Data allocation and application for time-dependent vehicle routing in city logistics

Jan Fabian Ehmke [1*], Dirk Christian Mattfeld [1]

[1] *University of Braunschweig – Carl-Friedrich Gauss Department*
Business Information Systems – Decision Support Group

Abstract

In city logistics, service providers have to consider dynamics within logistics processes in order to achieve higher schedule reliability and delivery flexibility. To this end, city logistics routing demands for time-dependent travel time estimates and time-dependent optimization models. We consider the process of allocation and application of empirical traffic data for time-dependent vehicle routing in city logistics with respect to its usage. Telematics based traffic data collection and the conversion from raw empirical traffic data into information models are discussed. A city logistics scenario points out the applicability of the information models provided, which are based on huge amounts of real traffic data (FCD). Thus, the benefits of time-dependent planning in contrast to common static planning methods can be demonstrated.

Keywords: City Logistics; Time-dependent; Data analysis; Data mining; Vehicle routing.

1. Introduction

City logistics is about logistics in urban areas. The focus is on concepts for fast and reliable transportation of goods in terms of cost-efficient and environmentally acceptable pickup and delivery routes. Nowadays, service providers have to consider dynamics within logistics processes, e.g. shorter delivery times, higher schedule reliability and delivery flexibility (Windt and Hülsmann (2007)). Furthermore, city logistics service providers compete against other road users for the scarce traffic space of inner cities. In conurbations, traffic infrastructure is often used to capacity.

In wide area networks, vehicle routing is usually based on distances. However, vehicle routing in city logistics networks demands for time-dependent travel time estimates for every route section. Vehicle routing based on actual travel time estimates requires empirical traffic data as a key input. Up to now, empirical traffic data has not

* Corresponding authors: J. F. Ehmke (j-f.ehmke @tu-bs.de)

been available to a sufficient extent due to prohibitive census costs. Recently, such data arises from modern vehicular communication networks.

In order to benefit from telematics based data collection, time-dependent travel time estimates have to be integrated into time-dependent vehicle routing frameworks. Whereas static approaches are well studied, time-dependent vehicle routing forms a field of potential research. Especially the preparation and the integration of time-dependent data into vehicle routing approaches is rarely focused:

- For city logistics, Fleischmann et al. (2004) design a traffic information system. Flow and speed data are collected in a field test with stationary measurement facilities and specially equipped vehicles in the metropolitan area of Berlin, Germany. The data is then aggregated and utilized in savings and insertion route construction methods. Here, the data collection methods used have surpassed by progress in technology.
- Eglese et al. (2006) refer to Floating Car Data (FCD) for time-dependent routing in a supra-regional road network in the UK. The FCD originate from a communication network consisting of trucks and coaches. Data is transmitted via text messages (SMS) and stored as a "road timetable" in a central database. In city logistics context, text messages are not appropriate for data collection due to high communication costs. The complexity of urban traffic and its variety of influences require a very thorough data collection method.
- Van Woensel et al. (2008) consider queuing theory to provide time-dependent travel time estimates. They refer to a tabu search approach to solve the time-dependent capacitated vehicle routing problem. Donati et al. (2008) focus on ant colonies to heuristically solve time-dependent vehicle routing problems. Both publications are more focused on large area networks.

In this paper, the process of allocation and application of empirical traffic data for time-dependent vehicle routing in city logistics is considered with respect to its usage. We discuss data collection and the conversion from raw empirical traffic data into information models (Section 2). An application example compares several information models based on real traffic data regarding its benefits for time-dependent route planning (Section 3). Then, the integration of information models into time-dependent vehicle routing frameworks is discussed (Section 4). Finally, the paper is concluded (Section 5).

2. From data to information models

The starting point for the estimation of time-dependent travel times for city logistics is the collection of traffic data. Reliable decisions must be derived from this raw data. Therefore, empirical traffic data has to be transformed into time-dependent information models. We shortly sketch the phases of the corresponding data chain and focus on the two main steps in terms of first and second level aggregation.

2.1. Data chain

Varying traffic flows require time-dependent routing decisions in cities. GPS based traffic data may be the source for the derivation of such decisions. The data chain ranging from GPS based collection of raw traffic data to time-dependent routing decisions is shown in Figure 1. Efficient decisions are enforced by the transformation of raw data into first level aggregated data into second level aggregated data. In particular, the elements involved are as follows:

- data collection: Taxi-FCD is a recent GPS based data collection method that provides raw traffic data in urban areas (cf. Section 2.2). Taxi-FCD result in a large data volume of city-wide traffic data, mainly based on the use of taxis as moving data sources.
- data cleaning: Erroneous data records are removed. E.g. obviously unrealistic speed observations due to GPS shadowing effects are filtered. Data cleaning is the precondition for reasonable data mining.
- data integration: The collected single Taxi-FCD records (empirical traffic data) are amended by a common digital roadmap (infrastructure data). The data is integrated into one database and aggregated for analysis purposes (first level aggregation, cf. Section 2.3).

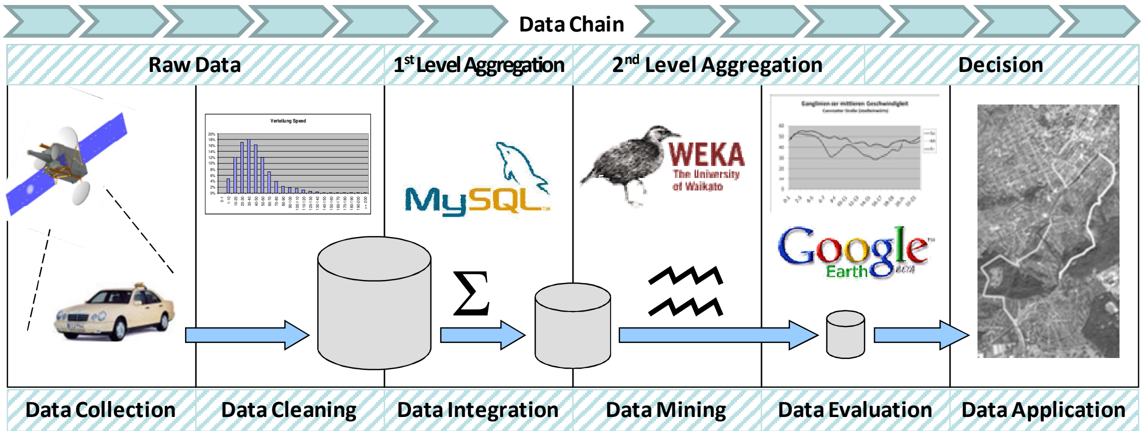


Figure 1: Data allocation and application in the context of GPS based data collection and time-dependent vehicle routing in city logistics (Ehmke et al., 2009).

- data mining: Aggregated Taxi-FCD are analyzed by means of cluster analysis. Cluster analysis is used for the allocation of time-dependent travel time estimates (second level aggregation, cf. Section 2.4). Widely-used algorithms are available in software packages like WEKA (Witten and Frank (2005)), also being used for this research.
- data evaluation: The travel time estimates are subject to evaluation and presentation because routing algorithms require realistic travel time estimates. To this end, the travel time estimates are visualized in daily courses in order to be compared with typical traffic patterns. Furthermore, geographical information systems like Google Earth (Google Earth KML 2.0) are involved. In this paper, we refer to a computational routing experiment in order to evaluate the information models from first and second level aggregation in contrast to travel time estimates based on a common digital roadmap (cf. Section 3).
- data application: The final step comprises the integration of time-dependent travel time estimates in time-dependent optimization frameworks (cf. Section 4).

2.2. Raw data

In order to derive travel time estimates for city logistics, city-wide data collection is necessary. To this end, conventional data collection methods are of limited use because they require a tremendous amount of effort (Gühnemann et al. (2004)). Typically, for vast parts of the city road network no data samples are available.

Traffic data can be collected by using telematics systems in terms of FCD. Recently, GPS have been propagated and are widely used for routing purposes. The Taxi-FCD project run by the German Aerospace Center (DLR) implements the idea of using taxis as mobile data sources for the collection of FCD. Here, a fleet of taxis characterized by typically high mileage is the basis of the system. The taxis are already equipped with GPS based navigation systems used for taxi disposition, hence causing no further costs for data transmission. Taxis transmit their current positions approximately every minute via digital radio trunking. A more detailed description of the data collection method can be found in Brockfeld et al. (2007). DLR has developed several map-matching and data handling algorithms. For the processing of the raw data, a general overview on Taxi-FCD and applications see Lorkowski et al. (2004) and Lorkowski et al. (2005). The structure of a resulting FCD record serving as input for the determination of time-dependent travel time estimates is shown in Table 1.

Table 1: Structure of Floating Car Data records.

TIME	*LINK*	*SPEED*
time of positioning	road segment ID	determined speed [km/h]
2003-08-01 07:01:22	54362718	50.73

2.3. First level aggregation

The raw Taxi-FCD records are filtered, integrated into a single database and then precalculated in terms of time-dependent aggregation. Here, raw traffic data evolve into planning data (first level aggregation). FCD speed averages are calculated for each link $\boldsymbol{l}$ according to the harmonic mean of the single measurements $v_{lw} = n / \sum_{i=1}^{n} \frac{1}{v_{lwi}}$, with $\boldsymbol{n}$ being the number of speed measurements for $\boldsymbol{l}$ within time interval $\boldsymbol{w}$ and $\boldsymbol{v_{lwi}}$ being the single vehicle speed. The result is a mean FCD speed $\boldsymbol{v_{lw}}$ for each link $\boldsymbol{l}$ and time interval $\boldsymbol{w}$, being the fundament for the derivation of time-dependent travel time estimates. If the data coverage is rather low, the median can be taken as a robust representation of $\boldsymbol{v_{lw}}$, alternatively.

The total number $\boldsymbol{W}$ of time intervals must be determined according to the requirements of the city logistic application. In the literature, several selections of $\boldsymbol{W}$ have been defined. E.g., Eglese et al. (2006) refer to 15 time intervals of different length per day ($\boldsymbol{W}$ = 15 x 7), whereas Ichoua et al. (2003) use 3 time intervals after “careful observation” of real traffic data.

Corresponding to common analysis methods from the area of traffic research (e.g. Pinkofsky (2006)), we refer to FCD by establishing 24 time intervals per day ($\boldsymbol{W}$ = 24 x 7). The resulting speed averages are referenced to as “FCD hourly average” (FH) (see Table 2). FH is supposed to cover expected fluctuations in travel times during 24 hours of the day and 7 days of the week.

Table 2: FCD hourly averages (FH) example for a link [km/h].

Time	*0-1*	*1-2*	…	*3-4*	…	*8-9*	…	*16-17*	…	*22-23*	*23-24*
Sun	47	52	…	55	…	50	…	42	…	47	50
Fri	47	52	…	52	…	33	…	31	…	44	46

For interpretation and analysis purposes, FH data can be transformed by normalization procedures in the following two ways. A min-max normalization into the interval [0,1] can be interpreted as daily course of speed variation for every link with focus on relative minimum and maximum speed. On the contrary, a normalization based on means describes daily courses of speed variation with focus on average speeds.

FH based planning algorithms for vehicle routing must cope with a huge amount of travel time estimates. However, limited memory capacities, complex vehicle routing algorithms and the desire for fast and reliable routing decisions require the reduction of the volume of input data without a significant decrease of reliability. The following cluster analysis approach responds to these requirements by providing weighted FCD averages in terms of second level aggregation.

2.4. Second level aggregation

In this section, we introduce data mining as important component of the data chain. We focus on weekday dependent clustering of FH data for the efficient allocation of time-dependent travel time estimates. The resulting clusters are supposed to characterize included links with similar speed variations. Thus, the data input for vehicle routing algorithms can be reduced by data mining (second level aggregation).

Table 3: Normalized FCD hourly averages example for a link.

Time	*0-1*	*1-2*	…	*3-4*	…	*8-9*	…	*16-17*	…	*22-23*	*23-24*
Sun	0.4	0.8	…	1.0	…	0.7	…	0.1	…	0.4	0.6
Fri	0.8	1.0	…	1.0	…	0.2	…	0.1	…	0.6	0.7

Normalized FH data (cf. Table 3) are clustered by the use of a $\boldsymbol{k}$-means algorithm (MacQueen (1967)). The $\boldsymbol{k}$-means algorithm is a partition-based clustering algorithm, requiring the number $\boldsymbol{k}$ of desired clusters and a distance function as input. The algorithm then iteratively minimizes the error sum of the data objects' distances to the cluster centers.

We parameterize the $\boldsymbol{k}$-means algorithm by a Euclidean distance function. In order to achieve a meaningful estimation of $\boldsymbol{k}$, we propose the determination of internal indices evaluating the quality of a clustering. The idea is to repeat a clustering with ascending size of $\boldsymbol{k}$, and then compare the results of the index (Tan et al. (2006), Jain and Dubes (1988)), e.g. the error amount of squares in partitioning clustering approaches. The general trade-off is as follows: On the one hand, $\boldsymbol{k}$ must be large enough to give a good approximation of the actual link travel time variations. On the other hand, $\boldsymbol{k}$ should be kept as small as possible in order to minimize the input data for routing algorithms.

Clustering of normalized FH data leads to a compact representation of time-dependent travel time estimates in terms of discount factors. The discount factors represent typical speed variations for the derivation of time-dependent travel time estimates. The main idea is to look up a link's discount factor and then weight a robust speed figure (e.g. average speed or maximum speed). Thus, the resulting information model is referenced as "Floating Car Data Weighted Averages" (FW).

Table 4: From the cluster analysis resulting discount factors (example with k = 4).

Cluster	*0-1*	*1-2*	*…*	*3-4*	*…*	*8-9*	*…*	*16-17*	*…*	*22-23*	*23-24*
1	0.71	0.76	...	0.79	...	0.25	...	0.23	...	0.59	0.65
2	0.64	0.65	...	0.63	...	0.52	...	0.53	...	0.64	0.64
3	0.43	0.46	...	0.55	...	0.30	...	0.27	...	0.38	0.41
4	0.78	0.76	...	0.77	...	0.48	...	0.45	...	0.71	0.75

An example result of clustering is given in Table 4. Each cluster represents a group of links. Each link is associated with its groups' representative vector of 24 discount factors. Thus, time-of-the-day specific link speeds can be derived by using the time-of-the-day specific discount factor and weighting it by its link's mean speed. This leads to enormous savings regarding input data for vehicle routing.

Table 5: Comparison of required volume of input data regarding different information models.

Information model	digital road map	FCD hourly average (FH)	weighted FCD average (FW)
Input data	n	t x d x n	n + (t x d x k)
Input data Stuttgart (t = 24, d = 7, n = 100 000, k = 4)	100,000	16,800 000	100,672
Input data Stuttgart per link	1	168	1.01

Notes: n = number of links, t = number of time intervals, d = number of days, k = number of clusters.

In Table 5, the resulting information models are compared in terms of the volume of input data for routing algorithms. Therefore, the resulting volume of input data is described formally and instantiated with specific figures from empirical FCD of Stuttgart, Germany (cf. Section 3). For a comparison of the information models from an algorithmic point of view, we point out the resulting amount of input data per link. Whereas a common digital roadmap would result in 100,000 travel time estimates or one travel time estimate per link, the FH information model amounts up to 16,8 millions travel time estimates to be considered. The FW approach results in 100,672 travel time estimates, indicating an upper bound of data reduction for the Stuttgart example from 168 travel time values (FA) to only 1.01 travel time values (FW) per link.

3. Example city logistics application

In this section, we provide time-dependent information models and evaluate its benefits in a city logistics context. The advantages of time-dependent planning for the reliability and the robustness of planned routes are demonstrated, contrasting disadvantages of common static planning methods. Therefore, a huge amount of FCD from the area of Stuttgart, Germany from the years 2003-2005 is processed as described before. All in all, about 230 million data sets are analyzed and aggregated in terms of first and second level aggregation, leading to FH and FW information models.

3.1. Experimental setting

The experimental setting comprises a fictitious city logistics application for the area of Stuttgart, Germany, as well as the usage of several information models and several traffic scenarios. We plan and simulate itineraries that are affected by a high fluctuation of travel times, thus requiring the consideration of time-dependency. The following questions are discussed:

- Which information model leads to the realization of the time-shortest itinerary?
- Which information model results in the most reliable travel time prediction?

To this end, several kinds of travel time estimates are used for route planning. We compare the resulting itineraries with respect to the realization of the fastest route and the most reliable travel time estimation. The itineraries' realization is done by simulation of the planned routes. Here, simulation means the recalculation of planned routes based on "true" travel times for specific days and time slots. The required "true" travel times have its seeds in a travel time database provided by the DLR, resulting from calendar date specific FCD of about 40 Mondays. Itinieraries are calculated using *Dijkstra's algorithm* (Dijkstra (1959)). Each route is denoted by its anticipated duration and its network links used.

The following *information models* provide travel time estimates for route planning:

- Digital Roadmap (DR): Travel time estimates associated with the links of the digital roadmap of Stuttgart serve as a static benchmark for the following FCD based travel times.
- FCD hourly averages (FH): Travel time estimates resulting from first level aggregation, depending on day of the week and time of the day.
- FCD weighted averages (FW): Travel time estimates resulting from second level aggregation, depending on day of the week and time of the day.

The following *representative traffic network scenarios* are considered for route planning. They have been identified by the observation of typical traffic states in urban areas (Ehmke et al. (2008)):

- "free flow network" (Monday 3-4 am),
- "early rush hour" (Monday 8-9 am),
- "average traffic" (Monday 11-12 am) and
- "late rush hour" (Monday 4-5 pm).

Itineraries are calculated for each combination of traffic network scenario and information model. Thus, the influence of time-dependency on planning quality can be demonstrated.

3.2. Example itinerary: airport – main station

Planning results for one interesting example itinerary are presented in Table 6. The given itineraries comprise routes between the outskirts of the city (airport) and Stuttgart main station (downtown). Generally, these routes are heavily frequented. Hence, a high fluctuation of the travel time is to be expected, enforcing the consideration of time-dependency in route planning.

Table 6: Extract from computational results for one route (airport – main station).

	DR: digital roadmap					*FH: FCD hourly averages*					*FW: Weighted FCD averages*				
	anticipated	simulated	diff	diff	std-dev	**anticipated**	simulated	diff	diff	std-dev	**anticipated**	simulated	diff	diff	std-dev
	(min)	(min)	(min)	(%)	(min)	**(min)**	(min)	(min)	(%)	(min)	**(min)**	(min)	(min)	(%)	(min)
3-4 am		14.8	3.4	29 %	1.89	**14.4**	14.2	0.9	6%	0.89	**14.1**	14.1	1.0	7%	0.95
8-9 am	**11.4**	22.2	10.8	95%	2.75	**19.7**	18.9	1.5	8%	1.31	**20.9**	23.2	3.4	16%	3.34
11-12 am		17.3	5.9	51%	2.27	**17.5**	17.2	1.2	7%	1.04	**18.4**	17.7	1.9	10%	2.06
4-5 pm		23.5	12.1	106%	3.17	**18.1**	18.1	1.4	8%	1.54	**20.6**	18.4	2.7	13%	1.37

In the digital roadmap case (DR), there is only one anticipated duration resulting from only one static travel time estimate per link, whereas FH and FW data allow for time-dependent travel time anticipations (cf. columns "anticipated"). The use of DR data for route planning leads to a time-independent travel time anticipation of 11.4 minutes and an average simulated route duration of e.g. 22.2 minutes (cf. "simulated", based on FCD "true" travel time) in the early rush-hour (8-9 am). The simulated route durations differ 10.8 minutes (or 95%, cf. "diff") from the anticipated route duration and are characterized by a standard deviation (cf. "std-dev") of 2.75 minutes. In contrast, FH based planning leads to a time-dependent travel time anticipation of 19.7 minutes and a simulated route duration of 18.9 minutes in average, which represents a significant decrease in differences and an increase in planning reliability (8-9 am). Differences are higher in the FW case (e.g. 16% in the early rush-hour), which is reasoned by a loss of accuracy due to the aggregation by cluster analysis.

Altogether, FH data leads to itineraries characterized by short durations and high planning reliability throughout all traffic scenarios. FW information models lead to slightly higher differences than in the FH case, but much lower than in the DR case. Static information models like the common digital roadmap can hardly reflect fluctuations in travel times.

3.3. Overall results

An aggregated overview on all computational results is given in Table 7. Experiments are based on 48 city logistics scenarios and about 4000 route simulations. Here, the information models are compared in terms of overall quality (cf. "mean difference") and reliability (cf. "mean std-dev"). The columns "mean difference" and "mean std-dev" denote the average difference from "true" travel times as a figure for planning quality and the mean standard deviation within this difference as a figure for planning reliability, respectively.

Table 7: Comparison of several information models.

	mean difference	mean std-dev	duration		difference		standard deviation	
			count	%	count	%	count	%
DR	72%	3.10	5	10%	0	0%	2	4%
FH	8%	1.47	48	100%	48	100%	47	98%
DR	72%	3.10	11	23%	1	2%	6	13%
FW	13%	2.08	42	88%	47	98%	42	88%
FH	8%	1.47	43	90%	44	92%	37	77%
FW	13%	2.08	15	31%	4	8%	11	23%

The DR information model features relatively high differences (72%) between anticipated and simulated routes in comparison with FH (8%) and FW (13%) information models. Thus, planning quality is relatively higher in the FH and FW case. The same is valid for the "mean standard-deviation", which is the smallest in FH case (1.47 min) and the highest in the DR case (3.10 min).

In the remaining six columns, the information models are compared to each other pair-wisely with respect to the provision of the number of actually fastest routes ("duration"), actually smallest differences ("difference") and actually smallest standard deviation ("standard deviation"). E.g. DR based planning results in the actual fastest route in only 5 cases (or 10 %), whereas FH based planning provides the relatively fastest route in all cases. The main conclusion is that FH as a comprehensive information model leads to relatively reliable and short routes. However, the cluster analysis based FW approach is useful for time-dependent planning, resulting in much more reliable and shorter routes (88 %) than in the static DR case and reducing effort for planning algorithms.

4. Integration of information models for time-dependent vehicle routing

As subsequent step to data analysis, this section provides an overview on the characteristics of time-dependent problem formulations. The main goal is to integrate the information models presented into a time-dependent optimization framework. In "static" city logistics networks, customers are usually represented by vertices. Vertices are connected by edges which represent shortest paths between customers in terms of distances or static travel time estimates. Each edge is associated with its static cost, duration or travel time.

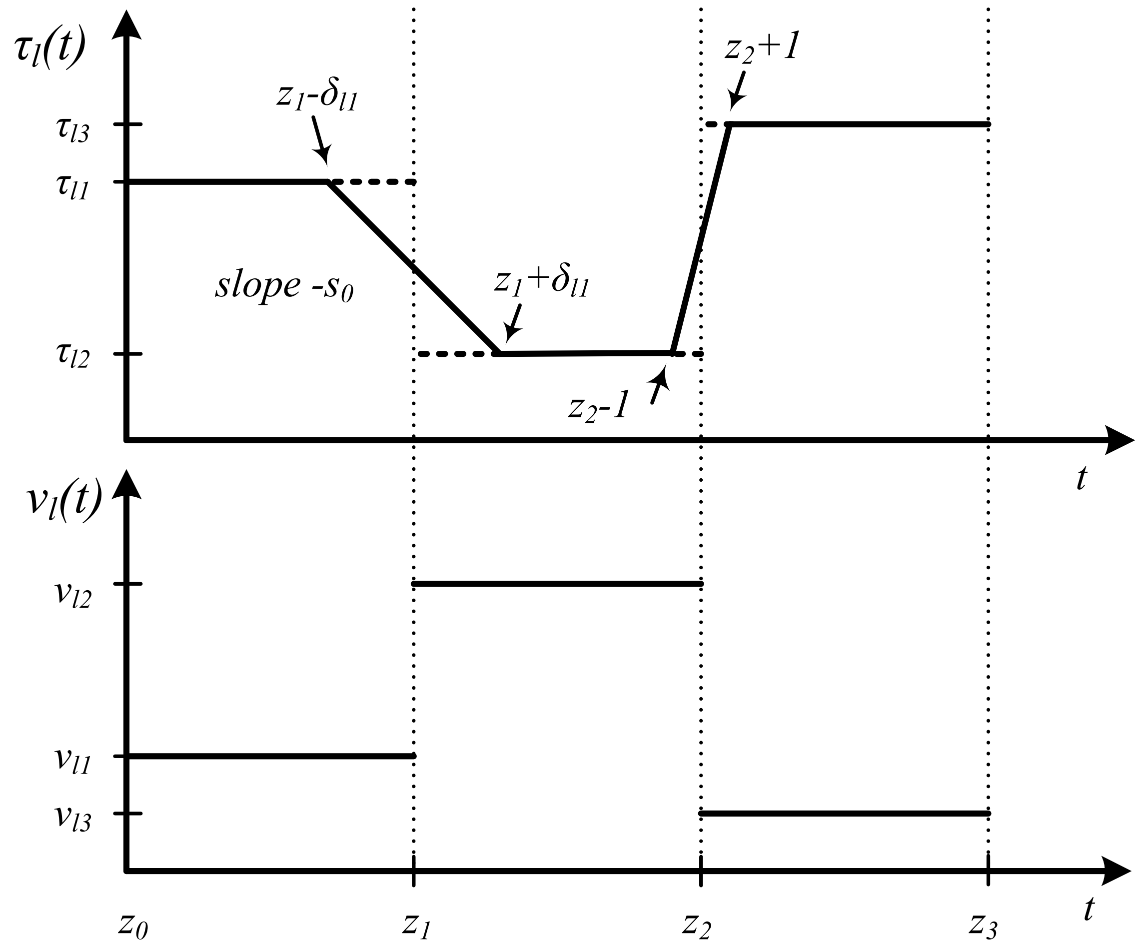
In contrast to static networks, time-dependent networks have to consider varying travel time estimates for each edge. Thus, the travel time is modeled as a function of departure time, which is to be determined in advance. Travel time functions are distinguished into integer and real valued functions, also known as discrete and continuous modeling (Dean (1999)).

In a discrete setting, time-dependent travel time estimates are usually approximated by piecewise-linear functions. Therefore, the time horizon considered has to be partitioned into an appropriate number of time intervals $\boldsymbol{W}$. In the continuous case, travel time functions are estimated based on e.g. empirical traffic data.

Travel time estimation based on empirical traffic data does not guarantee FIFO behavior. In FIFO networks, vehicles do not "pass" each other, i.e. vehicles arrive in the order they commence an edge ("non-passing condition", Kaufman and Smith (1993), Ichoua et al. (2003)). FIFO networks allow for time-dependent shortest path calculation in terms of a trivially-modified variant of any label-setting or label-correcting static shortest path algorithm like Dijkstra's algorithm (cf. Dijkstra (1959)). This is due to the following properties, resulting in reduced complexity of time-dependent vehicle routing (Dean (1999)):

- In FIFO networks, waiting at nodes delays arrival.
- In FIFO networks, one always finds shortest paths which are acyclic.
- In FIFO networks, one always finds shortest paths whose subpaths are also shortest paths.

Figure 2: Construction of the piecewise linear travel time function, including linearization at jumps z_1 and z_2.



The information models presented lead to piecewise-linear travel time functions, ignoring the FIFO property. Thus, the travel time function jumps between two time intervals, and passing may occur if the travel time decreases. Fleischmann et al. (2004) solve this problem by introducing a "smoothed" travel time function that transforms non-FIFO into FIFO networks. Therefore, the jump between two intervals is linearized.

In Figure 2, the derivation of travel times τ_{li} from average speeds v_{li} is illustrated for an example link l. The travel time function τ_l(t) results from the FH or FW information model and features several jumps at z_i. E.g. at z_1, the average speed changes from relative low to relative high speed, inducing a rather long or rather short travel time, respectively. This change is not FIFO valid; a vehicle starting shortly before z_1 would be overtaken by a vehicle starting shortly after z_1. Fleischmann et al. (2004) handle these jumps by linearizing the travel time function in the range $[z_i - \delta_{li}; z_i + \delta_{li}]$. The corresponding slope $-s_0$ is not allowed to become larger than $s = 1$, assuring the FIFO property. In the case of increasing travel times, the slope can be chosen freely.

More on basic concepts of time-dependent shortest paths can be found in a survey paper by Dean (1999). Pallottino and Scutella (1997) give an application driven overview on shortest path algorithms. A recent overview on algorithms for both the discrete and continuous case as well as a performance comparison is provided by Ding et al. (2008). Furthermore, Dell'Amico et al. (2008) introduce an approach for non-FIFO time-dependent networks.

5. Conclusions

This paper considers the allocation and application of travel time estimates for time-dependent vehicle routing in city logistics. Here, telematics based traffic data is converted into time-dependent planning data. An approach for sophisticated data analysis including data clustering and reduction of memory requirements for routing algorithms is introduced. The resulting information models are compared regarding usefulness for time-dependent route planning in terms of a fictitious city logistics example, based on real traffic data. The time-dependent calculation of shortest paths is identified as important part of an efficient and realistic representation of the road network typology for time-dependent planning problems.

The experiments underline the superiority of time-dependent information models (FH, FW) over common static data sets. FCD based time-dependent travel time estimates lead to more reliable routing results compared to route planning based on common digital roadmaps. The data mining approach presented provides time-dependent travel times in a memory efficient way without a significant reduction of the itineraries' reliability and robustness.

This paper focused data collection and application dependent provision of time-dependent travel time estimates. Based on this contribution, time-dependent optimization frameworks can be built, which integrate time-dependent information models and time-dependent optimization models. Thus, reliable delivery services can be provided.

References

Brockfeld, E., Lorkowski, S., Mieth, P. and Wagner, P. (2007) "Benefits and Limits of Recent Floating Car Data Technology – an Evaluation Study", 11th WCTR Conference, Berkeley, USA.

Dean, B.C. (1999) *Shortest Paths in FIFO Time-Dependent Networks: Theory and Algorithms*, Massachusetts Institute Of Technology.

Dell'Amico, M., Iori, M. and Pretolani, M. (2008) "Shortest paths in piecewise continuous time-dependent networks", *Operations Research Letters* 36: 688-691.

Dijkstra, E. W. (1959) "A note on two problems in connexion with graphs", *Numerische Mathematik*, 1: 269-271.

Ding, B., Yu, J.X. and Qin, L. "Finding Time-Dependent Shortest Paths over Large Graphs", Proceedings of the 11th international conference on Extending database technology: *Advances in database technology.*

Donati, A.V. (2008) "Time-dependent vehicle routing problem with a multi ant colony system", *European journal of operational research,* 3: 1174-1191.

Eglese, R., Maden, W. and Slater, A. (2006) "A road timetable to aid vehicle routing and scheduling", *Computers and Operations Research*, 12: 3508-3519.

Ehmke, J.F., Meisel, S. and Mattfeld, D.C. (2008) "Floating Car Data Based Analysis of Urban Travel Times for the Provision of Traffic Quality", Proceedings of *IWTDCS 08*, Barcelona.

Ehmke, J.F., Meisel, S., Engelmann, S. and Mattfeld, D.C. (2009) "Data chain management for Planning in City Logistics", *Int. Journal of Data Mining, Modelling and Management*, 1(4): 335-356.

Fleischmann, B., Gietz, M. and Gnutzmann, S. (2004) "Time-Varying Travel Times in Vehicle Routing", *Transportation Science*, 2: 160-173.

Google Earth KML 2.0. http://earth.google.com/kml/kml_toc.html.

Gühnemann, A., Schäfer, R.-P., Thiessenhusen, K.-U. and Wagner, P. (2004) "Monitoring Traffic and Emissions by Floating Car Data", *Working Paper ITS-WP-04-07*, Institute of Transport Studies, The University of Sydney, Australia.

Ichoua, S., Gendreau, M. and Potvin, J.-Y. (2003) "Vehicle dispatching with time-dependent travel times". *European journal of operational research*, 144(2): 379-396.

Jain, A.K. and Dubes, R.C. (1988): *Algorithms for clustering data*, Prentice Hall.

Kaufman, D. E. and Smith, R. L. (1993): "Fastest paths in time-dependent networks for intelligent vehicle-highway systems application", *IVHS*, 1: 1–11.

Lorkowski, S., Mieth, P., Thiessenhusen, K.-U., Chauham, D., Passfeld, B. and Schäfer, R.-P. (2004) "Towards Area-Wide Traffic Monitoring-Applications Derived From Probe Vehicle Data", In Kumares, C., Sinha. T.F, Fwa, Ruey, L., Cheu and Der-Horng, Lee (Eds.): *Proceedings of the 8th International Conference on Applications of Advanced Technologies in Transportation Engineering: ASCE*, 389–394.

Lorkowski, S., Mieth, P. and Schäfer, R.-P. (2005) *New ITS applications for metropolitan areas based on Floating Car Data*, Institute of Transport Research, German Aerospace Center (Berlin).

MacQueen, J. (1967). "Some methods for classification and analysis of multivariate observations", In: Le Cam, L. M. and Neyman, J. (Eds.): *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, 1: 281-287.

Pallottino, S. and Scutella, M. (1997) "Shortest Path Algorithms in Transportation models: classical and innovative aspects",University of Pisa, (TR 97-06).

Pinkofsky, L. (2006) Typisierung von Ganglinien der Verkehrsstärke und ihre Eignung zur Modellierung der Verkehrsnachfrage, *PhD. Thesis,* Technische Universität Braunschweig.

Tan, P.-N., Steinbach, M. and Kumar, V. (2006) *Introduction to data mining*, Pearson/Addison Wesley.

Van Woensel, T., Kerbache, L., Peremans, H. and Vandaele N. (2008) "Vehicle routing with dynamic travel times: A queueing approach", *European Journal of Operations Research*, 286: 990-1007.

Windt, K. and Hülsmann, M. (2007) "Changing Paradigms in Logistics – Understanding the Shift from Conventional Control to Autonomous Cooperation and Control", In Hülsmann, M. and Windt, K. (Eds.) *Understanding Autonomous Cooperation and Control in Logistics*, 1-72.

Witten, I. H. and Frank, E. (2005) *Data Mining. Practical machine learning tools and techniques,* Morgan Kaufmann, San Francisco.