AllAboard: a System for Exploring Urban Mobility and Optimizing Public Transport Using Cellphone Data

Michele Berlingerio, Francesco Calabrese, Giusy Di Lorenzo, Rahul Nair, Fabio Pinelli, and Marco Luca Sbodio

IBM Research, Dublin, Ireland
{mberling,fcalabre,giusydil,rahul.nair,fabiopin,marco.sbodio}
[at]ie.ibm.com

1   Introduction

The deep penetration of mobile phones offers cities the ability to opportunistically monitor citizens' interactions and use data-driven insights to better plan and manage services. In this context, transit operators can leverage pervasive mobile sensing to better match observed demand for travel with their service offerings. With large scale data on mobility patterns, operators can move away from the costly and resource intensive transportation planning processes prevalent in the West, to a more data-centric view, that places the instrumented user at the center of development. In this framework, using mobile phone data to perform transit analysis and optimization represents a new frontier with significant societal impact, especially in developing countries.

In this demo, we present AllAboard, a system for optimizing public transport using cellphone data. Our system uses mobile phone location data to infer origin-destination flows in the city, which is then converted to ridership on the existing transit network. Sequential travel patterns extracted from individual call location data are used to propose new candidate transit routes. An optimization model evaluates which new routes would best improve the existing transit network to increase ridership and user satisfaction, both in terms of reduced travel and wait time. The system provides also a User Interface that allows the interaction with results and the data themselves. The system in its whole is intended to be used by city authorities for improving their public transport systems, using cell phone data, which have a large penetration even in developing countries, and provide a cheaper, faster, alternative to costly surveys.

The system has been tested using Call Detail Record data from Orange for the city of Abidjan, Ivory Coast, with the focus to improve the existing transit network. Four new routes have been proposed by the optimization system, resulting in an expected reduction of 10% city-wide travel times.

Several projects deal with the analysis of mobility and mobile phone data [2, 1]. They present powerful mining engines but do not provide direct interaction with data and results. Other projects providing visualization and interaction, on the other hand, do not integrate the optimisation based on analytical results

[3, 4]. Our system integrates all these modules: the mobility analysis engine, the optimisation, and an interactive user interface, providing a new environment to extract and use information from mobile phone data.
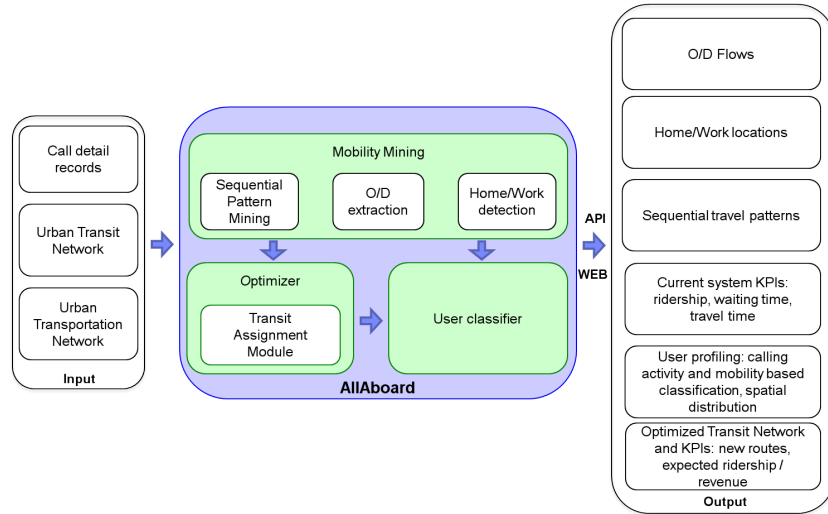
2 The AllAboard Platform

Fig. 1. Architecture of the AllAboard platform

The system is implemented with a modular architecture. We isolate our data models using an abstraction layer that separates algorithm implementations from data stores. The current version of the system includes two core modules: *Mobility Mining* and *Optimizer*. These modules are implemented as components within an extensible framework, and other components can easily be added in the future. Each component provides a lightweight REST service exposing its functionalities. The REST services are also used to implement the AJAX-based Web user interface. The main algorithms for the current two cores are described in Sections 2.1, and 2.2 respectively.

2.1 Mobility Mining

This module is able to process mobile phone data in the form of records: (userID, timestamp, cellID), where a location is associated to each cellID, and extract information about users' stops, trajectories, Origin/Destination (O/D) flows (i.e. number of people moving from the origin O to the destination D in a given 1 hour time interval), frequent travel sequences, and home and work locations for each user.

The methodology to process the data follows these steps:

1. we extract the location of the stops performed by users
2. based on those, we estimate the O/D flows, used to feed the optimization module presented in Section 2.2
3. we exploit the data to better understand the mobility of the users, and use the results as additional input to the optimization module. For this, we extract frequent sequential patterns from the sequences of stops
4. we identify for each user, when possible, most likely home and work locations based on both stops and mobile phone activity patterns

The results of this process are visualised by the framework, allowing the user to interact with them. For instance, the user can select a particular antenna and then the system visualises its O/D flows and the relative temporal profile. The system is also able to show the sequential patterns emerging over time from the mobile phone data, highlighting the most frequent ones. An example of the two visualisations is represented in Figure 2.
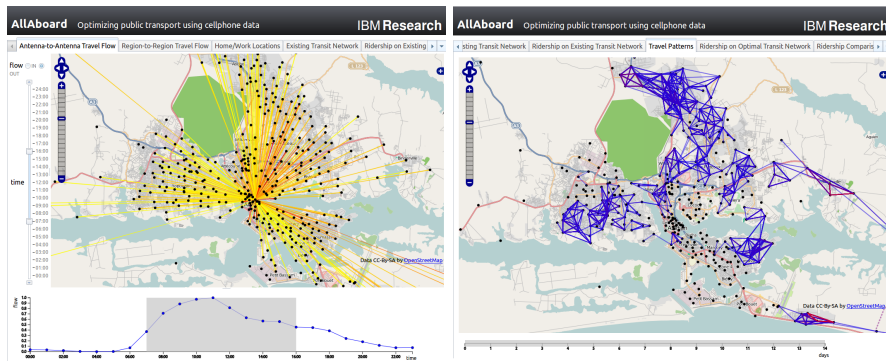

Fig. 2. Exploring mobility with AllAboard: O/D flows with temporal profiles (left); sequential patterns emerging from the data (right)

2.2 Optimal Transit Design

Given (a) an existing transit network, (b) O/D flows derived from mobile data representing travel demand, (c) a set of frequent sequences that serve as candidate new routes, (d) travel time estimates across the network, and (e) a resource budget in terms of fleet size, this module is able to determine an optimal set of new routes and their associated service frequencies, such that passenger journey times city-wide are minimized. A new route is defined by a sequence of transit stops and has an associated frequency.

The problem is strategic in nature as it represents a longer-term decision on the part of a public transport operator. The addition of new routes to the service
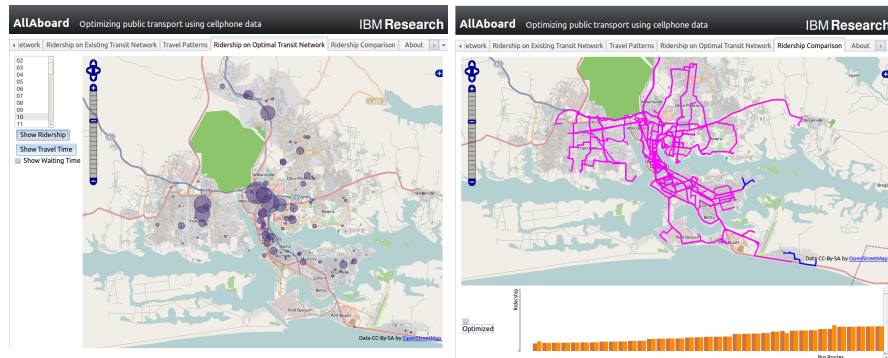
Fig. 3. Exploring optimization results with AllAboard: waiting times (left); comparing ridership on optimized network (right)

network are intended to match current supply with revealed demand. From a demand perspective, new routes will attract riders only if they offer competitive service to existing routes. The model therefore considers the user perspective, in terms of optimal strategies. From a supply side, new routes should fill service gaps and map user activity patterns observed in the data. The optimization routine therefore includes frequent sequences as potential new services. Taken together, a potential new route is only recommended if it directly addresses under-serviced demand, and does so by offering shorter journey times than the ones possible on the existing network.

The system offers an interactive User Interface that allows to visualise the results of the optimisation module, in particular in terms of ridership improvements, and decrease in waiting times, comparing the results between the existing network and the new extended one (see Figure 3).

References

1. F. Calabrese and C. Ratti. Real time rome. *Networks and Communications Studies*, 20(3-4):247–258, 2006.
2. N. Eagle and A. (Sandy) Pentland. Reality mining: sensing complex social systems. *Personal Ubiquitous Comput.*, 10(4):255–268, 2006.
3. F. Giannotti, M. Nanni, D. Pedreschi, F. Pinelli, C. Renso, S. Rinzivillo, and R. Trasarti. Unveiling the complexity of human mobility by querying and mining massive trajectory data. *The VLDB Journal*, 20(5):695–719, Oct. 2011.
4. Y. Zheng and X. Zhou, editors. *Computing with Spatial Trajectories*. Springer, 2011.