1. Problem Description

In highly frequent bus routes it is more important the headway regularity than the fulfilment of the arrival time at the bus stops. In fact, a small delay of a bus provokes the raising of the number of passengers in the next stop. This number increases the dwell time and obviously, it also increases the bus' delay. On the other hand, the next bus will have fewer passengers, shorter dwell times without delays. This will continue as a snow ball effect and, at a further point of that route, the two buses will meet at a bus stop, forming a platoon. This phenomenon is denominated as Bus Bunching (BB).

2. Goals

- To explore both historical and real-time Automatic Vehicle Location data by combining Offline and Online Supervised Learning techniques;
- To predict the certain short-term occurrence of such BB events in real-time;

3. Case Study

- City of Porto, Portugal with 1.3 million inhabitants;
- The biggest public transport operator in town (STCP) with 51 running lines;
- One year of data (2010) describing the trips of 3 lines (6 routes);
- Each data record contains a 1) route, 2) a time stamp, 3) the daytype and 4) the weekday, the respective 5) bus stop and 6) the Link Travel Time to arrive to it;

4. Methodology Illustration

Stepwise Framework:

1. To Perform Link Travel Time (LTT) prediction for all trips on a daily basis by employing the Random Forests for regression over historical trip-based data about the last 7 days;
2. These predictions are incrementally refined throughout the day using the averaged LTT residuals of the last trip to update all the LTT values for the next one (i.e. Greedy Exp. Delta Rule);
3. These predictions are used to infer the future headways - then, they are also incrementally refined using the stop-based residuals to slightly update the predictions (i.e. Exp. Delta Rule);
4. The headway predictions are used together with the most recent headway-based residuals to infer the headway's p.d.f. on each stop by assuming that they follow a Gaussian curve;
5. A BB likelihood is calculated for each stop based on such p.d.f. and combined to form a global BB score for the downstream stops;
6. A BB alarm may be triggered if this global BB likelihood goes over a certain frequency-based threshold;

1 Based on the Perceptron’s Delta Rule to update the weights on each neuron, these rule is used directly over the outputs with a dynamic α (i.e. the learning rate) - which reflects the current need of react the drifts on the underlying concepts (e.g. traffic jams provoking large delays on a given link).

5. Framework Description

6. Experimental Results

6. Discussion and Future work

By analysing the results, it is possible to conclude that the First-Order update rules are generally improving the convergence of our predictions to its true values. On other hand, the low Precision values shows that our model triggers more alarms than necessary - which is still an acceptable behavior in the preventive context that we are facing. This work can be extended on three distinct axis:

1. on the dataset, by including a larger dataset containing a set of lines more representative of the entire network;
2. on the parameter setting, by conducting a large-scale sensitivity analysis on their values;
3. on the corrective actions, by proposing a method to choose where and when a action should be took to avoid BB, as well as one to choose which is the best one to take in each case.

Funding

This work was supported by the VTL: “Virtual Traffic Lights” (PTDC/EIA-CCO/118114/2010), by MAESTRA (ICT-2013-612944), by I-CITY - "I-City for Mobility" (NORTE-07-0124-FEDER-009864) and by ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness) and also by the Portuguese Funds through the FCT (Portuguese Foundation for Science and Technology) within project FCOMP-01-0124-FEDER-037281.