The SmartDeliveries project in Lyon: optimizing professional urban mobility leveraging ITS

Baudel, Thomas*1; Aguiar-Melgarejo1,2, Penelope; Laborie, Philippe1; Aston, Jean1,3; Depetris, Alain1
1. baudelth@fr.ibm.com, IBM France Lab, France
2. INSA Lyon, France
3. Ecole Polytechnique, France

Abstract
SmartDeliveries is a city-wide delivery rounds optimization system aiming at leveraging city traffic information to optimise professional vehicle rounds. This system has been developed as part of the OptimodLyon project. One of the key features of the system is to integrate information of short and longer term expected mobility demand. In this paper, we provide a detailed presentation of SmartDeliveries system and some evaluation results that show its potential to both optimize urban deliveries (18% savings in distance and 11% in time) and, if adopted at the scale of a city, significantly contribute to globally improved traffic (5% reduction in traffic and corresponding reduction in emissions).

Keywords:
Urban logistics, round optimization, traffic monitoring.

Introduction
Context
Traffic management in major cities faces well-known but major challenges. While mobility is directly tied to economic growth and prosperity, road networks are saturated. Current approaches in Intelligent Transportation Systems (ITS) favor, among others, to rely on better information processing from all the actors involved to develop alternate mobility patterns, for instance encouraging the use of public transports and cycling, or deferring rush hour travels.

OptimodLyon (http://optimodlyon.fr/en/) is a three year project supported by the French agency for the environment, in which eight industry partners, among which IBM, four academic institutions and the metropolitan authority of Lyon (called Urban Community of Lyon) work together to improve urban mobility through better collection, processing and distribution of mobility information. The flagship application of OptimodLyon is a real-time, multi-modal journey planner targeted at personal uses.
Planned, professional urban mobility

Despite environmental concerns, there is a large portion of road traffic that will not lend itself easily to modal transfer. According to the (Academie des Technologies, 2009) report, 35% of road trips carry goods rather than people. If we account for the important portion of movements that involve moving both goods and persons to deliver a value added service (such as maintenance and construction visits, catering, sales visits…), an even larger portion of existing road traffic needs to be accounted for. We call this portion of the traffic we focus on “professional urban mobility” and we estimate (counting campaigns performed internally) that it covers at least 50% of daytime urban traffic. Most policies aimed at limiting private vehicle use affect almost equally professional urban mobility, which could impact economic activity negatively. Any policy aimed at limiting the negative externalities of urban car travel must therefore target this portion of the traffic properly.

Businesses are also acutely aware and affected by the cost of mobility and have direct incentives to rationalize it. For instance, fuel represents 18% of the total costs of a small freight operator employing 200 persons (undisclosed personal communication). Current estimates (CERB, 2007) consider that 20 minutes lost in urban traffic amount to about eight euros for such an operator, which represents, by-and-large, its gross financial margin for a half-day vehicle round. We can therefore consider there is a synergy of goals between public authorities and the private businesses in curbing congestion and optimizing trips. The shared objective of both public and private stakeholders is to tend towards the minimization of vehicle-kilometers and time spent on the road.

Our work stems from the observation that a large portion of professional mobility demand is actually planned. Somewhere, in the information system of the companies, the intent to move is known in advance, at a close informational distance from an internet connection. If the city can securely obtain this information and use it to improve its traffic forecast, providing in return relevant suggestions that can reduce the cost of mobility, both parties could benefit from this exchange of information.

Synergy between cities and freight operators

The expected gains of a cooperative data exchange for cities and road users are as follows:

- Freight trip planning data constitutes a new source of information for traffic regulation, providing an advance view of freight mobility, which represents a sizable portion of traffic.
- Data exchange provides a new regulation tool, enabling to divert a sizable portion of traffic outside of congested areas.
- Because the service may measurably save time and money to its users, it can be turned into a revenue source for the city to leverage its traffic regulation investments.

For its users, we expect a high profitability, as distance and time savings directly impact their revenue. Finally, the costs of exploiting such a solution is fairly small, as all it requires is an information infrastructure which is already in place: internet, mobile phones, and traffic data. In the remainder of this paper, we present an implemented prototype, then discuss the evaluations we have performed with three companies and 4000 tours. These tests led us to identify three major technical challenges, as well as helped us quantify the potential benefits of the system. In the final section of this paper, we discuss the remaining challenges, which are divided between technical and organizational considerations.
Overview of the proposed system

SmartDeliveries (Baudel et al, 2013) is a software service and a set of applications comprising the following elements (Figure 1):

A situation monitor maintains a model of the current and forecast traffic conditions for the whole city. It connects to the city information services (such as data.grandlyon.com) to gather planned events and public works information, real time and forecast traffic conditions, weather data. It provides means to plan routes and current and future travel time estimates.

A demand monitor gathers users’ mobility plans. It is accessed either via a web application (for occasional users) or a dedicated web service (connecting to the firm’s information services). The plans comprise a list of destinations, together with constraints on the trips, such as time and precedence constraints.

A routes monitor creates, optimizes and follows all planned trips in real time, adapting to the observed route taken (as the driver stays in control of the plans), and provides optimization suggestions and alerts when there are traffic events from the city or deviations from the initial plans. It connects to a mobile application that acts as a route planner and provides traffic alerts.

Once fully implemented, this architecture shall enable the computation of global optimization plans, enabling to spread traffic in time and space when potential delivery parking spots congestion is detected. The key implemented components for now are the travel time estimator and the time-dependent, robust optimization service, which will be described hereafter.

Usage scenarios

The system is meant to be used first in the morning, once the rounds have been assigned to individual drivers. The fleet manager can either upload route sheets on a web interface in a straightforward csv file format, or, for larger companies, the IT system can be directly interfaced via web services to the route
The SmartDeliveries project in Lyon: Optimizing Professional Urban Mobility Leveraging ITS

optimizer. The system in return provides route recommendations, including public works and events that may impact the trajectory. For ease of integration, those can either be printed and given to the driver for guidance, or they may be uploaded on a smartphone application, which the driver may use to perform amendments to the suggested route, often adding constraints that are known only by them, such as the opening and closing hours of the delivery locations. Route changes and constraints on the delivery locations are stored for future reuse, which provides a degree of knowledge sharing and simplifies the daily use of the application for repetitive rounds.

When the planned tour has been downloaded on a GPS-enabled driver’s mobile phone, the application with follow the driver, and manage to guess the course of the round, issuing planning corrections if the driver decides not to follow the initial recommendations. There can be many good reasons for a driver to modify the proposed tour, as we will see later, so the system is designed as a “recommender” rather than a prescriptive system. We believe this is a key issue to ensure that users accept to be tracked by a GPS enabled device. As an additional benefit, traffic events (such as accidents and congestions) are relayed to the mobile phone and trigger alerts and route change proposals, possibly re-ordering the stops so that problematic locations can be avoided. To the drivers, and the city, this later feature is seen as quite novel over common navigation aids: the ability of the planner to foresee several steps ahead allows issuing more complex re-routing than one-step ahead suggestions. Of course, it is possible to limit the amount of re-ordering, to account for heavy cargo that cannot be moved easily.

Although fairly new, some commercial systems do provide somewhat similar functionalities, but the integration of the services, the interfacing with the city’s ITS, and the deliberate design as a recommender rather than a prescriptive system represent, to our knowledge, innovations.

Resources and methods

We performed extensive studies and projections to assess the suitability of the system to its purpose. We could identify a number of technical hurdles to overcome. Once they are, the gains seem fairly substantial. The city provides six minutes real time and historicised traffic data over 600 detectors spread throughout the urban area, advance data of the public works impacting road usage, as well as a DATEX II event stream manually fed to alert of ongoing traffic events such as accidents and major congestions (http://data.grandlyon.com/). Several studies (Godinot & Bonnel, 2008, Bousquet, 2010, Leclercq & Coldefy, 2013) have shown that historic traffic data together with known planned events allow for providing a somewhat accurate picture of traffic, and thus create a travel time model usable for route planning. Our goal is to see if this knowledge can be used to optimize vehicle rounds at the scale of the city.

We first conducted a survey, interviewing about 50 drivers, fleet managers and logistics managers in France, from large to small freight carriers, in order to collect needs and assess workflows. When we started the studies, in 2012, all freight carriers had plans to leverage mobile technologies to improve agility and service to customers, but no system was fully operative or functioning as expected. Drivers mostly relied on their field experience to organize their rounds, using simplistic but somewhat effective nearest-neighbor heuristics to optimize their route. Existing routing system seemed to lack precision in travel time estimation to provide usable guidance, at least in the companies we surveyed. In 2015, all major companies are
The SmartDeliveries project in Lyon: Optimizing Professional Urban Mobility Leveraging ITS

equipped with GPS vehicle tracking, and they are somewhat satisfied with the ability to better inform their customers of expected delivery time, and to better measure the productivity of their drivers. Yet, route optimization still appears of little use in urban round planning.

In order to assess why, and to see if accurate traffic information received from the city would remediate this reluctance to optimization, we have obtained two full months of complete delivery rounds, time-stamped and geo-localized, from three freight carriers in Lyon. This represents over 80,000 individual moves, 4,000 rounds, averaging 20 positions per round, and 10 minutes on average per trip. After various stages of data cleansing (described below), we obtained a test set of 1715 rounds that were suitable for evaluating the potential gain of a city-scale round optimization system (Figure 2).

Using traffic and event data from the city for the corresponding months, we simulated our system to assess whether we could have provided the drivers with better routes. The key findings started with the identification of three major technical challenges:

- Understanding the mobility intent from the set of addresses that had been provided.
- Estimating travel times correctly.
- Optimizing algorithms that leverage accurately the available city data.

We will examine those issues in detail before assessing the potential value of the system.

**Understanding the mobility intent**

While seemingly trivial, figuring where a delivery is to be made given a written address turns out to be a challenge. 40% of the addresses that were provided to us in the pilot did not conform to a standard format.
Geocoding is a difficult problem, which is highly dependent on local conventions, and which involves careful trade-offs to avoid overfitting while still providing enough error recovery (Goldberg et al. 2007). After fuzzy-match geocoding, 16% of the addresses are still imprecisely located, and 1% wrongly located. Note that our results compare favourably to off-the-shelf solutions such as Google maps, as we specialized our algorithm for the data at hand. Examples of imprecisions include: providing only a cell phone number as address, or providing a phrase such as “Street XX, construction site in front of the bakery.” Quite often, the point of delivery is different from the postal address provided by the sender; for instance, the address is that of a shop, but deliveries are to be made in a back alley or at another location. All in all, after manually resolving the uncertainties, we estimate that about 8% of the addresses cannot be precisely resolved before the round, but will instead require some on-site searching by the driver. Depending on the type of shipment (express or standard deliveries, business or individual shipping), about 1-2% of the deliveries will return at the warehouse for lack of finding the proper delivery location.

Freight carriers are well aware of this issue. Right now, they partially alleviate it by performing systematic geotagging of all delivered addresses, so that the same loss of time shall not happen twice. In order to provide robust route planning under these circumstances, our system must take into account the uncertainty of delivery locations, adding extra padding time in proportion to the uncertainty of the interpreted location. Addresses that are completely unrecognizable are discarded from the round optimization, and extra time is added to the total time estimate of the round to account for the time to search for the delivery address. The driver is reminded of this uncertainty in the mobile application, and asked to notify when the proper location has been found to allow further occurrences of this address to be taken into account properly.

This particular technological issue highlights the importance of the driver expertise: experienced drivers have usually much lower return rates and higher productivity. This guides our choice to function only as a recommender system and be flexible about the deviations of the driver from the established path. Still, we believe proper interpretation of the rounds is a major issue to address to make the Smart Deliveries system bring its full potential.

Estimating travel time

The travel time estimation model used initially (Leclercq & Coldefy, 2013) relies on a set of rules which derive average speed for all road segments on the network based on a small number of variables: period of the year, time of day, measured or predicted congestion and various other road characteristics. While this model fits the available data overall (6% difference between observed and predicted total travel time), it is hardly usable due to the huge observed variance for individual trips: the average error for individual trips of 10 minutes reaches 55%. Worse, the use of a rule-based model introduces internal biases, which distort the error distribution outside of a normal law. As a consequence, errors along a full round do not compensate each other as well as the law of large numbers should allow. The result is that the overall standard deviation for the rounds is much higher than the estimates drivers provide when polled informally (23% vs. 15-20%).

While a standard deviation amounting to 50% for trip durations of 10 minutes is of little incidence on the perception of the system reliability, it is more critical that the estimate for the overall tour match the expectation of the users. If we were to provide the system as-is to drivers, the discrepancy between forecast
and observed travel time for the full round would generate distrust in the system reliability and reduce acceptance.

Improving on our travel time estimation required a novel approach, detailed in (Ashton and Baudel, 2014), which we summarize here. We first ran supervised learning regression methods to build a decision-tree-based predictor of travel time based on route features. The predictors took as inputs various attributes of the tour, including the geocoordinates of the origin and destination, the time of day, day of week, distance between origin and destination (along the shortest path), the eccentricity of the trip (angle of the travel direction taken from the city center), the observed congestion along the travelled path, the round identifier (drivers tend to serve the same neighbourhoods every day, and are specialized for one type of delivery service, with a specific type of truck. All those are summarized by the round ID). Using a 5-fold method for cross validation, we found slightly better results for travel time estimation: 41% for individual routes, and 21% for the full rounds. The interesting lesson however is that the predictor assigns a strong predictive weight to the round id, which encapsulates many endogenous aspects of the round, while other exogeneous factors, such as the observed congestion (Occupancy attribute) have far less influence on travel time (Table 1).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Distance</th>
<th>Round ID</th>
<th>Static travel time forecast</th>
<th>Occupancy</th>
<th>Angle</th>
<th>Time of day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gini importance</td>
<td>77</td>
<td>11</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 1. Gini importance of the main attributes of the decision tree predictor.

This observation yields to the following interpretations:

- The relatively low impact of occupancy rate suggests that congestion, while definitely noticeable, is fairly quickly remediated by the traffic regulation center, and has – overall – only a low impact on total travel-time. This has been confirmed by the regulation authorities, in the following terms: “Our goal is not to get rid of all congestion, this is impossible. Rather, it is to make sure that on any trip, congestion won’t impact you by more than five minutes”. This does not mean in any way that our model can get rid altogether of city traffic information. On the contrary, awareness of the accidental or forecast (closed roads) congestions is of high value to drivers, even if it is used only 20 times per year. Provided this information is accompanied with an alternate route, it offers the opportunity to save some of what some drivers call “jours de galère” (bummer days), when a slight delay at the beginning of the round triggers cascading effects that impact the full round, resulting in a net loss for the day’s work.

- The high impact of the “round ID” attribute means that different drivers/round types have differing average speeds. In other words, endogenous aspects of the round may have a much higher impact than exogenous factors such as global road conditions. Interviews with drivers have confirmed that experienced drivers are far more productive than beginners, and of course, drivers may adopt less civil behavior when they are under stress, or take a time off if
they see that their progress has been good. This suggests a fairly important degree of elasticity in the travel time estimate, and that our estimates shall be improved by a travel time prediction model that includes historical knowledge of the driver and vehicle past driving patterns in its input variables.

These observations have led us to propose a new travel time prediction model (Ashton & Baudel, 2014), based on estimating a trip duration with a weighted average of k-nearest neighbours taken from historical data. When the distance between the trip to perform and its nearest neighbours is too large, then the model relies on the simulation-based model defined by the traffic authority. The model parameters have been determined by cross-validation on our dataset. This new model enables reaching the precision wished, with an average error of 43% on individual routes, and 13% on rounds, with 83% of the rounds being inside a 20% margin error. This roughly matches the estimates drivers provide when interviewed.

In summary, while progress is still needed, it appears of utmost importance that a travel time estimation model take into account endogenous aspects of the round to be performed to perform adequately, which is not the case of most off-the-shelf travel time estimators.

**Optimizing vehicle rounds**

The core problem at the heart of round optimization, the Traveling Salesman Problem (TSP) under constraints, is a well-known “hard” computational problem. In addition, our optimization model involves precedence and time window constraints, time-dependency (the expected duration of a trip varies with the time of day), and uncertainty (the travel time is known only within certain bounds). The approach we have chosen to solve the TSP relies on constraints-programming and scheduling (Baptiste et al. 2006), (Laborie & Rogerie, 2008), (Laborie et al., 2009). Our first implementation (Aguiar-Melgarejo et al, 2012), using state-of-the art but straightforward approaches, would require on average 90 seconds of computation to provide a solution for a round. While this is sufficient for an individual round and for evaluating the potential impact of the system, it would prove highly impractical if used in production. If we assume thousands of rounds to optimize and that each round requires about three recomputations in the course of the day, this means hundreds of hours of intensive processor time per day. Because most rounds are planned in the morning hours between 6 and 8 am, and the response time should not exceed a few minutes per request, implementing a city-scale solution would require a significant datacenter (hundreds of processors) to deliver results in a reasonable timeframe. Given the expected returns, even if the results were allowing substantial savings to its users, the solution would not be economically feasible.

Fundamentally, the model for optimizing an urban vehicle round under constraints is the Time-Dependent Traveling Salesman Problem (TDTSP). The solution we provide involves implementing a global “TimeDependant no-overlap” constraint, that takes as parameter a distance tensor (origin-destination temporal matrix, wherein each of the element is a function of time indicating the distance from an origin to a destination over time). This new constraint and its benchmark are described in (Aguiar et al, 2015). It allows leveraging the distance tensor to further propagations to drastically cut the search space.

This new constraint enables our algorithm to find the optimal solutions in 0.1s of computation for rounds of 20 positions, and reach a minimum at 5% of the optimal solution in five seconds on average (Figure 3).
This 13-fold performance increase makes our approach scalable, requiring only a dozen or so processors to serve the potential needs of a city the size of Lyon. Further work is required to adapt the model to the uncertainty of transition times.

Optimization results (projection)
Results for the freight companies

After filtering for poorly resolved delivery locations and rounds whose travel time forecast did not match the recorded data correctly, we obtained the test set described above consisting of 1715 rounds. Applying our optimization algorithms, we found the possibility of saving 18% in distance on average, and 11% in time. The following example (Figure 4 and Table 2) is a round performed on June 12, 2014. It started at 8:50, consisting of 38 positions (some positions being at the same address). The difference between forecast and observed duration is 7%, and the truck has been on its delivery locations for 53 minutes.
### Table 2. Results from Figure 4 example.

<table>
<thead>
<tr>
<th></th>
<th>actual</th>
<th>optimised</th>
<th>gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance travelled</td>
<td>63km</td>
<td>47km</td>
<td>25%</td>
</tr>
<tr>
<td>Time in movement</td>
<td>3h25</td>
<td>2h58</td>
<td>20%</td>
</tr>
<tr>
<td>End of tour</td>
<td>12h19</td>
<td>11h48</td>
<td>29 min</td>
</tr>
</tbody>
</table>

**Results at the scale of the city**

As our initial goal is to study the feasibility and potential impact of the solution for city traffic regulation, we have further projected what those individual gains could provide if the system was adopted at the scale of the city.

In partnership with the Grand Lyon climate protection service, we have obtained data after running the Freturb model (Routhier et al., 2001). According to Freturb, there are on average 7630 rounds per day over the city, among which 2400 (31%) serve more than 10 positions and could be optimized. In a rough approximation, if our number of 50% of daytime traffic qualifies as “professional urban mobility,” and 31% may benefit from our optimization service, this means SmartDeliveries offers the perspective of a reduction of 2.8% of traffic, with a net positive economic gain, as the productivity of the drivers is increased.

**Challenges and future work**

SmartDeliveries is now in preparation for conducting live tours with a few selected participants. Two main types of issues need to be carefully addressed: technological and organizational.

**Technological challenges**

Between 2012 and 2015, conditions have drastically changed, as many freight companies and drivers have acquired off-the-shelf hardware and information systems that they can leverage to use the proposed solution, at a very small cost. However, the key technological challenge remains the ability to gracefully manage the uncertainties in delivery locations and travel-time estimations. This is to be addressed partly by optimization algorithms that take these uncertainties into account, but also via user interfaces and proposed workflows that acknowledge these uncertainties, adapt to field constraints that the driver may not be able to inform the system about, and leave the user in charge. Finally, once sufficient volume can be reached, global optimization integrating delivery parking spots availability should be a significant challenge.

**Organizational challenges**

Urban freight distribution is made by a very large variety of freight operators (Dablanc, 2007). Just in the parcel and express transport sector, for example, some very big transport companies (Chronopost, UPS, DHL for France) work alongside very small ones, with few employees, or no staff at all other than the drivers themselves (own account drivers). These small companies, which can be more than several thousands operating in a single large city such as Lyon or Paris, often work for the large ones, carrying the goods and providing the final deliveries on their behalf, as contractors. This specific end of the freight market suffers from very low profit margins and difficult working conditions. These small players may find it difficult to consider investing in a new technology, preferring to deal with delivery tour incidents with their own human resourcefulness. They will need to see actual gains of the system before joining it. Also,
The SmartDeliveries project in Lyon: Optimizing Professional Urban Mobility Leveraging ITS

when embedded in highly dependent relationships with large transport companies (Reme Harnay et al., 2014), small companies will not always be able to decide for themselves if they can join the system or not. Large operators, although more technologically savvy than small ones, can be reluctant to export their data on systems they do not fully control and they may raise security issues.

However, the framework in which urban freight deliveries are currently operated may be changing rapidly, specifically in French cities, allowing for more favourable conditions for the adoption of a delivery optimising system. Access regulations in city centres such as low emission zones are currently being adopted (Dablanc & Montenon, 2015). Municipalities increase bicycle use and promote roadway usages competing with delivery activities. Also, the rise in “residential deliveries” (deliveries for e-commerce consumers, at home or in pick-up points spread around residential and commercial neighborhoods) in cities is adding a significant layer of complexity to delivery round planning. All these new urban conditions may make it more difficult to operate “low cost” last mile deliveries and may generate interest in optimization systems integrating real time data from all stakeholders.

**Conclusion**

There are good reasons to believe that no human can find near optimal solutions to the time-dependent traveling salesman (TDTSP) problem for more than a few positions: the problem is proven to be hard to solve. The heuristics commonly employed by drivers, such as always moving to the nearest non-visited position, are quite robust and rarely yield very bad rounds. But those solutions are still far from the optimum. It is very reasonable to expect that 10 to 20% of productivity increase can be found if state-of-the-art algorithmics can applied successfully to the urban delivery round problem. Furthermore, this productivity increase would come with improved traffic conditions, benefiting all city residents.

Freight carriers and vehicle fleet operators are well aware of this opportunity, and numerous attempts have been made in the past 30 years to materialize this potential. In North America, several companies such as routific.com are starting to propose services that come close to our proposition. Yet, in practice, at least in Europe, drivers still prefer to rely on their experience, as the provided “optimized” plans often prove to be not so optimal in practice and fail in providing the needed flexibility.

Our hope is that injecting city traffic information to the optimization process, correctly assessing the travel time by taking the driver’s experience into account and designing the full workflow of the optimization as a recommender system rather than a prescriptive system, we could see those expectations finally come to fruition.

**Acknowledgements**

The OptimodLyon project, of which SmartDeliveries is a part of, is partly funded by ADEME and the French Investissements d’Avenir programme. We thank the OptimodLyon project’s partners; Jean Coldefy, Diana Diziain and Luce Ponsar from the Grand Lyon; the freight companies who provided data and answered surveys; Christine Solnon from INSA Lyon for advising optimization research, Philippe Sahjau, Yves Daumas, Jean-François Puget and Christian de Sainte-Marie for project support, the France Lab CIS development team, and the interns for their contributions: Pierre André, Clément Mathieu and Lexi Weng.
The SmartDeliveries project in Lyon: Optimizing Professional Urban Mobility Leveraging ITS

References
Dablanc, L. and Montenon, A., 2015. Impacts of environmental access restrictions on freight delivery activities, the example of Low Emission Zones in Europe. Presentation at the Transportation Research Board 94rd Annual Meeting, Washington DC, USA.