tailbacks up to 10 miles on the highway A1 towards the stadium. An interesting aspect is that the traffic due to PSEs has a quite typical behavior, having two subsequent waves of congestion [11]. The first one is caused by people going to the event, while the second one is due to people leaving the venue, and may be even bigger that the first wave, since the most of the

* Corresponding author.
E-mail addresses: simon.kwoczek@volkswagen.de (S. Kwoczek), sergio.di.martino@volkswagen.de (S. Di Martino), nejdl@l3s.de (W. Nejdl).

http://dx.doi.org/10.1016/j.jvlc.2014.10.028
1045-926X/

attendees are leaving at the same time. Despite these numbers, to date very few research attention has been devoted to predict the congestion due to a PSE. As a consequence, even the most advanced available commercial systems are incapable of predicting this kind of nonrecurring traffic ahead of time.

To address this open issue having huge social costs, in this paper we describe a solution we developed to predict and visualize the spatio-temporal impact of the second wave of traffic due to a PSE around its venue. In particular, the proposed approach is meant to be executed while the event is happening, and takes as input the category of the PSE, like "Concert" for instance, and the information on the first wave of traffic, coming from traditional traffic providers. Then, using an adaptation of the K-Nearest Neighbors (K-NN) algorithm, we look for the most similar past PSEs (cases) among historic observations, in order to derive a prediction of the impact of the second wave of traffic. Such a prediction is done in terms of average delay over the road segments around the venue that have been found to be highly correlated with the congestions due to PSEs. To graphically visualize the results, we developed a Google Earth-based tool,[1] similar to the ones developed in [14,4] or [6]. Thanks to this integrated solution, on one hand it is possible for Decision Makers to understand how the traffic situation is influenced by a PSE, in order to take actions to improve mobility. On the other hand, the predicted impact can be directly fed to a route calculation engine, in order to reroute passing drivers away from congestions.

To assess the proposed approach, we used data on traffic and events from June to December 2013 in the inner city of Cologne, Germany. The traffic data has been provided by one of the most prominent traffic providers in Europe. It comes mainly from Floating Car Data, i.e. data collected from GPS sensors in vehicles, and it covers most of the streets within the inner city. As for the PSEs, we considered all the 29 events hosted in the Cologne LANXESS arena, in the same temporal span. By performing a leave-one-out cross validation on the event dataset, we compared our proposal with the current state of the art, intended as real time traffic information, and with a baseline consisting simply of replicating the impact of the first wave as predicted second wave. Results show that our proposal outperforms both alternative solutions, providing prediction that are up to 35% better compared to the current state of the art solutions and up to 41% better compared to the introduced alternative baseline.

The remainder of the paper is structured as follows: Section 2 presents related work within the field of traffic predictions and PSEs and some background terminology. In Section 3 we present the approach to predict the congestions caused by outbound traffic after an event, plus the GUI to visualize the results. In Section 4 we describe the Research Questions and the evaluation protocol to assess the proposed approach. In Section 5 the results of the empirical assessment are presented and discussed. Finally in Section 6 conclusions are outlined, together with some future research directions.

[1] http://earth.google.com

2. Background and definitions

Traffic congestion predictions have been widely studied within the research communities of Intelligent Transportation Systems (ITS), Smart Cities and Territorial Intelligence, leading to a rich body of literature. Indeed, by knowing in advance the traffic patterns, route calculation engines can compute more efficient routes, saving also time for the drivers. Due to the lack of real mobility data, early approaches for traffic predictions were mainly based on simulations and theoretical modeling (e.g. [2,3]). Nowadays thanks to the spreading of GPS and high-speed cellular networks, new massive datasets on traffic are available. Consequently, several statistic and data driven approaches for traffic predictions have been presented to the community. Examples include generalized linear regression [22], nonlinear time series [9], Kalman filters [12], support vector regression [21], and various neural network models [18,12,20]. A combination of these approaches is also used by current commercial navigation solutions, able to predict recurring congestions by learning characteristic traffic flow patterns for street segments from historical data. On top of that, these commercial systems can also optimize the route planning based on the real-time traffic situation [19].

In general, traffic congestion can be divided into recurring congestions, usually caused by a mobility demand that exceeds the capacity of the road network (e.g. due to rush hour), and congestions that are nonrecurring (e.g. due to incidents or special events) [10]. The effects of nonrecurring traffic congestions and their prediction are a widely investigated topic within the research community (e.g.[16,17,15]). Although these approaches showed significant improvements in prediction, their focus lies on highways that are one-directional road segments, whereby usually in the inner cities the impact of traffic is a multidimensional problem, evolving in a 2D, more complex route network.

Previous researches have highlighted that also PSEs are a possible influencing factors [10,8], since they may lead thousands (or even hundreds of thousands) of people to travel towards the same destination in a very limited time interval, and then to leave the venue again in a very short time span. To the best of our knowledge, the only work available that has its focus on the influence of PSEs on traffic is presented in [11]. The authors report a generic overview about the influence of PSEs on the road network, derived from an event classification defined by the Chinese State Council. They also introduce management plans for the different types of events, but there is no quantifiable solution for the prediction of the traffic.

2.1. Definitions

A PSE may have an impact on the traffic behavior in a specific region over time. Such an *impact region* can be defined as the list of congested segments of the road network. A *congested segment* is a piece of road network