Dynamic Travel Time Provision for Road Networks

Dieter Pfoser
Sotiris Brakatsoulas
RA Computer Technology Institute
University Campus Patras
26500 Rion, Greece
+30 210 6930 700
{pfoser|sbrakats}@cti.gr

Petra Brosch
Martina Umlauft
Institute of Software Technology and Interactive Systems
Vienna University of Technology
Favoritenstraße 9-11/188-3
1040 Vienna, Austria
+43 (1) 58801-18804
{brosch|umlauft}@big.tuwien.ac.at

Nektaria Tryfona
Giorgos Tsironis
Talent SA
Karytsi Square 4A
10561 Athens, Greece
+30 210 3217720
{tryfona|yiorgos}@talent.gr

ABSTRACT
The application domain of intelligent transportation is plagued by a shortage of data sources that adequately assess traffic situations. Typically, to provide routing and navigation solutions map attributes in the form of static weights as derived from road categories and speed limits used for road networks. With the advent of Floating Car Data (FCD) and specifically the GPS-based tracking data component, a means was found to derive accurate and up-to-date travel times, i.e., qualitative traffic information. FCD is a by-product in fleet management applications and given a minimum number and uniform distribution of vehicles, this data can be used for accurate traffic assessment and also prediction. This work showcases a system that facilitates the collection of FCD, produces dynamic travel time information, and provides value-added services based on the dynamic travel times. The essential components that will be discussed are a Web-services-based data collection approach, sophisticated map-matching algorithms, a data management architecture and an online visualization platform.

Categories and Subject Descriptors
H.2.8 [Database Applications] - Spatial databases and GIS

General Terms
Algorithms

Keywords
Spatiotemporal databases, floating car data, probe vehicle data, FCD, PVD, traffic management, road networks

1. INTRODUCTION
With the availability of cheap positioning technology and the penetration of asset tracking applications such as fleet management applications, vehicle tracking data, as a component of floating car data (FCD)[1], becomes an important tool for traffic assessment and prediction [13]. Floating car data (FCD) refers to using data generated by one vehicle as a sample to assess to overall traffic conditions ("cork swimming in the river"). Having large amounts of vehicles collecting such data for a given spatial area such as a city (e.g., taxis, public transport, utility vehicles, and private vehicles) will create an accurate picture of the traffic condition in time and space [12]. The whole motivation for this work and the related research project TRACK&TRADE it presents is derived from this fact.

This work presents the technology needed (i) to collect FCD from vehicle fleets, (ii) to derive road-network-related travel times (map-matching), (iii) to provide efficient data manipulation means and (iv) to provide services that utilize these respective datasets. A critical aspect when using FCD to assess traffic is the amount of data that needs to be available and consequently the number of vehicle collecting the data. *Increasing the vehicle penetration* is thus of critical importance and this work provides a flexible, *Web-services-based architecture to simplify the connection of new FCD sources*. *Map-matching* the tracking data produces travel time data related to a specific edge and allows us to derive travel time profiles for a road network. Figure 1 gives example travel times in minutes for major roads in the metropolitan area of Athens, Greece. To derive such *travel time profiles from collected travel times* (dynamic travel times), efficient data manipulation methods are needed. Here a simple data warehouse is used to aggregate collected travel times and to derive dynamic travel times for the road network (link-based speed types).

The approach most closely related to this work is from Dash [6]. The company markets an off-board navigation device that retrieves online traffic information based on the submitted FCD from all Dash devices. In that it closely resembles the TRACK&TRADE approach of collected FCD from a vehicle fleet. However, currently little is known about the vehicle penetration of the service and thus the quality of the provided information. Other examples of related applications include Inrix [9], whose data is also used in MS Live Maps [10] and in Google Maps [8] to assess traffic situations. Cityrouter [7], a prototype installation, provides routing application based on FCD data

[1] A synonym for FCD is Probe Vehicle Data (PVD).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ACM GIS '08, November 5-7, 2008. Irvine, CA, USA
(c) 2008 ACM ISBN 978-1-60558-323-5/08/11...$5.00.

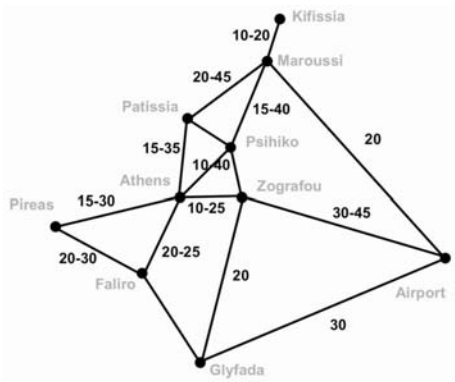

Figure 1: Travel time fluctuation in road network

collected from taxi fleets for various cities in Germany and Austria. A basic problem with all these applications is the limited vehicle penetration and insufficient data coverage.

2. SYSTEM OVERVIEW

Our overall system design addresses the tasks of (i) FCD data collection, (ii) data management and (iii) and service provision. Figure 2 visualizes the overall system.

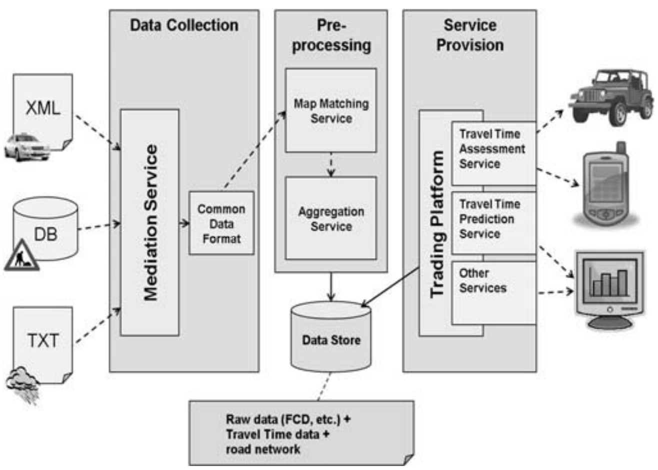

Figure 2: Architecture overview

Data collection focuses on connecting as many FCD data sources to the data mart as possible. In that various Web-based technologies have been developed to simplify the overall effort of connecting a new source. *Data management* includes components for map-matching the tracking data and for data management of (i) FCD data streams and (ii) resulting travel time data. Finally, in *Service Provision* value-added services will be provided that utilize the available dynamic travel time data.

The following sections discuss each system component in greater detail.

3. DATA COLLECTION

For valid and significant dynamic travel time values, a good coverage of the traffic data of the underlying road network is essential. This leads to the need for collecting and integrating various traffic data from different sources.

Floating car data is collected mostly by GPS-devices installed in vehicles and measured as they move. The vehicles send their position and status periodically with a typical sampling rate ranging from 30s to 1min to a central site, typically for disposition purposes. In our running FCD collection system data is collected from taxi fleets in Berlin and Vienna, which use a dispatch system manufactured by Austrosoft [1]. The Berlin/Vienna data is provided as XML and contains a unique ID, timestamp, position at start of measurement, time elapsed, position at end of measurement, and a status code. For Athens, FCD from about 100 vehicles (delivery trucks) is provided in comma separated value format and contains an ID, timestamp, position, and speed. Note that speed is provided as measured by the GPS-device which is in contrast to the Berlin/Vienna case where the speed must be computed from two position samples.

Data collection involves collecting data from these various sources in their different formats and transforming the input data into a common format by a mediation service. For this purpose, various data sources have been analyzed and a respective *TRACK&TRADE data model* has been derived.

In our data collection approach, data can either be pushed or polled. For the push approach a user uploads the data to the Upload Web Service, while in the pull approach a Poller service collects the data from a passive data source and uploads it to the Mediation Service. The user is validated and the data is converted to TRACK&TRADE data model by a transformation engine if necessary. After transformation a configurable Interceptor sends the data to the Map Matcher for further processing. To achieve a maximum of flexibility we offer three data collection alternatives: (i) connection of *new data sources*, (ii) a *push aproach* for legacy data sources, where a Web service wrapper is installed at the source, and (iii) a *pull approach* for legacy data sources where the a server collects the data from a passive source.

The transformation of input data is specific to each user sending their proprietary data format and concrete transformation rules have to be specified. While these rules are implemented by hand, the object representation of the underlying (semi-)structured data schema can be generated by the Java Architecture for XML Binding (JAXB) if the data is described in XML.

4. DATA MANAGEMENT

The data management aspect in this work is related to (i) the use of map-matching algorithms to derive travel times from FCD and (ii) to store and manipulate all collected and derived data, and, specifically, to aggregate travel time to provide a meaningful dynamic travel time database.

4.1 Map Matching

Tracking data is obtained by sampling movement using typically GPS. Unfortunately, this data is not precise due to the *measurement error* caused by the limited GPS accuracy, and the *sampling error* caused by the sampling rate, i.e., not knowing where the moving object was in between position samples [11]. A processing step is needed that matches tracking data to the road network. This technique is commonly referred to as *map matching*. Figure 3 gives two examples of measured GPS positions and the possible trajectory the vehicle could have taken.

The algorithm we utilize is based on the global map-matching algorithms of [3], which employ the Fréchet distance measure for curves [1]. For a detailed description of the algorithm, the reader is referred to [3].

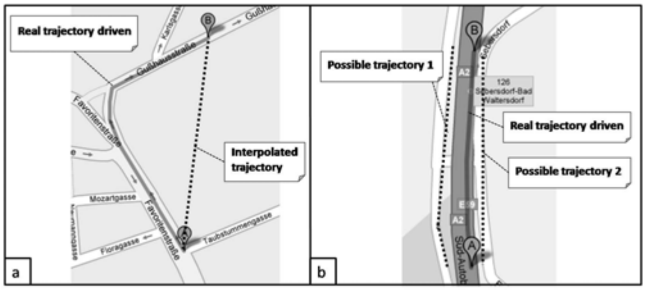

Figure 3: Map-Matching example

Travel times can be derived from FCD by mapping travel times from the trajectory described by the FCD tracking data to the road network. The approach we employ is to *uniformly map the time* recorded between two consecutive GPS position samples to the respective edges of the road network. With all travel times being recorded in relation to the edge and time in question, the following section discusses the most efficient way to store the data and to derive dynamic travel times.

4.2 Data Management

We will adopt a data management approach that will address the data manipulation and storage means of the following items:

- The *FCD* (GPS samples) from the mediation service.
- The *travel times* as they are produced by the map matching process.
- The *travel times in an aggregated form* that enables fast historical travel time queries.

The storage framework described in the following has been implemented by means of an Oracle relational database system.

The collected *FCD* is stored in a single relational table INPUT_FCD. This table is queried at regular intervals (typically 5 minutes) by the map-matching process and if any new data is found, a respective map-matching task is executed. The *travel times* produced by the map-matching process are stored by means of a single table, TRAVEL_TIME, as well.

The collection of historical travel time data provides a strong basis for the derivation of dynamic weights provided that one can establish the causality of travel time with respect to time (time of the day) and space (portion of the road network). The advocated approach will be based on *daily courses of speed*, i.e., the variation of speed and thus travel time for a given road segment with the (i) hour of the day and (ii) day of the week. This resembles a *simple prediction* approach which asserts causality between travel times and uses the observations from the past to predict the travel times of the future.

As travel times are recorded with respect to a specific edge in the road network and a specific time, *travel time aggregation* now refers to *averaging travel times* on a per-edge basis based on the time they were recorded since we consider them related. An example here would be to average all travel times recorded on Mondays from 9:00 – 9:15. Essentially, averaging travel times reduces an error that might exist in a single travel time measurement [4].

To illustrate the power of aggregated travel times consider the example of Figure 4 showing travel times (in minutes) for a road segment. The figure shows aggregated travel times and a static travel time, with the latter being the standard attribute data available with road network datasets. In this context, static travel times are also referred to as *static link-based speed types*. While static travel times remain constant independent of the time (of the day), aggregation produces travel times that vary. It is clear that such *dynamic travel times* are an interesting means to improve the weight database for road networks as they would improve the quality of routing solutions.

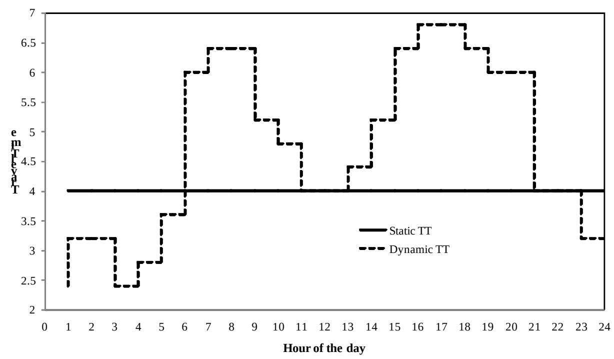

Figure 4: Dynamic vs. static travel time example

To facilitate aggregated travel time computation, a simple data warehouse infrastructure was implemented consisting of one dimension table and one fact table (cf. [12]).

The *fact table* contains travel times on a per-edge basis. The *dimension table* implements the <YEAR, MONTH, MONTH_WEEK, WEEK_DAY, DAY_HOUR, HOUR_15MIN, HOUR_5MIN> time hierarchy. The dimension table stores every possible meaningful combination of values of its columns, e.g., at the first level of granularity (HOUR_5MIN, 5 min interval), there is a row for every year, every month, every week of month, every week day, every hour, every hour-quarter and every 5-minute-interval of the total time period for which the fact table contains measures.

Using the *star schema* in connection with *Bitmap indices* on the time dimension columns and on the foreign key column of the fact table, we can achieve smaller query response times when compared to querying directly the TRAVEL_TIME table.

4.3 Collected Data

The geographic areas for which travel time data has been derived in the TRACK&TRADE project are Vienna, Austria, Berlin, Germany and Athens, Greece.

To illustrate the size of the data that has been collected, Table 1 gives the data that was collected in each city on a typical day (Feb. 4, 2008). It is evident that Vienna has the largest vehicle fleet collecting FCD. On a typical day FCD data corresponding of 500k GPS samples is collected. In contrast, for, both, Vienna and Berlin only 1/10th of the data is available. Feeding this data to the map-matching algorithm produces roughly 50% more travel time data, since the time between two GPS position samples is mapped to more than one edge in the road network. Overall, during a 1.5 year period up to 220 million FCD samples have been collected for a city.

Table 1: Collected data – daily amount

4/2/2008	FCD	TT
Athens	46365	73523
Berlin	38955	66953
Vienna	481479	718408

In addition, various dynamic travel time datasets have been computed by in each case map-matching up to 12 months of FCD for a city (cf. Table 2). Do note that only travel times for roads of higher categories were computed as the recording of travel times by means of FCD for, e.g., local roads poses significant problems and does not produce a reliable travel time base.

Table 2: Dynamic travel time datasets

City	Size road nw [edges]	Higher cat. roads [edges]	Size TT dataset [tt entries]
Athens	207,000	16k	~5M
Berlin	163,000	21k	~6M
Vienna	75,000	10k	~7M

5. VALUE-ADDED SERVICES

The dynamic travel times form a good basis for the definition of a range of more or less advances services utilizing this database.

Travel Time Maps illustrate the traffic situation in urban areas by indicating current speeds for road links in a map using a simple coloring scheme (green – yellow – red). Figure 5 gives a screenshot of our prototype that shows the travel times that are being collected and the traffic situation they represent.

Figure 5: User interface for travel time visualization

The *Travel Time Assessment* and *Travel Time Prediction services* provide current and estimated travel times for a portion of the road network and a time specified. Such services were introduced as a means to encapsulate access the travel time database.

Having available dynamic travel times they provide a excellent basis for improving the results of *routing solutions*. Here, shortest-path algorithms that typically rely on static travel times supplied by the map provider can now use dynamic travel times as collected by vehicle fleets! Overall, having more precise weights available should considerable improve the quality of such routing solutions.

6. CONCLUSIONS AND FUTURE WORK

The live assessment of traffic conditions is a hot topic for the provision of accurate routing solutions. While several commercial solutions started to emerge (cf. [8],[9],[10]) little has been known behind the actual data collection and data processing that leads to such real time traffic services. This work provides an overview of the ambition and the solutions produced by a European research effort that explores the use of Floating Car Data (Probe Vehicle Data), to assess the overall traffic conditions and to create a valuable resource for related services such as accurate routing and navigation solutions.

ACKNOWLEDGEMENTS

This work was partially supported by the TRACK&TRADE project (www.trackandtrade.org) funded by the European Commission under contract number COOP-CT-2006-032823.

REFERENCES

[1] H. Alt and M. Godau. Computing the Fréchet distance between two polygonal curves. Int. J. Comput. Geom. Appl., 5:75–91, 1995.

[2] Austrosoft Weiss Datenverarbeitung Ges.m.b.H. Company homepage. http://www.austrosoft.net, current as of June 2008.

[3] S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk. On map-matching vehicle tracking data. In Proc. 31st VLDB conf., pages 853–864, 2005.

[4] E. Brockfeld, P. Wagner, B. Passfeld, Bert. Validating travel times calculated on the basis of Taxi Floating Car Data with test drives. In Proc. 14th World Congress on Intelligent Transport Systems, 2007.

[5] P. Brosch. A Service Oriented Approach to Traffic Dependent Navigation Systems, In Proc. IEEE International Conference of Services Computing, Honolulu, USA, 2008

[6] Dash Inc. Dash Express - Traffic powered by the Dash Driver Network. http://blog.dash.net/2008/03/18/dash-express-traffic-powered-by-the-dash-driver-network, current as of June 2008.

[7] DLR. Cityrouter. http://www.cityrouter.net, current as of June 2008.

[8] Google Inc. Google Maps. Web page http://maps.google.com, current as of June 2008.

[9] INRIX Inc. Company homepage. http://www.inrix.com, current as of June 2008.

[10] Microsoft Inc. Live Search Maps. http://maps.live.com, current as of June 2008.

[11] D. Pfoser and C.S. Jensen. Capturing the Uncertainty of Moving-Object Representations. In Proc. 6th SSD conf., pages 111-132, 1999.

[12] D. Pfoser, N. Tryfona, A. Voisard. Dynamic Travel Time Maps - Enabling Efficient Navigation. In Proc. 18th SSDBM conf., pages 369-378, 2006.

[13] R.-P. Schaefer, K.-U. Thiessenhusen, and P. Wagner. A Traffic Information System by Means of Real-time Floating-car Data. In Proc. ITS World Congress, Chicago, USA, 2002.