The value of real-time traffic information in urban freight distribution

Marta Flamini¹, Marialisa Nigro², Dario Pacciarelli³,

RT-DIA-180-2010 October 2010

(1) Data Management S.p.A., 1, Largo Lido Duranti, 00128, Rome, Italy

(2) Dipartimento di Scienze dell’Ingegneria Civile, Università degli Studi Roma Tre, Via Vito Volterra 62, 00146, Rome, Italy

(3) Dipartimento di Informatica e Automazione, Università degli Studi Roma Tre, Via della Vasca Navale 79, 00146, Rome, Italy

This work is partially supported by the Italian Ministry of Research, Grant number RBIP06BZW8, project FIRB “Advanced tracking system in intermodal freight transportation”.
ABSTRACT

Routes optimization in urban freight distribution is usually an off-line process based on the knowledge of historical conditions on the network. Real-time data provided by Intelligent Transportation Systems (ITS) enable on-line re-optimization on the basis of actual traffic conditions. This paper evaluates the benefit achievable when route optimization is based on off-line data characterized by different degree of reliability. Also, the added value generated by re-optimizing the solution with real-time information is estimated. The study is carried out for a practical application to the freight distribution of perishable goods in the city of Rome (Italy). The off-line problem is formulated as a vehicle routing problem with soft time windows while in the on-line problem it is also allowed to skip some customers. Both versions are solved with different algorithms and with different data sets. The practical performance of each solution is then assessed by using a traffic simulator reproducing the typical traffic conditions of Rome. From the results obtained the marginal value of different types of information is then computed. Such value can be used to evaluate the potential return on investment on the acquisition of reliable data. At the same time, the results of this paper can be of interest also for information providers, to fix the price of off-line and on-line information and/or to estimate the associated potential market share.
1 Introduction

In the last 30 years private mobility has increased by 140% in the UE and by 214% in Italy. The European freight road transport is expected to increase by close to 55% by 2020 [4]. This increase is expected to cause a strong rise of congestion, especially in urban areas, since urban freight transport plays a critical role in the quality of urban transportation [6]. The traditional measures to accommodate the growth of freight transport, such as the expansion of the existing transport networks, cannot be pursued, at least in urban areas. Innovative approaches are required to this aim.

Intelligent Transportation Systems (ITS) can play a key role to optimize the organization of urban distribution, thus reducing the impact of freight traffic on urban congestion. For example, Wang et al. [20] use real-time information from different ITS such as RFID and GPS to optimally route and schedule vehicles in logistic and distribution services. Jarugumilli and Grasman [9] use RFID technology to enable efficient control of inventory distribution by exchanging real-time information upon arrival at each location. In order to apply ITS solutions to the optimization of freight distribution, reliable measures and forecasts of road network conditions are needed. Kim et al. [10] propose effective algorithms for data estimation, to be used once measures from the field have been collected. On the other hand, it is necessary to assess that the investment for collecting and processing reliable information produces sufficient Return On Investment (ROI). This key aspect has not been sufficiently investigated in the literature, so far.

The aim of this paper is to quantify the value of information for a practical case study. We define the marginal value of information as the cost reduction achievable through the use of a more reliable source of data with respect to a reference data set. Estimating the value of information is useful both for logistic operators, to evaluate the ROI of reliable data, and for information providers, to fix the best service price and/or to estimate the associated potential market share. Specifically, in this paper we analyze the marginal value of several types of information:

- off-line data (based on historical information),
- real-time data (based on real-time measures of the traffic),
- static data (if measures are assumed to be constant over all the planning horizon),
- time-dependent data (if measures vary over the planning horizon).

In general, time dependent data present a higher reliability with respect to static data since road traffic is typically time-dependent, due to the variability of travel demand in the different time slices of the day. Moreover, real-time data acquisition allows re-optimization of the delivery plan, thus reducing the transport cost with respect to the delivery plan computed by using historical data. Real-time information can be also enriched by short-term forecasts on traffic conditions by using traffic simulators. Therefore, we also analyze the marginal value of short-term forecasts, which may contribute to reduce transport costs while increasing service quality (see, e.g., [12]).

The assessment is carried out for a practical case study arising in the urban freight distribution of perishable goods from an intermodal platform to several retailers located in the city center of Rome, in Italy. The off-line distribution problem is formulated as a Vehicle Routing Problem with soft time windows (VRPTW) for the deliveries, in which
the objective function includes the transportation costs and the cost of late deliveries. The real-time problem arises after that the vehicles have left the platform. This problem can be viewed either as a Traveling Salesman problem with soft time windows (TSPTW), if all the retailers have to be served, or as a Team Orienteering Problem (TOP), if some customers can be skipped in order to reduce transportation cost and penalty cost for missed or late deliveries. The latter problem has been studied by several authors in the last years, including Tang and Miller-Hooks [18], Archetti et al. [1], and Boussier et al. [2].

Among the papers addressing the static VRPTW we cite Solomon, 1987 [17]; Russell, 1995 [16]; Bramel and Simchi-Levi, 1996 [3]; Potvin et al., 1996 [15]; Taniguchi et al., 1998 [19]; Cordeau et al. [5]. As for the time-dependent VRPTW, Ichoua et al. [11] develop a tabu search algorithm, as well as Flamini et al. [7]. The latter algorithm is used in this paper for off-line generation of vehicle routes, based on static as well as time-dependent travel times. As for the real-time problem, Li et al. [13] study the case of trucks involved in severe accidents, when the fleet plan needs to be adjusted in real-time depending on the current state of the network. Lin and Cai [14] develop an adaptive fastest path algorithm capable of efficiently accounting for the network with real-time traffic data. In this paper, two real-time re-optimization algorithms are evaluated for reducing the solution cost based on the acquisition of real-time information with or without updated forecast of traffic evolution. The paper is organized as follows. The assessment of the marginal value of information is described in Section 2. Section 3 deals with models and algorithms used to collect data and to solve off-line and real-time problems. Computational results and the practical assessment for the case study are reported in Section 4. Some conclusions follow in Section 5.

2 The problem

This section describes the practical case study used for the assessment and the methodology adopted in the paper to quantify the cost reduction achievable when the route choices are based on different types of information on traffic network conditions.

2.1 Case study

We consider an intermodal logistic platform LP located in the south area of Rome (Italy), near the Big Ring Road, that distributes fresh goods to retailers spread in the city center. The center of Rome is characterized by narrow streets and high density of small commercial activities. Thus, the distribution in the city center is quite decoupled from the rest of the city since specific small vehicles have to be used in this area (smaller than 3.5 tons).

The perishable goods (e.g., fresh cheese and vegetables) are characterized by the short life time of products, sometimes limited to few days. Thus, each retailer prefers to sell products as soon as possible after they are received and therefore he/she can refuse a delivery after the best time of the day to sell the products. The distribution must therefore comply with strict restrictions on the delivery times. If the delivery is refused, the merchandise is returned to LP and delivered the next day at a smaller price. The depreciation can be a significant fraction of the goods value, since the time elapsed from the production to the expiry date can range to a week or less for some fresh products like the
mozzarella cheese.

In order to plan its deliveries, LP needs data on traffic conditions (specifically, link travel times), that are heavily time-dependent in urban contexts like Rome, due to traffic congestion occurring during the peak hours. Information about traffic conditions are usually collected, elaborated and made available to the market, according to the preference of customers, by service providers.

2.2 The types of information

In Figure 1, a possible architecture for data collection and communication is depicted. Two main categories of data can be provided: off-line data and real-time data. Off-line data can be divided into (i) off-line static data, (ii) off-line time-dependent data.

Let $t_{ij}^{h}$ be the off-line static value of travel time on link $(i, j)$, computed as the mean of historical measures collected on the link during different days in a certain time interval. $t_{ij}^{h}$ can be more or less representative of network conditions depending on the time interval considered to compute the mean value. The off-line time-dependent data is the vector $(t_{ij}^{1h}, t_{ij}^{2h}, ..., t_{ij}^{nh})$ of historical travel times on link $(i, j)$ for each interval $k = 1, \ldots, n$ in which is divided the planning horizon $T$. Clearly, time-dependent data can be more reliable than static data for representing the travel time evolution of link $(i, j)$.

Both the off-line static and the off-line time-dependent travel times do not take into account stochastic events that can change the usual traffic conditions. When such events occur, real-time data can provide more reliable information than off-line data.

Two types of real-time data can be provided, as reported in Figure 1: (iii) the current travel time $t_{ij}^{kr}$ measured on link $(i, j)$ at the $k$-th time interval, and (iv) the vector...
\((t_{kr}^{ij}, t_{ij}^{(k+1)f}, \ldots, t_{ij}^{nf})\) including the real travel time measured on link \((i, j)\) at the \(k\)-th time interval, and a forecast on the travel times occurring during the subsequent \(n-k\) time intervals, updated at time \(k\). The additional forecast service on the traffic evolution can be offered to \(LP\) by the same subject providing real-time information or by a different one, which makes use of dynamic simulation tools.

In our model, the logistic operator plans the deliveries using off-line data, either static or time-dependent, and then possibly re-optimizes the route of each vehicle in real-time based on the current traffic conditions and vehicle position. In the latter case the vehicles must be equipped with GPS/GPRS systems in order to communicate their real-time position and to receive updated delivery plans. If forecasts of the traffic evolution are available in addition to real-time data, routing algorithms based on time-dependent travel times can be used to better estimate the delivery times.

### 2.3 The value of information

In our assessment, we evaluate the benefit of using off-line data characterized by different degree of reliability and the added value generated by re-optimizing the solution with real-time information. Specifically, we analyze four types of data sets, denoted in Figure 1 as: \((i)\) off-line static, \((ii)\) off-line time dependent, \((iii)\) real-time, \((iv)\) real-time with forecasts. Denoting with \(C_x\) the monetary cost of a solution computed by using data set \(x\), we compute the following marginal values:

\[
V_{TD-ST} = C_{(i)} - C_{(ii)}
\]

\[
V_{RT-ST} = C_{(i)} - C_{(iii)}
\]

\[
V_{RT-TD} = C_{(ii)} - C_{(iii)}
\]

\[
V_{F-RT} = C_{(iii)} - C_{(iv)}
\]

These expressions represent the economic benefit achievable if the initial solution is computed with off-line static (ST) or off-line time-dependent (TD) data, and if the solution is then re-optimized in real-time with (F) or without (RT) taking into account forecasts of traffic conditions. In this paper we carry out the assessment for our case study. However, the methodology is quite general and can be applied as well to other distribution contexts.

### 3 Models and algorithms

This section describes models and algorithms used to collect data and to solve off-line and real-time problems

#### 3.1 Network description

The network includes the historical center of Rome, where retailers are located, and the south area until the Big Ring Road, where the logistic platform \(LP\) is located. The network is shown in Figure 2 and consists of 250 centroids, 425 nodes and 2346 oriented links.

The historical center of Rome is characterized by many narrow streets and by a large number of small activities, which translate into specific problem characteristics such as the
dimension and the number of customers and low link capacity values. The distribution of merchandise takes place from 4:00 am to 11:00 am (planning horizon). In order to model the traffic conditions within this time window, about 280,000 vehicles are generated on the network considering the variable demand profile shown in Figure 3.

![Figure 2: Rome network](image)

Figure 2: Rome network

![Figure 3: Demand profile](image)

Figure 3: Demand profile

Link travel times resulting from simulation are obtained every 30 minutes using dynamic assignment model where transport demand and supply characteristics can change during the simulation interval. For the dynamic simulation we use the DYNAMEQ model: this is a dynamic traffic assignment model which exploits variants of gradient like directions and the method of successive averages to determine pre-trip dynamic equilibrium path choices [8]. As a consequence, the travel times between each pair of retailers, as well as between each retailer and the logistic platform $LP$, are time-dependent and can be represented by a vector where each component is associated to a certain time slice.

In our study, we consider these travel times as the historical traffic conditions (i.e., the off-line time-dependent data, thereafter called as TD data). In order to simulate perturbed conditions on the network that can occur during a typical day (i.e., the real condition of the network, thereafter called as RE data), fictitious traffic signals have been activated at 5:00 am and later at 7:00 am in order to generate congestion.
3.2 Off-line problem

The off-line distribution problem is formulated as a vehicle routing problem with soft time windows of [earliest, latest] delivery times, and referred to as VRPTW in the following. The logistic platform LP must distribute the required amount of perishable goods to a given set \( R \) of retailers by using a given set \( V \) of vehicles of given capacity. We assume that an unlimited amount of merchandise and number of vehicles is available at LP. Each retailer \( r \) requests a certain quantity of goods \( d_r \) to be delivered within a given time window \([t_r, T_r]\).

A feasible solution of the problem consists of constructing a route for each vehicle starting and ending in LP such that (i) the demand of each retailer is satisfied, (ii) each retailer is served by exactly one vehicle, and (iii) the capacity of each vehicle is not exceeded. In our model, a vehicle arriving early at a certain retailer will wait until its earliest delivery time.

A vehicle arriving at time \( a_r > T_r \) at retailer \( r \) incurs a penalty cost \( w_r \) for late delivery. This penalty depends on the probability \( p_r \) that the delivery is refused by the retailer. We assume \( p_r = 0 \) for on-time deliveries, i.e., for \( t_r \leq a_r \leq T_r \). The probability that a delivery is refused is \( p_r = 1 \) for a delay \( a_r - T_r \geq \tau_{max} \) and increases linearly from 0 to 1 when the arrival time is in the time window \([T_r, T_r + \tau_{max}]\), as in Figure 4.

If the delivery is refused, the merchandise is returned to the logistic platform and delivered the next day, with an associated cost corresponding to the goods depreciation \( \gamma_r \) and the cost of a new delivery to be performed in the following day. This quantity can be a significant fraction of the goods value, since the time elapsed from the production to the expiry date can range to a week or less for some fresh products. We model \( \gamma_r \) as a function of the demand \( d_r \) to be delivered to retailer \( r \), of the good unit price \( u \) and of the number of days remaining to the expiry date \( g \) as follows: \( \gamma_r = \frac{d_r u}{g} \). To compute the cost of a new delivery we use a rough estimation equal to the total delivery cost of historical data divided by the number of deliveries. We call \( \xi \) this quantity.

The penalty cost \( w_r \) is therefore:

\[
w_r = (\gamma_r + \xi) p_r
\]

![Figure 4: Probability of refusing delivery](image)

Let \( \rho \) be the set of routes in a solution, each associated to the vehicle \( v(\rho_i) \) used for the \( i \)-th route. The cost of route \( \rho_i \) is given by three quantities: (i) the fixed cost \( f_{v(\rho_i)} \) associated to the usage of vehicle \( v(\rho_i) \), (ii) the variable cost \( c_i(\rho_i) \) associated to length of route \( \rho_i \), (iii) the penalty cost \( \sum_{r \in \rho_i} w_r(\rho_i) \) for late deliveries. The objective function
of the VRPTW is therefore:

\[
\min \sum_{i=1}^{\left| \rho \right|} \left[ f_{v}(\rho_i) + c_i(\rho_i) + \sum_{r \in \rho_i} w_r(\rho_i) \right]
\]  

(1)

The algorithms used for computing an initial solution to the VRPTW are the initial greedy and the advanced tabu search algorithm (TS) described in [7]. The TS is executed by using off-line data, either static or time-dependent.

### 3.3 Real-time problem

The best solution found to the off-line VRPTW consists of a vector of routes \( \bar{\rho} \), one for each vehicle, and it is the initial solution used for real-time re-optimization. The real-time problem consists of re-routing each vehicle on the basis of its current position and the traffic conditions. We assume that the freight carried by a vehicle cannot be transferred to another vehicle after that the vehicle has left the platform LP. Therefore, the problem to be solved in real-time is a Traveling Salesman Problem with Time Windows, or TSPTW, in which the cost for the \( i \)-th route using vehicle \( v(\bar{\rho}_i) \) can be expressed as:

\[
f_{v(\bar{\rho}_i)} + \min_{\bar{\rho}_i} \left\{ c_i(\rho_i) + \sum_{r \in \rho_i} w_r(\rho_i) \right\}
\]  

(2)

We compare two algorithms for solving the real-time re-optimization problem TSPTW, called resequence retailers (RR) and skip retailer (SR).

With both algorithms, the initial solution is represented by a set \( \bar{\rho} \) of routes. The \( i \)-th route specifies the sequence of retailers to be served by vehicle \( v(\bar{\rho}_i) \). Let us suppose that at time \( t \) the vehicle has just completed the delivery at the \( j \)-th retailer of \( \bar{\rho}_i \). Note that \( t \) can be different from the value estimated by the initial solution, due to the usage of off-line data instead of real data. In order to improve the current solution from time \( t \) on, we allow the following two real-time strategies:

- **resquence retailers (RR).** The order in which the remaining retailers are served by the vehicle may change due to the updated traffic situation, which makes more convenient following a different route for the same set of remaining retailers.

- **skip retailer (SR).** As in a team orienteering problem, the vehicle can decide to skip the \( (j+1) \)-th retailer in the sequence \( \bar{\rho}_i \) and to serve directly the following one. This strategy can be useful, e.g., if the vehicle is late and serving the next retailer and all the following ones in the original sequence would yield a larger cost with respect to skipping the next retailer and serving the remaining ones. We estimate the new solution cost by using the updated information on the current time and position of the vehicle, and the updated travel times for moving from a retailer to another.

We apply the two re-optimization algorithms to each vehicle whenever it completes a delivery.
4 Numerical experience

This section reports on the performance of the algorithms described in sections 3.2 and 3.3 for the network of Section 3.1. The codes are implemented in C++ and run on a PC equipped with a Intel 2 GHz processor and 2 GB of RAM.

4.1 Data sets generation

In this section, we describe the instances used for the assessment and the data used to simulate the traffic behavior.

In the test cases we consider instances with 50 and 100 retailers. We fix each vehicle load to 70 items. The other parameters, namely the demand of each retailer $r$, $d_r$, the item unitary cost $u$, and the number of days $g$ remaining to the expiry date for each item, are randomly generated with uniform distribution in the following intervals:

- $d_r \in U(1, 10)$ items;
- $u \in U(4, 10)$ euro;
- $g \in U(4, 8)$ days.

About the delivery time windows, two values have been considered in order to evaluate the weight of the delay penalty cost on the objective function: 20 minutes (small time windows) and 2 hours (large time windows).

The input data set, characterized by different degree of reliability and adopted to generate the initial solution of VRPTW, can be summarized as follows (see Figure 5):

- $S4$: off-line static with no congestion (measured at 4:00 am),
- $S9$: off-line static with congested traffic conditions (measured at 9:00 am),
- $TD$: off-line time-dependent,
- $RE$: real data.

Off-line static data corresponds to taking the values in $TD$ at a certain time slice and considering them constant along the whole planning horizon. The real data are considered only as a reference case to allow comparison of the ideal solution with other solutions. Given a demand from the retailers, the advanced tabu search algorithm produces a different initial solution for each of the four input data sets ($S4, S9, TD, RE$). This solution can be re-optimized in real-time by using the SR or the RR algorithms after each delivery. Both SR and RR make use of update information (real-time information). We assume that real-time information of traffic conditions are provided three times in the planning horizon: at 4:00 am (starting routing time), at 5:00 am and at 7:00 (i.e., immediately after perturbations generated by fictitious traffic signals). We consider two types of information available (see Figure 6): (i) real-time updates (at 4:00, 5:00 and 7:00), and (ii) real-time updates plus the forecasts of traffic evolution after each perturbation.

In the former case, the travel times are assumed to be constant and equal to the last value received, until a new real-time update is transmitted (Figure 7). In the latter case, a forecast on the travel times from the last real-time update to the end of the planning horizon is provided. The forecast is computed by a traffic simulator every time a real-time update occurs (Figure 8).
4.2 Computational results

Figure 9 reports the comparison between the estimated and the real value of the objective function for the different initial solutions, computed by the tabu search. By estimated value we mean the value of the objective function (2) for the best solution found by the tabu search algorithm when using the different input data sets of Section 4.1. By real value we mean the value of expression (2) for the same routing solution but computed with the data set $RE$.

As expected, there is a remarkable gap between the estimated and the real costs. The solutions computed with $S_4$, $S_9$ and $TD$ data are apparently better than the ones computed with $RE$ data but the real cost occurring when these solutions are implemented in practice is larger.

The solutions computed using $TD$ data exhibit the smallest gap between estimated and real cost and the smallest real cost with respect to the static $S_4$ and $S_9$ data. Moreover, the solutions obtained with $TD$ data exhibit the smallest real cost, besides the ones obtained with $RE$ data. In fact, $TD$ data are more representative of the network traffic evolution with respect to the rough static data $S_4$ and $S_9$. $TD$ data are therefore the most suitable to build the best off-line solution. On the other hand, $S_4$ data are the cheapest data to collect, since they refer to the empty network conditions. If we consider $S_4$ as the reference case, the marginal value of $S_9$ and $TD$ are $V_{S_9-S_4} = -16\%$ and $V_{TD-S_4} = 22.5\%$, respectively. In other words, there is no convenience in measuring the link travel times in the peak hour $S_9$ with respect to $S_4$. $TD$ is convenient if the cost of collecting such data set is smaller than 22.5% of the total transportation costs computed with $S_4$.

Let us now assess the value of real time re-optimization algorithms. In this case, the four off-line solutions of Figure 9 are the input of the real-time re-optimization algorithms SR and RR.

In Figure 10, the solutions computed by SR are evaluated in terms of real costs, both
Figure 6: Real-time data and forecasts provided for re-optimization

Figure 7: Use of real-time data during the re-optimization procedure

Figure 8: Use of real-time data with forecasts during the re-optimization procedure
for small and large delivery time windows. For the instances with small time windows of Section 4.1, the results are described in Figure 10(a). In this case, SR causes a remarkable reduction of the real cost with respect to the initial solution with and without real-time forecast. It is interesting to observe that SR produces a cost reduction also when starting from the initial solution computed with $RE$. Clearly, this cost reduction does not depend on the availability of real-time data since real-time data coincide with $RE$. Therefore, the reduction is totally due to the possibility of skipping retailers in real-time, while this is forbidden in the off-line problem. This means that there would be a cost reduction also by allowing the solution to skip some retailers in the off-line problem, i.e. by formulating the off-line problem as a TOP rather than a VRPTW.

As for scenarios S4, TD and S9, we observe that SR produces a mean cost reduction of 32.90% with respect to the off-line solution, and in particular of 32.02% for SR without forecasts and 33.77% for SR with forecasts. The mean number of skipped retailers is 12.7. The strongest cost reduction is obtained in case of S9 data, in which SR algorithm is very effective in improving the rough off-line solution, achieving a final performance quite close to the other data sets. The marginal value $V_{F-RT}$ of the real-time forecasts is very small in all cases, ranging from 7 euro (S9 data) to 17 euro TD data).

The results for large delivery time windows are reported in Figure 10(b). In this case, the solutions are characterized by small delays and the total cost (2) is dominated by the fixed vehicle cost and by the variable cost associated to route lengths. Therefore, the
Figure 10: Effect of the SR algorithm
effect of SR algorithm is mainly limited to reducing the cost related to the length of the routes. The mean percentage cost reduction generated by SR with respect to the off-line solution is limited to 8.37% on average and the mean number of skipped retailers is equal to 2.3.

Also for large time windows, the marginal value $V_{F-RT}$ of real-time forecasts is quite limited. It is always below 3% of total delivery cost with respect to using real-time data without forecasts.

Figure 11 shows the real performance of RR algorithm when starting from different off-line solutions and for the two cases of small and large delivery time windows, denoted with (a) and (b) in figure. The percentage cost reductions with respect to the off-line solutions are lower than those obtained with SR algorithm, both for small and large time windows. In particular, in the former case RR improves the off-line solution by 10.52% (9.35% without forecasts and 11.68% with forecasts). In the case of large delivery time windows, the mean percentage improvement decreases to 6.38%. Also for RR algorithm, the marginal value $V_{F-RT}$ of using real-time forecast is at most 3% of the total delivery cost without forecast.

Summarizing the above results, with both SR and RR the best solution is always computed by starting from $TD$ data and using the real-time forecast. However, the marginal value of $TD$ with respect to the reference case $S4$ with small time windows is $V_{TD-S4} = 39.6\%$ for SR and only $V_{TD-S4} = 25.7\%$ for RR. The largest part of this value is due to the use of TD data set and only a minor part is due to the real-time re-optimization algorithm.
4.3 The information value

This section deals with the computation of the marginal value of each combination of data set and real time re-optimization strategy. In order to build a feasible delivery plan, the cheapest data set is S4, which can be determined by simply estimating the travel time of each link in the empty network conditions. For the case of small delivery time windows, an average solution cost of 821 euro is obtained from our experiments. When replacing S4 data with the most reliable and accurate TD data, the solution cost decreases to 636 euro, with a marginal value of 185 euro (22.5% of the daily distribution cost). It is interesting to analyze the S9 dataset, for which the data collection is slightly more expensive than S4. Replacing S4 with S9 produces an increase of the distribution costs, from 821 to 956, so the latter data set is strictly dominated by S4. The same dominance relation can be observed in figures 10 and 11 when using S9 in combination with all real-time re-optimization algorithms. Therefore, S9 will not be further analyzed.

Algorithm SR in combination with S4 makes possible a cost reduction of 280 euro (without forecast) and 295 euro (with forecast) for small delivery time windows, thus resulting in a marginal value of 34.1% for SR plus an additional value of 2.85% for the short term forecast.

It is interesting to observe that the marginal value of SR depends on the data set chosen for the off-line solution. When using the reliable TD dataset, the marginal value
Table 1: Percentage difference between the cost of each strategy and TD, using SR

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>initial</th>
<th>without forecasts</th>
<th>with forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4</td>
<td></td>
<td>+22.53%</td>
<td>+39.59%</td>
<td>+41.66%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>+8.32%</td>
<td>+11.46%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>+8.94%</td>
</tr>
<tr>
<td>TD data</td>
<td></td>
<td>0%</td>
<td>+22.01%</td>
<td>+24.69%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>0%</td>
<td>+3.43%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>0%</td>
</tr>
<tr>
<td>S9</td>
<td></td>
<td>+33.47%</td>
<td>+48.12%</td>
<td>+49.90%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>+13.44%</td>
<td>+14.11%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>+15.37%</td>
</tr>
<tr>
<td>RE</td>
<td></td>
<td>-27.97%</td>
<td>-0.20%</td>
<td>-3.62%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>-10.71%</td>
<td>-6.92%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>-8.86%</td>
</tr>
</tbody>
</table>

Table 2: Percentage difference between the cost of each strategy and TD, using RR

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th>initial</th>
<th>without forecasts</th>
<th>with forecasts</th>
</tr>
</thead>
<tbody>
<tr>
<td>S4</td>
<td></td>
<td>22.53%</td>
<td>+25.70%</td>
<td>+27.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>+18.88%</td>
<td>+20.48%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>+18.08%</td>
</tr>
<tr>
<td>TD data</td>
<td></td>
<td>0%</td>
<td>+4.09%</td>
<td>+5.97%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>0%</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>0%</td>
</tr>
<tr>
<td>S9</td>
<td></td>
<td>+33.47%</td>
<td>+36.19%</td>
<td>+37.45%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>+24.41%</td>
<td>+25.90%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>+23.63%</td>
</tr>
<tr>
<td>RE data</td>
<td></td>
<td>-27.97%</td>
<td>-22.74%</td>
<td>-20.32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>without forecasts</td>
<td>-22.74%</td>
<td>-20.32%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>with forecasts</td>
<td>-</td>
<td>-20.32%</td>
</tr>
</tbody>
</table>

of SR without forecast is only 22% for the instances with small time windows. Therefore, if TD data is used by the logistic platform to develop off-line routes, SR is convenient only if this value is strictly greater than the cost needed to equip all vehicles with a GPS-GPRS system plus the cost for collecting real time information.

Table 1 reports the systematic comparison between the costs of every combination of dataset and real-time strategy with the cost of TD dataset, both with and without real-time re-optimization using algorithm SR. A similar comparison is carried out in Table 2 when the real-time re-optimization is based on algorithm RR.

The values reported in tables 1 and 2 can be used by the logistic operator to define the best configuration for equipment, data and planning algorithms, given their respective costs. At the same time, from the point of view of service providers, tables 1 and 2 suggest
the maximum value of the information and therefore the maximum price at which the data can be provided.

5 Conclusions

This paper evaluates the benefit achievable when route optimization is based on different configurations of off-line data sets and planning algorithm, and real-time data sets and re-optimization algorithms. The methodology is carried out with reference to a practical case study arising in the freight distribution of perishable goods, but it can be extended to other distribution contexts as well.

The numerical experience shows that the marginal value of a certain data set can be almost independent from its cost. For example, $S_9$ data is likely more expensive than $S_4$ but the marginal value $V_{S_9-S_4}$ is negative. Therefore, special attention should be devoted to estimate the value and the potential ROI of information before investing in a specific ITS technology. We believe that the results of this paper can be of interest also for information providers, to estimate the potential market share and the best price for a specific type of information.

This paper is only a first attempt to define an effective methodology to quantify the value of information in the distribution market. A number of issues remain that need further research. Different models and algorithms should be tested for solving the off-line problem with the same data set, to better quantify the value of information. Moreover, it would be interesting to quantify the value of different real-time re-optimization algorithms for different practical contexts and different information reliability.

References


