D2.2 Report of new types of data for transport systems and services

Project:	ROADIDEA 215455
Document Number and Title:	D2.2 Report of new types of data for transport systems and services
Work-Package:	WP2
Deliverable Type:	Report
Contractual Date of Delivery:	30.11.2008
Actual Date of Delivery:	30.11.2008
Author/Responsible(s):	Rene Kelpin (DLR)
Contributors:	Antje Parnitzke (DLR), Auli Keskinen (FORECA), Mikko Tarkiainen, Pekka Leviäkangas (VTT), Igor Grabec (AMANOVA), Marjo Hippi (FMI), Andrea Rossa, Franco Zardini (ARPAV), Jörgen Bogren (KLIMATOR), Djurdica Markovic (METEO) Jörg Dubbert (PÖYRY), Pirkko Saarikivi (FORC), Lulu Hyvätti (FORC),
Approval of this report:	Coordinator, Technical Committee
Summary of this report:	Report summarises the relevant data needs for the transport service ideas resulted from the first Innovation Cycle
Keyword List:	Data collection, data classification, data sources, mediation, data workshop, Clearing House of transport data
Dissemination level:	Public

Co-funded by the European Commission under the 7th Community Framework Programme for Research and Technological Development

Change History

Version	Date	Status	Author (Partner)	Description
0.1	26.09.2008	Draft	Rene Kelpin (DLR)	First draft
0.2	31.10.2008	Draft	Rene Kelpin (DLR)	2. draft version, Introduction added
0.3	07.11.2008	Draft	Rene Kelpin (DLR)	3. draft version, added text from VTT and AMANOVA
0.4	12.11.2008	Draft	Antje Parnitzke (DLR)	4. draft version, added text from DLR, FMI and KLIM
0.5	18.11.2008	Draft	Rene Kelpin (DLR)	5. draft version, added text from PÖYRY and ARPAV
0.6	21.11.2008	Draft	Rene Kelpin (DLR), Antje Parnitzke (DLR)	6. draft version, added text from FORECA, referencing
1.0	28.11.2008	Final	Rene Kelpin (DLR)	1. final version, added Executive summary and Conclusion
1.1	28.11.2008	Final	Rene Kelpin (DLR), Pirkko Saarikivi (FORC)	Final additions and quality check
1.2	26.05.2009	Final	Pirkko Saarikivi (FORC)	Front page with flag

Distribution list

European Commission	Emilio Davila Gonzalez
	Wolfgang Höfs
ROADIDEA partners	E-mail list
	www.roadidea.eu

Executive summary

Since the first innovation seminar, held in Prague in May 2008, fourteen ROADIDEA ideas are listed. Four ideas were decided to be developed further as project pilots, whereas it has to be distinguished between two practical pilots, which are ready for implementing and which are expected to be up and running within the project run time, and two theoretical pilots. The latter need further theoretical background consideration, before they might be decided as piloting candidates in a later project phase. In this report all pilots and ideas are described as detailed as reasonable, and due to the objective of this report, with special respect to the data needs.

Hence, this report summarizes the relevant data needs for the mentioned pilots and ideas - and synchronizes the data needs with the data availabilities investigated in the task 2.1 of the data collection and data source investigation work package.

It can be pointed out generally, that almost all needed data for implementing and feeding project pilots and ideas are available. But, given data coverage and according availability varies strongly as a function of the kind of data, of the structure of urban development, and of the necessity to face issues, that might be addressed with specific data. That means that productive and problem-oriented transport and information systems are commonly provided with necessary data only, where it is really needed. Respective data can only be valid for the underlying spatial context. For instance, very good traffic data coverage for the Gothenburg city area can neither describe traffic conditions in similar Swedish or European cities nor the traffic conditions in the surrounding rural regions.

All over Europe the availability of road and grid weather sources is much better than the availability of reliable road traffic observations. Problems to be solved with tools to be developed in the ROADIDEA project vary from the south to the north of Europe very strongly. With partners from the weather monitoring and forecasting sector a very good experience in terms of acquisition and utilisation of weather data can be applied immediately.

In order to implement some ideas, specific piloting areas, countries or cross-border regions have to be identified, where needed data are available as necessary. For cross-border regions possibilities to utilize overlapping data sources from the other region have to be elaborated. European regions with similar problems and similar data availability have to be found and analysed concerning the possibility to transfer know-how from one region to the other. These issues are to be considered during the WP2 task 2.3 on integration of data.

Utilising so far available data does not make it possible to implement one specific pilot or idea in all European countries.

It has to be concluded, that the data source investigation must be continued. The table in the results chapter of this report gives an overview of data needs and data availabilities. Although only CAN-bus and mobile phone tracking data are mentioned to be further investigated, a general necessity for more data can be derived from the table. It should be tried to find data source descriptions for the same data classes in all participating ROADIDEA countries. Besides that, in order to use similar data sets for cross-border ap-

proaches, data sets have to be made comparable. This can be done by converting relevant data into a common data format, which remains to be decided. Within the ROADIDEA project, this can be done for only a few data examples.

Especially on the field of CAN-bus and mobile phone tracking data the given data availability is very weak or not existing. Here a further investigation of available data sources is very urgently needed.

Table of contents

1. Introduction

The data source investigation of ROADIDEA WP2 task 2.1 revealed a quite comprehensive overview of existing data in the European transportation system. With respect to the project objectives a decisive subset of relevant data was described. Data samples and additional materials such as codebooks were partly uploaded to the data collection platform. A reasonable data classification was developed and all data sets were assigned to this data classification. The results of this task have been described in the report D2.1 [ROADIDEA 2008a].

But, as long as the services and pilots to be implemented in the ROADIDEA project were unspecified, nobody was able to answer one crucial question: Have we collected the right data? In other words: Is it possible to provide the necessary data for those piloted services that where decided to be developed further. This dilemma had been affecting the data source investigation and was discussed several times.

With the first innovation seminar the ROADIDEA idea and pilot discussion process began. At the innovation seminar in Prague in May 2008 the ROADIDEA consortium supplemented by some experts from relevant sectors of the public and private transportation community (such as mobile phone companies, road authorities, logistics service provider etc) were creating service ideas in an at the first instance free brainstorming process. Hence, the already heterogeneous ROADIDEA consortium has been enriched with further expertises, view points, innovation requirements and business approaches which were very fruitful for the idea generation process. After the brainstorming and idea collecting all ideas were WACSVOD analysed and basketed accordingly, see D5.2 Results of the first innovations seminar [ROADIDEA 2008c]. At the end of the first innovation seminar a couple of ideas were given. Idea care taker and idea teams were defined.

For the further development of the data providing platform and the initial ROADIDEA database, specification of the data needs for the given ideas is needed urgently. It has to be decided whether relevant data can be provided for the ideas directly or which data has to be searched, ascertained, generated or requested.

That was the main objective of the task 2.2 of the work package 2. It was to be accompanied by the generation of new data by ROADIDEA work package achievements.

The ideas and its data needs are described in the chapter 2 of this report. This chapter also includes some examples for achievements which have been obtained in the ROADIDEA work packages. According the ROADIDEA description of work some research activities were performed in the preceding project phase. This work aimed at the provision of needed data by self-generated new or so far untapped data sources. The data relevant results of these activities are described in chapter 2.16.

Afterwards some ideas were needed to be decided as candidates for piloting in the ROADIDEA project run time. This decision has been made in several project and subgroup meetings. Four ideas got the pilot status while it has to be distinguished between two practical pilots

- Pulp Friction and
- Gothenburg Pilot

and two theoretical pilots

- Hamburg Port Pilot and the
- Italian Fog Pilot.

The pilots and its data needs are described in the result chapter 3 of this report.

Besides that, the mentioned data classification was considered with regard to empty or sparsely filled data classes defined in D2.1 [ROADIDEA 2008a]. Although the data need is determined by the ideas, it might be useful to know whether empty data classes are not needed for the ROADIDEA data classification or if it is necessary to emphasize the data source investigation for those data classes as well. According to the results of this consideration the data classification may be adapted.

The work and achievements described in this report are to be considered as a base for the further development of ideas, pilots, and, hence, ROADIDEA innovations. With the second innovation seminar in May 2009 all ideas have to stand further tests and critical considerations; business models for all ideas have to be designed and assessed according their reliability and viability. New business ideas will be born. They also have to be described and analysed if nothing else, according their data needs. Thus, the investigation and description of data needs for new ideas, new pilots and further ROADIDEA innovative business cases have to be considered as an ongoing process during the entire project runtime. This might also cause new data source identification needs.

2. Methodology

2.1 Finding new ideas – The ROADIDEA innovation process

Methodology

Innovations in ROADIDEA will be created using well-tested brainstorming methods such as Charrette and Futures Workshop described by the WFUNA Millennium Project covering both seminars, the time in-between for running pilots and the processing of the ideas for the Road Map after the seminars, see D5.1 Innovation Plan [ROADIDEA 2008b]. Charrette is an intensive face-to-face process carefully designed to bring people from various disciplines and expertise to create new solutions to problems within a short period of time. The first round (the innovation seminar) will produce the ideas to the focus groups (work packages) that will take on the ideas and develop their work accordingly. The second round will further develop the original ideas and create new ones for further consideration.

Objectives

The WP5 Innovation procedures and management implements the innovation procedures in a systematic way in two major cycles and continuously in between. The main goal is to produce radical ideas to develop safer, more secure, efficient and environment-friendly ICT-based transport solutions and services.

Results of the 1st innovation seminar, May 2008

The best ideas were sought for during the seminars using first individual brainstorming, grouping of ideas, evaluating the ideas and lastly "basketing" the ideas according to their characteristics. The short-listed ideas were also screened against questions concerning weak signals, actors, collaborators, structures, values, obstacles and drivers, called WACSVOD analysis. To ensure the full exploitation during the project life cycle of the ideas created, a 1-3 person team (IDEA Team) was appointed to take care of each idea.

Summarising, the results include 34 fully studied ideas, of which 19 were short-listed after the evaluation. From these, 12 ideas were chosen for further work with dedicated IDEA teams, consisting of 6 piloting ideas, 3 modelling ideas and 3 general development ideas.

Many more additional ideas were presented that either complemented the short-listed ideas or did not get enough evaluation points. However, all ideas were basketed and stored to the ROADIDEA web site for further use.

The seminar survey results (50 % responded from 36 participants) indicate that the seminar was well received. The majority of participants were of the opinion that no truly radical ideas were created, however, they could name the most radical ideas from their point of view. The overall problem was that the concept "radical" was varyingly understood and no consensus prevailed. Also, what was radical to one person was not so to another. However, the main target of ROADIDEA is to study the overall innovation potential of European transport services, and especially analyse the barriers for their further exploitation. In this respect the seminar provided ample material for studies and for the first pilot product development phase.

In the following all ideas taken into further consideration (WACSVOD Analysis) during the innovation seminar in Prague are listed regardless the baskets they were put in.

2.2 Pulp Friction

2.2.1 Description

The idea

Finnish road weather observation network consists of about 475 stations (situation in November 2008). 90 of those stations are new types of optical sensors which give an estimation of prevailing friction on the surface, too. The road weather station network provides good monitoring system for road maintenance personnel as well as meteorologists. There is several road weather products developed which help monitoring the road weather.

In the road weather warning service, which is operated by FMI and Finnish Road Administration together, the road conditions are divided into three categories: normal, bad and very bad road conditions. There is a link between friction and road conditions: friction above 0.3 means normal road conditions, friction 0.15…0.3 bad road conditions and friction below 0.15 is linked to very bad road conditions (by Yrjö Pilli-Sihvola, Finnish Road Administration, see Figure 1). Estimated friction data from road weather stations is not included to the weather warning service system so far, but in the Pulp Friction pilot the usability will be studied.

The Pulp Friction pilot is mainly considering of road weather and friction. FMI will do two separate studies. On the first research the relationship between observed friction value and the classified road condition will be studied. In this pilot FMI is evaluating the road classification criteria against the "subjective" value for road conditions given by the personnel in the Road Monitoring Center of Finnish Road Administration (and also against other RWIS data). If the results look good we could have a more automated analysis of road conditions based on friction measurements.

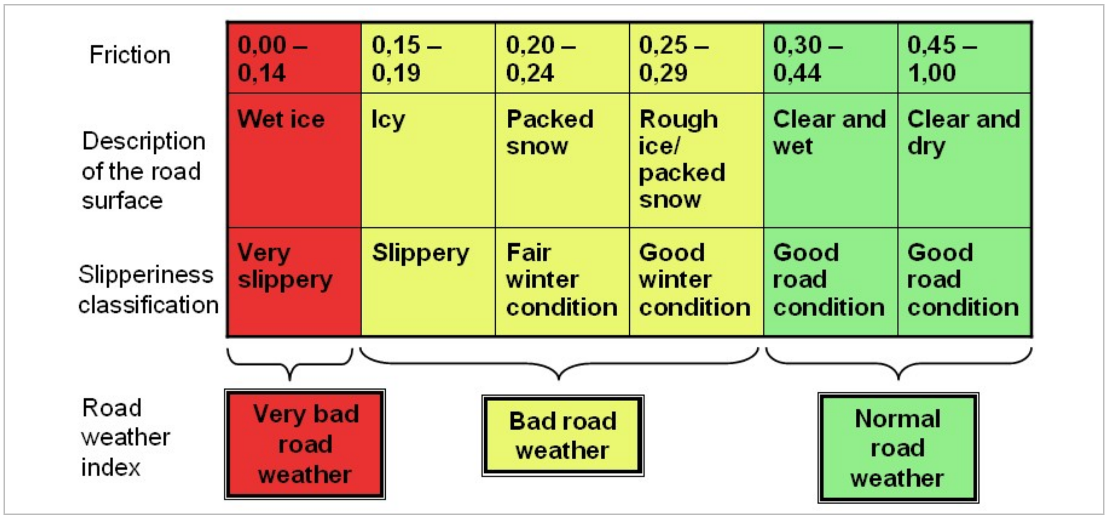


Figure 1: Correlation between friction, description of the road surface, slipperiness classification, and road weather index.

The main idea of the other research is to study the correlation between road condition and road weather observation and create a friction model based on that information. The friction model could do an estimation of prevailing friction based on prevailing and past road weather observation and weather. Also, the friction model can predict the coming friction based on the forecasted road weather model data. The friction model will probably be part of FMI's existing road weather model.

The needed input data is road weather observations (including friction observations) and road weather models inputs and outputs. FMI's road weather model is an operative product already, so it won't need any new inputs. Road weather classification indexes for last winter are compared to the prevailing friction value and the connection between those values will be studies. The road weather classification index data is stored in FMI for previous winters.

The outcomes of these two separates studies will be aimed for the road maintenance personnel and meteorologist to help monitoring prevailing road weather and to estimate the prevailing and forecasted road weather. Also, some products based on friction or processed friction can be developed for drivers.

<u>Technical approach</u>

This pilot will provide two different kinds of services; an (semi)automatic road weather classification system based on the road weather observations (e.g. fiction) and a friction model which models friction based on road weather observations and forecasts. Both services will give information about road weather for meteorologists and road maintenance personnel as well as generally for all drivers.

Friction data is more or less new input and output when developing road weather tools. Friction values from road weather stations are taken into account at first. Later it is probably use friction values also from different sources (e.g. from cars) if the values are comparable.

Geographical scope

Results of the Pulp Friction pilot may be used anywhere despite of the geographical location. The friction model will be developed by using Finnish data. However, if the input data is available, the model can be run for any certain place. Road weather classification study follows Finnish classification settings made by Finnish Road Administration.

2.2.2 Data Needs

Pilot Pulp Friction needs **road weather observations** as an input data. Also, **road weather classification indexes for previous winters** will be in use when studying the relationship between friction and road weather. Friction model will be developed by studying the correlation and relationship between friction values and road weather. All needed data is stored at FMI's database. Road weather observations are also available from Finnish Road Administration.

2.3 Cross Border Weather Alerts

2.3.1 Description

For multi-modal and long distance travel, the accessibility to recent and up-to-date warning information is essential. The accessibility is just one important requirement. The form of the service, i.e. language and other user features of the user interface are also important (see Chapter 2.13). Weather warnings form one of the most useful set of information. There are recent cases of severe traffic accidents, e.g. in the Czech Republic, resulting from the lack of good warning information to the truck drivers in their own language (personal communication with ROADIDEA Advisory Committee member Michal Najman /Meteopress during the first ROADIDEA Innovations Seminar).

Efficient access to warning information and its further processing must be secured in all European countries. A centralised service has been formed already among the National Meteorological Services (www.meteoalarm.eu). A similar, more extended service with all other kinds of traffic warning information would be useful to all cross-border traffic. This would require administering the access to warning sources and integration of these various data sources to a well-maintained operational traffic warning information server.

2.3.2 Data Needs

Data needs for the aforementioned centralised warning information service cover all warnings, which are valid for transport users, whether car drivers, passengers in public transport or professional transport users. **Weather warnings** are already well organised and the first centralised service exists. For **road weather observations and warnings** the situation is already more complex and needs further analysis and planning to achieve a more centralised database, which would be openly accessible for service providers.

First intention of this idea was to concentrate just on weather information and weather alerts. However, there is plenty of other **additional warning information** valid for road users containing large variety of data from different sources. Those include i.e. information on **traffic jams**, **road construction works**, **damaged and/or closed road sections**, moose or other **animals on road**, just to name a few. No centralised system exists to get this kind of warning information easily to the car driver, while driving or during the pre-trip phase. There are many sources of information: public authorities and data providers, i.e. fire brigade, police, construction departments of responsible government authorities. Relevant data should be collected from all these sources and integrated in one central server for efficient production of transport warning services.

2.4 Mobile Phones as Sensors

2.4.1 Description

This idea has close contacts to the Gothenburg pilot, see Chapter 3.1.2. Not only the car itself can act as a sensor while driving, but also the car drivers with their mobile phones can act as sources as well as receivers of information to and from various transport services. This idea connects also to the previous Chapter 2.3 Cross boarder weather alerts. Car position tracking could be done if the position information was available from teleoperators regardless who the operator is. This case is comparable to that of CAN-bus data in the Gothenburg pilot. At the moment car manufacturers do not give their specific data to others outside the circle of their own car users.

Additionally, drivers (whether private of professional) could systematically send relevant warning information (e.g. road structure damaged, moose on road) if the collection and further processing of the information would be well planned and managed, and as easy as possible for the driver. Very popular, old-fashioned way is to call a local radio station. This kind of activity should be conveyed to a more organised and (semi)automated way to send information from the car, and back to the car.

Recently new transport services have been emerged, based on the active data gathering of users. Good example is the service for bicycle users in the Netherlands (see the demo on the ROADIDEA website www.roadidea.eu).

2.4.2 Data Needs

If car drivers are given the opportunity to send their **position tracking or other driver observed information** to transport services, the main challenge here is to organise the collection of data, to take care of the reliability and quality of information aspects, and the further processing of information to best possible end-user products and services. Utilising position tracking data from mobile phones is technically possible, and many new transport services could be innovated if these data were available. But unfortunately, the use of position data has many obstacles starting from the legal constraints dealing with data security and privacy issues.

2.5 My Route Mobile Pocket Guide

2.5.1 Description

The idea:

My Route Mobile Pocket Guide is a system for providing travel information and updates over a mobile network. It gives the traveller/driver comprehensive real time traffic information needed for well-informed travel decisions (pre-trip information) as well as information during the journey (on-trip).

There are many ways to get traffic information (Internet, Teletext, TV, radio). But the information the traveller can get from those services usually are:

- either already out of date by the time driver gets them or can actually use them,
- it's passive information that's often irrelevant to traveller's personal route
- most of these services only cover parts of the highway network,
- delays are given in kilometres instead of minutes, thus the driver / traveller still has no idea how long the delay would be

Combining real-time information about events (incidents, accidents, construction sites, etc), real-time information about traffic conditions, information about traffic forecast (potential road traffic conditions), travel time information, speed limit information, weather information and available parking and rest areas, it would support the traveller in finding the best, the most efficient and safest way to reach wished destination. It would influence the modal choice, route selection and the time of departure as well as contribute to the reduction of incidents and emissions from vehicles.

The first part of the service is its **pre-trip part**, which enables user to choose the best route according to the driver's criteria such as: quickest, shortest, environmentally most friendly, avoiding toll roads and congestion charge areas, by required arrival time etc. It also inform the driver about traffic signs, prohibited manoeuvres, vehicle restrictions, points of interest, tourist information, speed camera data and much more on the chosen route.

The second part of the service is its **on-trip part**, which enables user to stay informed and in control when he / she's on the road. Whenever some significant situation occurs on the chosen rout, by using voice messaging alert system, My Route Mobile Pocket Guide service will inform the driver of a new situation and suggest the smartest possible action in order to achieve required arrival time and budget.

My Route Mobile Pocket Guide is suitable as much for people (private drivers) as for businesses. It can be used in the field of transport, urban planning end economics, by policymakers and consultants dealing with mobility issues, by authorities dealing with traffic flow, it could supply traffic information to (local) governments.

Technical approach

Implementation of My Route Mobile Pocket Guide requires at least further elements:

- Easily accessible **application with road map** of assigned area (town, region, country, cross - border)
- **Central database system** (traffic and weather; current and forecasted) with **update function** and the **search function** providing information along a specified route and connected to an **SMS alert centre**
- The **GPS** receiver with Bluetooth technology to ensure the best possible reception wherever the driver / traveller goes
- **Subscription model** to information flow

Geographical scope:

My Route Mobile Pocket Guide is suitable as well for long distance journeys as for daily commuting, and applicable in local, regional and national as well as European level. The level itself depends directly on the Central database system scope.

2.5.2 Data Needs

Full realization of My Route Mobil Pocket Guide idea mostly depends on combination of current and forecasted traffic and weather data. There is a need for both, temporal and spatial data coverage. Raw known data can not be used directly – data fusion or integration is necessary as well as value adding aggregation.

Data for pre-trip part of the service are mostly spatial. Database should include:

- Road selection: road, highway, roads with no toll, roads with no congestion charges;
- Route selection: shortest, quickest, cheapest, environmentally most friendly, safest;
- Resting places selection: position, purpose, size, working time
- Travel time information
- Weather and traffic forecast data

Real-time data or on-trip data are mostly temporally:

- Real-time weather data
- Weather nowcasting data
- Traffic real-time data
- Traffic nowcasting data including “breaking news data”

Pre-trip data should be modified and corrected during the trip by fresh real-time data or "breaking news data" – information about incident events on the route like car crash caused jam or similar.

Such service requires permanent communication between personal communication device (cellular phone for example) and Central database which could be local or regional. Communication frequency should be in minutes or Central base should have information about all users in its region and send fresh information including "breaking news" when they occur.

Some real-time data could be "two-way" data, means that device collects some data from vehicle sensor and sends to Central database. But, such activity has some legal problems which should be solved in advance.

2.6 My Travel Toilet-TomTom

2.6.1 Description

The idea

My Travel Toilet Tom-Tom service is offered to all drivers and travellers. It gives a driver/traveller the possibility to find information on availability of toilets along a planned route, with particular impact on places for disabled persons and mothers with small children. It could also give the driver/traveller possibility to check the other resting possibilities at chosen place. Although My Travel Toilet Tom-Tom idea could seem a little bit frivolous, the service will have impact on the state of mind of the driver, and thus reduce driver's nervousness, inattentiveness, uneasiness, and so reduce the risk of incidents. My Travel Toilet Tom-Tom could be integrated in some other Tom-Tom applications regarding travel planning.

Technical approach

Tom-Tom mobility solutions offer their users possibility to deal with many mobility issues at lower cost and without need for a complex implementation to carry out.

Implementation of My Travel toilet Tom-Tom idea requires at least further elements:

- Easily accessible application with **road map** of assigned area (town, region, country... cross - border)
- **Database** with all available toilets / resting facilities along the assigned area **with update** function and the **search function** providing information on available toilets / resting facilities along a specified route
- The **GPS** receiver with Bluetooth technology to eliminate the need for connecting cables and a high performance GPS antenna to ensure the best possible reception wherever the driver / traveller goes
- **Subscription model** to information flow

Geographical scope

My Travel Toilet Tom-Tom service's assignment is mostly local and regional. But, as a part of some other Tom-Tom services, it can be used even more widely.

2.6.2 Data Needs

There is a need for spatial data coverage for the application which would be used for potential My Travel Toilet Tom-Tom service. Database should consist of **facilities** by further categorization:

- by position
- by purpose
- by size
- by working time

Raw known data can be used directly, but has to be updated regularly.

2.7 In-Vehicle Information

2.7.1 Description

The idea and existing data in Vehicles

Modern society is equipped with lots of sensors. Sensors are embedded in vehicles, in road infrastructure, in aviation and satellites.

A vehicle contains a lot of existing information used by the different units inside the vehicle. This information is in most cases only used for its designed purpose and not so often used to gain new information outside the vehicle. The vehicle contains information for example about: temperature, road condition, speed, air pressure, air quality etc.

About 90% of the new cars sold in Sweden 2008 have the safety system ESP, Electronic Stability Programme.

The SRIS project

Semcon has previous used existing signal in a project regarding slipperiness (SRIS) with floating car data together data from infrastructure such as road weather information. Semcon has also provided data from sources not used before to this purpose such as airport weather and traffic information. All these different datasets of historical data have been prepared in WP2, ready to be used.

The result from SRIS (Slippery Road Information System), is that the information from the vehicles can be used as real information regarding the road condition and by combining with road weather stations, a higher quality of the information can be available. The result also showed, that in most cases both the vehicles and the road weather station indicated slipperiness, but where were also events of slipperiness reported by the vehicles without any warnings from the infrastructure.

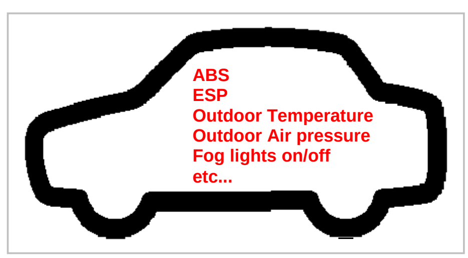
The challenge is to combine data from today existing separate system to gain new and vital information.

Technical approach

The approach Semcon has used is to only use data from existing sensors in a new way. To transfer these existing signals, equipment is in some cases needed. But the future idea is to not need specific equipment in each vehicle to receive data.

Background SRIS hardware

Semcon has a wide area of knowledge regarding vehicles. This knowledge has been used to develop own hardware for CAN-bus communication



and to send the information with telematics to a central database.

To test the existing sensors and safety system in vehicles, first test-drive was made with a Volvo V70 from 2001. Test drive took place in north of Sweden in the winter. Turing the test-drive, a person with a laptop connected to the OBD connection (Onboard Diagnostic) and CANalyzer was used to monitor the upcoming signals. The result showed a good correlation between the real, provoked events and the observed signals. And this, with a relative old car!

The winter 2007/2008, the SRIS field test took place with 100 cars reporting events of slipperiness to a central database. The main part of the cars was located in Gothenburg area but also a number of cars in Stockholm used the SRIS system. In the field-test, it was both Volvo V70 (2007) and SAAB 9-5 (2007) used. The hardware was designed to be connected to the standardised On-Board Diagnostics (OBD) connector and transmit with a telematic unit of the same sort Volvo has used in previous models.

<u>Geographical scope</u>

The geographic area is for vehicle data only depends where the vehicles are driving. By increasing the number of vehicles transmitting data, a wider area with better coverage can be established. In the future, we see the possibility to increase the area to a larger area if good covers can be established by vehicle transmitting floating car data. The SRIS system field-test 2008 was based on 90 cars in Gothenburg and 10 in Stockholm.

2.7.2 Data Needs

For the Gothenburg pilot, following datasets have been prepared:

- Can bus vehicle data
  - SRIS data, data regarding slipperiness and position information
- Road traffic counting
  - 15 places in Gothenburg have been selected, if necessary, it’s possible to get access to information from more locations.
- Road weather data acquisition
  - Access to approximate 700 road weather stations
- Airport weather observations
  - Historical data from 8 selected airports in west Sweden, access to live weather from the major big airports in the world.

For all these datasets, Semcon has historical data stored, ready to be used. But it should also be possible in the future to establish a link to each data source for real time information.

2.8 EUroadmap

2.8.1 Description

The idea

The pilot is aimed at the very heterogeneous field of given road weather data in Europe. The availability of road weather data ranges from well-described, free available and intensively applied data sets in Scandinavia or Germany to non existent data sets in some east-European countries.

While for Scandinavian countries as Finland and Sweden a reliable knowledge of the road conditions derived from road weather information is very important not only in the winter season, the necessity for daily road weather information in central and south European countries seems to be negligible. However, not only northern European countries may increase road safety by using accurate and actual road weather information. Also for instance strong wind regions as the Croatian east-Adriatic coast or alpine regions in the Alp's countries could use road weather information for forecasting road conditions more intensively. There are multiple areas in Europe that have exceptional weather conditions that need to be taken into account by road users.

First of all the pilot needs a thorough European wide investigation of available road weather data sets, their formats, and utilizations. Measurement devices and communication technologies have to be considered as well as the quality and reliability of measured weather data. For this purpose a contribution of all ROADIDEA project members is needed. The ROADIDEA WP2 data source investigation methodology can be adapted and used for this first step of the pilot implementation. By and large the necessary data is there, at least for problematic regions but the problem is with integration and distribution. For example, the TEMPO-programme has reported a substantial increase in coverage of traveller information services (traffic situation via internet, VMS, RDS-TMC) in Europe during the recent years. [TIS, 2007]

In order to give a European wide overview of existing road weather data a general format has to be designed, which allows storing all data sets in one EUROADMAP data base. This overview is to be done on a web site, which accesses the mentioned data base and which could be use as a trading platform as well.

---

[1] Traffic Information Services on the Trans-European Road Network. Success stories, current practice and perspectives. Traveller Information Services (TIS) Expert Group. May 16, 2007, version 5.3.

In order to carry this out, a business case need to be pointed out so that wide-coverage distributors, such as BBC and similar global information service enterprises could work as integrators. This is already today done for weather information services and a similar business case needs to be addressed.

A similar approach was followed in the EU research project Track&Trade, in which vehicle probe data from taxi fleets (FCD) was investigated, visualized on a web site in traffic state maps and traded directly. A business model has been developed which included different aggregation levels of the raw data and different billing models accordingly.

The EUROADMAP pilot might follow the Track&Trade example directly.

A special challenge in case of EUROADMAP is the merging of overlapping road weather data sets with inconsistent attributes and values, i.e. coverage overlappings of two neighbour weather stations with different formats at inner-European borders.

For the project run time of ROADIDEA a visualization of available road weather data sets all over Europe would be reasonable. This includes the data set investigation, a format consideration, a data base implementation as well as the visualization on a web site.

The web site could link to existing web applications (DESTIA) based on the visualized road weather data sets.

Big information services companies should be contacted and the business case should be inquired and made explicit. A demonstration should be interesting to them.

Technical approach

The following technical components are needed for the pilot:

- data base
- web server with data base and map interfaces

Only standardized communication channels are to be used. Data base accesses using SQL commands. The web site communication (only visualization) is to be established via the HTTP protocol.

It is recommended to use the ROADIDEA platform (includes at least a data base as well as a web server) for this pilot.

Geographical scope

The geographical scope could be

- 'bilateral' between selected two countries on specific road stretches
- a corridor–specific, e.g. Trans-Alpine routes, or Adriatic coastal corridor or based on Euro-regional division as in EasyWay project.

2.8.2 Data Needs

During the ROADIDEA data source investigation a couple of **road weather data** as well as **grid weather data** was identified and described. This data is needed from the ROADIDEA platform. As described above a further investigation of available road weather data has to be performed. These data sets are to be stored in the ROADIDEA data base as well and are to be provided for the EUROADMAP pilot, too.

In practice, there needs to be an access to national road authorities' (or their service providers') road weather data bases and integrate this information on digital map information.

Road weather data could be obtained from the national road administrations in Finland, Sweden and Norway as well as from the road administrations in the North-German Federal States. Available DATEX-nodes could be used in order to obtain road weather data referenced in RDS-TMC location codes.

2.8.3 Other necessary steps

For demonstration purpose, the access to road weather data, which is a combination of road data from road authorities and weather data available now globe-wide, would be sufficient.

For real business case, there must be the 'integrator' and 'distributor' of road weather information. For professional services, such as drivers of large transport companies, the companies themselves could act as a client to integrators and distributors who in turn could be SME service providers, targeting their services to dedicated segments of professional users that have willingness to pay for the services.

In consumer market, i.e. the ordinary road users in their passenger cars, the distributors and integrators must be larger entities, precisely as it is today for global weather information provision. This is done for free just because it provides an access to offer these consumers other types of services.

2.9 EYEAR

2.9.1 Description

The idea

The idea is centred on friction data collection and transmission. It is supposed to improve the detection of road friction data by introducing measurements based on floating vehicles. In this sense, EYEAR is a form of an extended floating car data detection technology. The vehicle is used to carry optical friction sensors and brake sensors which detect the degree of the local friction on the road.

a)

Combined with the GPS-position, it would be possible to detect road sections which low friction values and – based on this – to generate a map-based overview of temporarily dangerous road sections and respective warning messages which can be distributed to the road users as a service and to road maintenance staff as an improved basis for taking counter-actions like salt spraying or snow ploughing.

b)

Based on car-to-car communication, a vehicle which detects low friction values could inform the immediately following vehicles and warn the driver of the hazard.

Technical approach

The friction data need to be location referenced (GPS) and transferred from the sensors to the onboard unit of the vehicle via the CAN-bus system of the vehicle.

Approach a)

In order to implement such an application, a fleet of a larger size (5% of all vehicles in the network) need to be equipped with friction sensors, on-board units and data communication. A quick data transmission to a central computer system is needed in order to be able to process the friction data sets from several vehicles in order to generation warning messages to the general road user via information services and road maintenance information to the road authorities. For this, a larger vehicle fleet is necessary.

Approach b)

Using car-to-car information technology, messages about slipperiness could be given from one car to the other by short-range communication. If one car detects a slippery road section, the warning message could be given to the following vehicles. This requires the equipment of cars by respective intelligent on-board computers and a reliable short-range communication.

Geographical scope

The application is relevant in all countries, but with a stronger focus on countries with harder winter weather conditions (e.g. in Northern Europe).

2.9.2 Data Needs

The application needs to detect the following data:

- measured **friction values** from single vehicles
- **vehicle positions**

Approach a)

It is necessary to equip a larger fleet of vehicles in order to obtain a good geographical coverage and to guarantee a high credibility and quality of the database.

Those vehicle fleets should be preferred which often operate in the area in question and which have a digital mobile data communication method on board which allows a low-cost data communication procedure.

The application could be improved by predicted friction values coming from models.

The application generates location-referenced warning messages on slippery roads due to adverse weather conditions (ice, snow, rain). The messages are sent as information content to service providers and road managers.

Approach b):

The data needs are in principle the same. However, the application can already be used by a few vehicles. The idea concentrates on local warnings within a group of vehicles. Data need to be transferred reliably on a short-range communication line.

2.10 Traffic Forecasting Model

2.10.1 Description

The idea

Road traffic is a stochastic dynamic phenomenon whose description has to be performed statistically based upon data provided by traffic monitoring and information about activity of population and environmental conditions. The basic problem is the modelling of traffic flow time series generator and forecasting of future traffic properties under changing environmental conditions. Traffic data generally represent stochastic time series. [Helbing, 1997], [I. Grabec et al., 1997]

Hence, the leading idea is to extract from past data a statistical model of the time series generator and apply it in forecasting. Our goal is to propose a non-parametric statistical method for this purpose and to demonstrate its applicability. In relation to this goal there generally appears a problem with blank spaces in traffic records. To solve this problem available data can be used in modelling of a statistical predictor of missing data which in fact coincides with prediction.

In the next subsections the fundamentals of non-parametric statistical modelling and their validation are presented together with some characteristic examples. More detailed description of non-parametric modelling is described in the Report of WP3, while here the fundamentals of new data generation are mainly pointed out.

Fundamentals of non-parametric modelling

As the basic tool for description of traffic we utilize a time series of rate *Q* of vehicles passing an observation point at time t. Since the traffic is often comprised of several categories of vehicles moving on several lanes, the flow has to be generally described by a multi-component vector **Q** = (q1, q2, …). This vector depends on time t and a point of observation **r**, therefore the traffic flow generally represents a dynamic field **Q**(t, **r**). [Helbing, 1997], [I. Grabec et al., 1997]

We next describe variables influencing the traffic by a vector **V**=(v1, v2,…). Its components represent weather conditions, calendar, properties of population activity, etc. Similarly as **Q** it should also be considered as a dynamic filed: **V**=**V**(t,**r**). The traffic is then treated as a non-autonomous phenomenon and characterised by the generating equation:

$$\mathbf{Q} = \mathbf{G}(\mathbf{V})$$

Equation 1

Since **Q** and **V** are stochastic variables this equation is interpreted as an expression of a statistical mapping relation, while the traffic generating function **G** is considered as a statistical estimator. [I. Grabec et al., 1997]

There are generally two approaches available for specification of statistical estimators: a parametric and a non-parametric one. [I. Grabec et al., 1997] A disadvantage of parametric approach is that complex phenomena, such as traffic, hardly permit for a simple analytical modelling of the estimator **G**. In opposition to this a non-parametric approach renders possible a simple and quite general formulation of a statistical model. It is called *general regression* and is based on experimental data only. Here we apply general regression to modelling of traffic under variable environmental conditions. [I. Grabec et al., 2008-04], [I. Grabec et al., 2008-05] An advantage is that this modelling can be performed completely automatically by modern data acquisition systems. [I. Grabec et al., 1997], [I. Grabec et al., 2008-04], [I. Grabec et al., 2008-05]

The non-parametric approach starts with an estimation of the probability distribution of joint variables **Q** and **V**. For this purpose a kernel function estimator is generally applicable. [I. Grabec et al., 1997] The mediator between the measured data and the joint probability distribution is a kernel function g(**Z**-**Z**n) corresponding to the scattering function of experimental system. [I. Grabec, Website] Most often the scattering function corresponds to Gaussian distribution. In the next step of modelling the probability distribution is converted into the estimator of the mapping relation in Equation 1. For this purpose we assume that the traffic observation system has provided a set of N joint statistical samples: {**Z**n = (**Q**n, **V**n) ; n = 1, 2, … , N} and describe the probability density function f(**Z**) in terms of statistical samples by the kernel estimator: [I. Grabec et al., 1997]

$$\mathrm{f}(\mathbf{Z}) = \frac{1}{N}\sum_{n=1}^{N} g(\mathbf{Z}-\mathbf{Z}_n)$$

Equation 2

Determination of f(**Z**) corresponds to a complete statistical description of the phenomenon and further leads to the definition of statistical estimators characterizing it. [I. Grabec et al., 1997]

From the theory of optimal statistical estimators it follows that function **G** can be interpreted as the conditional average of variable **Q** at given condition **V**: **G**(**V**)**=**E[**Q|V**]. From this definition we obtain from Equation 2 the following model for the predictor of the mapping relation: [I. Grabec et al., 1997], [I. Grabec, Website]

$$\mathbf{Q}_p = \mathbf{G}(\mathbf{V}) = \sum_{n=1}^{N} \mathbf{Q}_n \, \mathrm{S}(\mathbf{V}-\mathbf{V}_n)$$

Equation 3

Here the function $\mathrm{S}(\mathbf{V}-\mathbf{V}_n)$ denotes the measure of similarity between given condition **V** and a stored datum **V**n. It is given by:

$$\mathrm{S}(\mathbf{V}-\mathbf{V}_n) = \frac{g(\mathbf{V}-\mathbf{V}_n)}{\sum_{m=1}^{N} g(\mathbf{V}-\mathbf{V}_m)}$$

Equation 4

The more similar is the given condition **V** to the n-th sample **V**n , the more its complement **Q**n contributes to the estimated value $\mathbf{Q}_p$. The estimator specified by Equation 3 and Equation 4 represents an optimal approach to simple and quite general determination of traffic generating function **G** in terms of joint statistical samples **Z**n=(**Q**n, **V**n) and the kernel function g(**Z**-**Z**n).

The description of variable **Q** by predicted $\mathbf{Q}_p$ is generally not exact. Their agreement is most simply described by the correlation coefficient r. In order to determine it we need *modelling data* {**Z**n = (*Q*n, **V**n) ; n = 1, 2, … , N} and *testing data* {**Z**k = (*Q*k, **V**k) ; k = 1, 2, … , K}. The model Equation 3 and Equation 4 is built yy the first ones and after that its performance is evaluated using the second ones. From the sets of actually observed

Qk and predicted data Q_p : { (Qk, Q_p k) ; k = 1, 2, … , K} the correlation coefficient is then calculated by the standard formula from statistics.

In order to proceed along the proposed route of modelling one has to provide records of past traffic flow field **Q**(t,**r**) and related influencing variable **V**(t,**r**). These data are usually contained in different data bases but need very little additional pre-processing in order to adapt them to the modelling algorithm.

2.10.2 Data Needs

In order to apply the non-parametric modelling in creation of new data we have to provide a data base of joint statistical samples {**Z**n = (Qn, **V**n) ; n = 1, 2, … , N}. In the most simple case the components of vector **V** represent the hour H, the day-code D, and the weather state class W : **V** = (H, D, W). If we are interested in the traffic rate Q at a given hour H of a certain day D under given weather state W in the future, we apply these values in model Equation 3 and apply the model to predict the corresponding traffic rate: Q_p (H, D, W). However, for this purpose we must provide joint samples {**Z**n = (Qn, Hn, Dn, Wn) ; n = 1, 2, … , N} by past observations.

From this consideration one can derive the following data classes which are needed for this approach:

- road weather data
- statistical traffic data as loop detector data, Floating Car Data (FCD) or other traffic flow observations

2.10.3 Examples

Estimation of missing data

As an example we consider a one-component flow rate Q at Delsjön station in Göteborg district in Sweden (Figure 2). The data have been provided by Roadidea Partners Semcon and Klimator in Sweden and refer to the period from 1 Nov. 2007 to 31 March 2008. The records of time series are given in Figure 3 and Figure 4a. The values of day-code denote: 1 – day after weekend or holiday, 2 – normal days (Tuesday, Wednesday, Thursday), 3 – day before weekend or holyday, 4 – weekend or holiday. Similarly the values of weather class denote: 1 – dry, 2 – wet, 3 – heavy rain, 4 – frost formation, 5 – light snow on cold surface, 6 – light snow on warm surface, 7 – drifting snow, 8 – heavy snowfall.

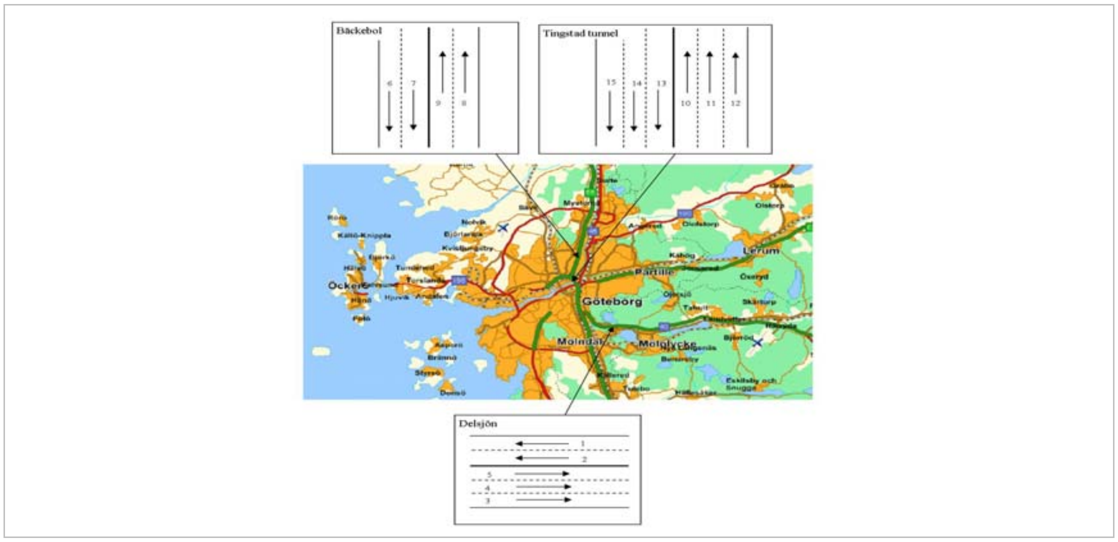


Figure 2: Map of Göteborg district

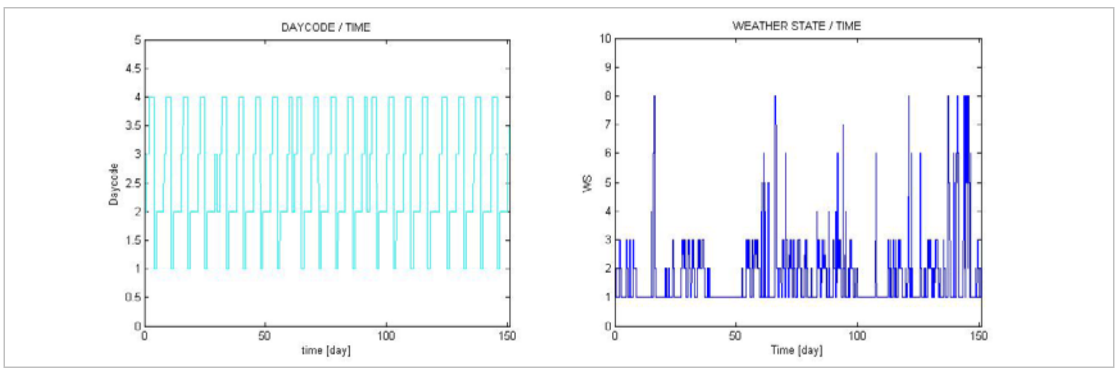
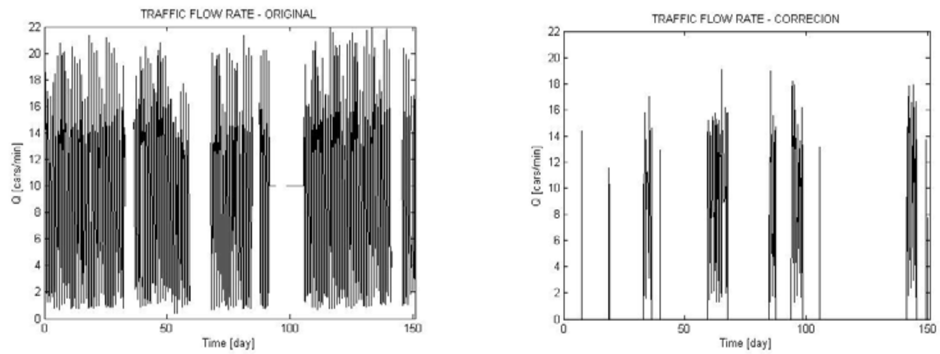


Figure 3: Day-code and weather-state-class series corresponding to Delsjön Sweden.

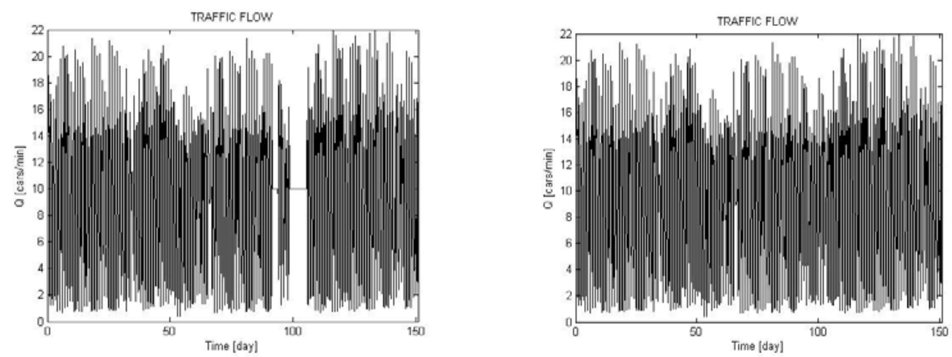
In order to demonstrate applicability of non-parametric modelling in estimation of missing data we utilize data provided by traffic observation at Delsjön station. Figure 4a shows a record of as received traffic flow rate *Q* in dependence of time. Blank gaps in this time series correspond to missing data. Beside these an interval with constant data takes place there. The prediction model was first built based upon reliable past data only. The model was then used to predict the missing data. The corresponding record is shown in Figure 4b. These data were further joined with data presented in Figure 4 to get the record shown in Fig.4a.

The modelling accuracy was tested based upon data shown in Figure 5a. At a certain time the model was built from the past traffic flow data and further applied to predict the traffic flow in the next time moment. The predicted value was compared with the actually measured one and the correlation coefficient between both data was estimated from a record having a length of one week. The dependence of the calculated correlation coefficient on time is shown in Figure 6a. The correlation coefficient in the interval with constant traffic flow rate is rather low $r \sim 0.3$, while otherwise it is $r \sim 0.85$. This indicates inconsistency of constant data with the predicted ones.



a – as received data	b - predicted missing data

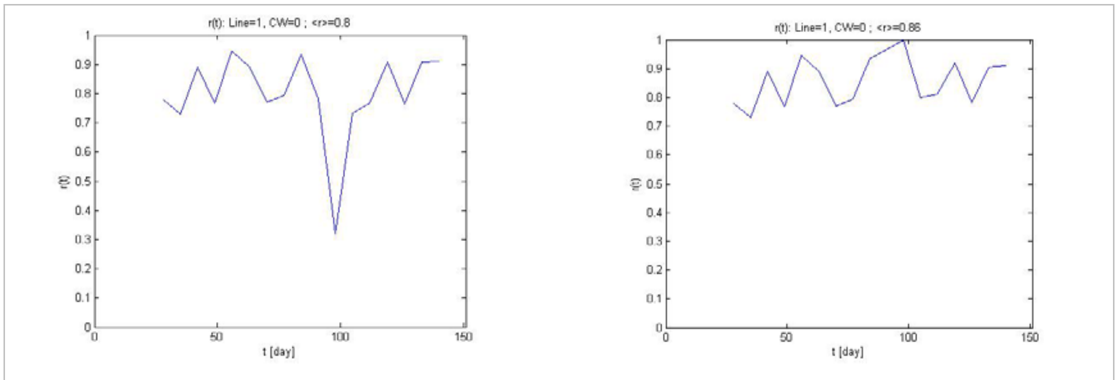
Figure 4: Time series of traffic flow rate on line 1 at Delsjön in Sweden.



a – with unreliable constant data	b – with predicted unreliable data

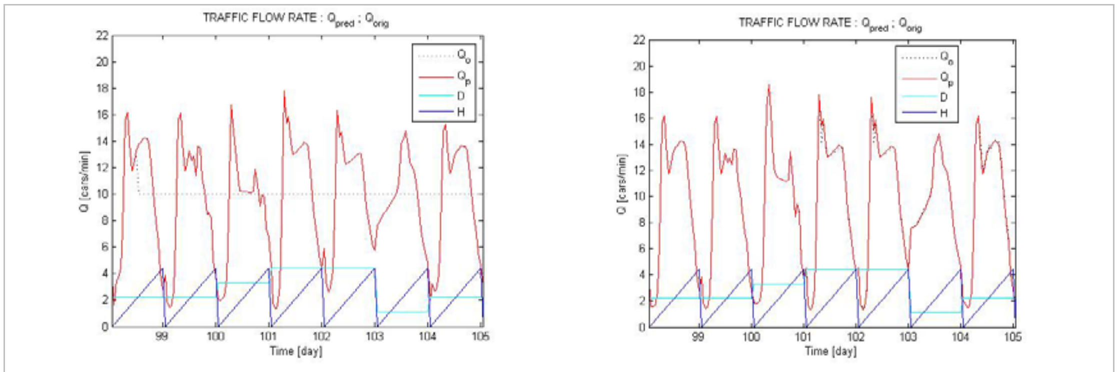
Figure 5: Time series of corrected traffic flow rate.

Figure 7a shows the predicted (solid red line) and as received data (dotted line black) in the corresponding interval. In order to avoid the evident inconsistency the constant data were replaced by the predicted ones. The time series record corresponding to this case is shown in Figure 5b. The corresponding dependence of correlation coefficient on time is shown in Figure 6b, while record of predicted data and those from the basic record in the unreliable interval is shown in Figure 7b. The record in Fig.4b appears rather homogeneous with no outstanding inconsistencies. This is also quantitatively confirmed by a high value of correlation coefficient (Figure 6b) and the agreement of predicted and basic data in the unreliable interval Figure 7b.



a– with unreliable constant data b – with predicted unreliable data

Figure 6: Dependence of correlation coefficient on time for traffic flow rate.



a – comparison with constant data, b – comparison with corrected data

Figure 7: Predicted traffic flow in region with non-reliable data of traffic flow rate.

<u>Predicting the influence of weather on traffic flow</u>

In order to predict the influence of weather on traffic flow we next apply the corrected time series shown in Figure 5b. Using data from the joint time series {$\mathbf{Z}n = (Qn, Hn, Dn, Wn)$; n = 1, 2, … , N} a model for traffic flow rate Q was built sequentially at increasing time values. The purpose of our modelling was to provide for forecasting of traffic flow based upon given data about hour, day and weather. Consequently, just past data were utilized at each selected time. In the prediction two possibilities of weather-state class specification were tested: in the first one (dented by class 0) the weather-state class was taken as it was really observed, while in the second one a supposed class value (from 1 to 8) was utilized. The second possibility was examined in order to show how the class obtained from weather prediction could be utilized in the prediction of traffic flow properties.

Similarly as in the previous case the modelling accuracy was tested by predicting the traffic flow rate in successive weeks. Figure 8 shows the record of the predicted and actually observed flow in the last week when using actually observed weather-sate class as the condition. From such a record the correlation coefficient between the predicted and actually observed traffic flow rate was estimated in each week of prediction.

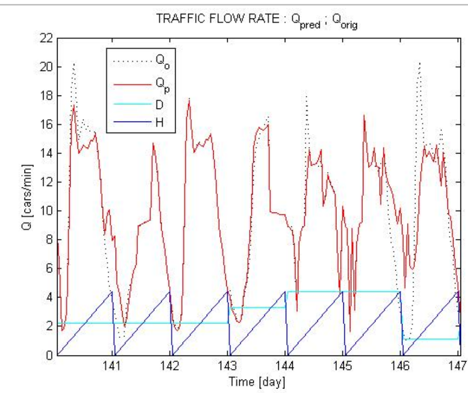


Figure 8: Records of predicted (solid red) and measured (dotted black) flow rate during the last week when using actually observed weather state class as the condition (W=0). The remaining two curves denote the hour (saw-tooth) and the day-code (step-like) variable.

Figure 9 shows the correlation plot corresponding to the records given in Figure 8. The value of correlation coefficient $r \sim 0.9$ indicates rather good prediction accuracy. Figure 10 shows dependence of correlation coefficient on time for actually observed weather-state class (W=0). Similar performance as in the last week was observed also over the complete time interval, so that the mean value of correlation coefficient is $<r>\sim 0.86$.

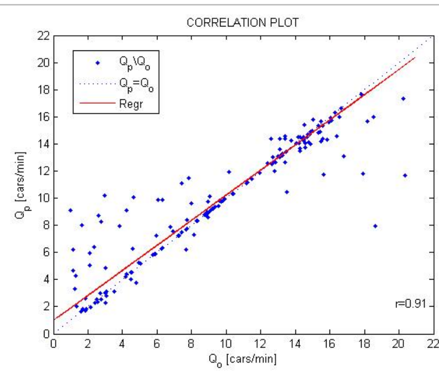


Figure 9: Correlation diagram corresponding to the records given in Figure 8.

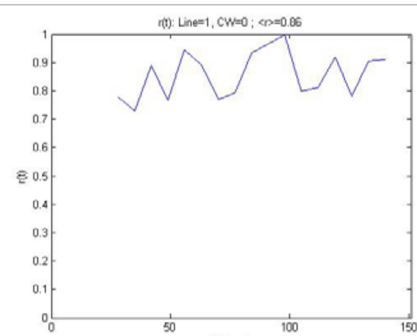


Figure 10: Dependence of correlation coefficient on time for actually observed weather state class (W=0).

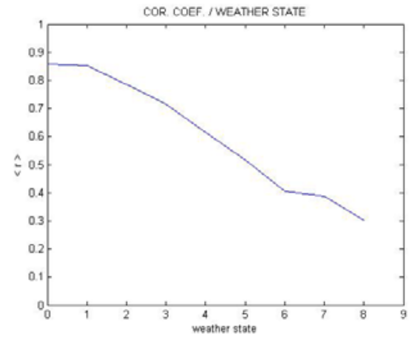


Figure 11: Dependence of the mean correlation coefficient on weather state class.

Similar calculations as with actually observed weather condition were performed also with supposed conditions. Figure 11 shows the dependence of the mean correlation coefficient on the weather-state class. As could be expected, a supposed condition generally does not correspond to a real situation and consequently, the prediction accuracy is diminished. With the increasing weather severity, the probability of coincidence between the supposed and actually observed class is decreasing and consequently, the accuracy of prediction based upon a wrong supposition is also decreasing. However, at low values of weather-state class, the probability of coincidence between the supposed and the actually observed weather class is high and therefore the correlation coefficient is there not significantly smaller than in the case when a true weather condition is utilized.

2.10.4 Conclusions

In this chapter of the report the focus is on description of non-parametric statistical methods for creation of new data based upon traffic flow modelling. Since the complete non-parametric model stems from recorded data only it is rather generally applicable. In spite of the complexity of traffic phenomena, non-parametric statistical modelling of relations between characteristic variables by conditional average estimator appears quite promising for generation of new data. Examination of various examples [3-8] has shown that this method can be simply utilized in prediction of traffic field distribution and related variables such as integral of traffic activity and an optimal travelling time interval. Beside this, non-parametric approach renders possible a simple joining of weather and traffic data in modelling. Data generated by such a treatment can be simply transmitted to participants in the traffic over existing communication devices such as mobile telephones or internet.

2.11 Port

2.11.1 Description

The idea

Main Ports in Europe benefit considerably from a remarkable increase in the growth of global trade. Transport flows between East-Asia and Europe dominate international trade relations. The Port of Hamburg is no exception. In the recent years the container transport volumes have risen dramatically, and prognoses say that the volumes will again double in the coming eight years. If these prognoses are true, several main ports in Europe will face considerable problems handling the hinterland connection, and the hinterland connection will become the bottleneck in the performance and competitiveness of the ports. Innovative traffic planning and management measures can support the mitigation of this problem. Hamburg will face challenges in rail, inland navigation and road transport. Although the main hinterland transport is carried out by rail, road transport plays an important role in handling local and regional container transport, e.g. in the so-called "Wet Triangle" made up by the port cities of Hamburg, Bremen and Bremerhaven. This calls for suitable solutions for road connections and for a closer look at the correlation between ship arrivals/departures and the development of road freight traffic. ROADIDEA could consider the road hinterland connection depending on the traffic demand generated by the activity of maritime transport. This has to be done by a closer look at the transport chain of containers from/to the ships including the ship arrivals and departures, the transhipment process, the container handling and the road traffic. The normal road traffic conditions need to be taken into account as well, because road container traffic naturally mixes with it.

The road network of the Port of Hamburg and in the wider surroundings of the port is characterised by a high share of HGV traffic. A part of this traffic is generated by the loading and unloading business of the container terminals and eventually by ship movements. If it is possible to predict the correlation between ship arrivals/departures and generated HGV traffic on the roads, a better traffic prognosis could be made. This could be done in a first step by a model for road traffic planning. Depending on the availability of online data, this approach could be made dynamic for traffic management in a second step by involving measured online data.

The availability of predictions would allow road traffic planners to design the transport infrastructure according to the future demand, to consider alternative transport modes and schemes or to plan more intelligent technologies for traffic control and traffic information.

The availability of online predictions of the impact of maritime transport may lead to planned short-term traffic management measures in order to be able the handle peaks in road freight transport.

Envisaged Result: Theoretical concept for a prediction model for generated road container traffic based on the analysis of container ship arrival.

Technical approach

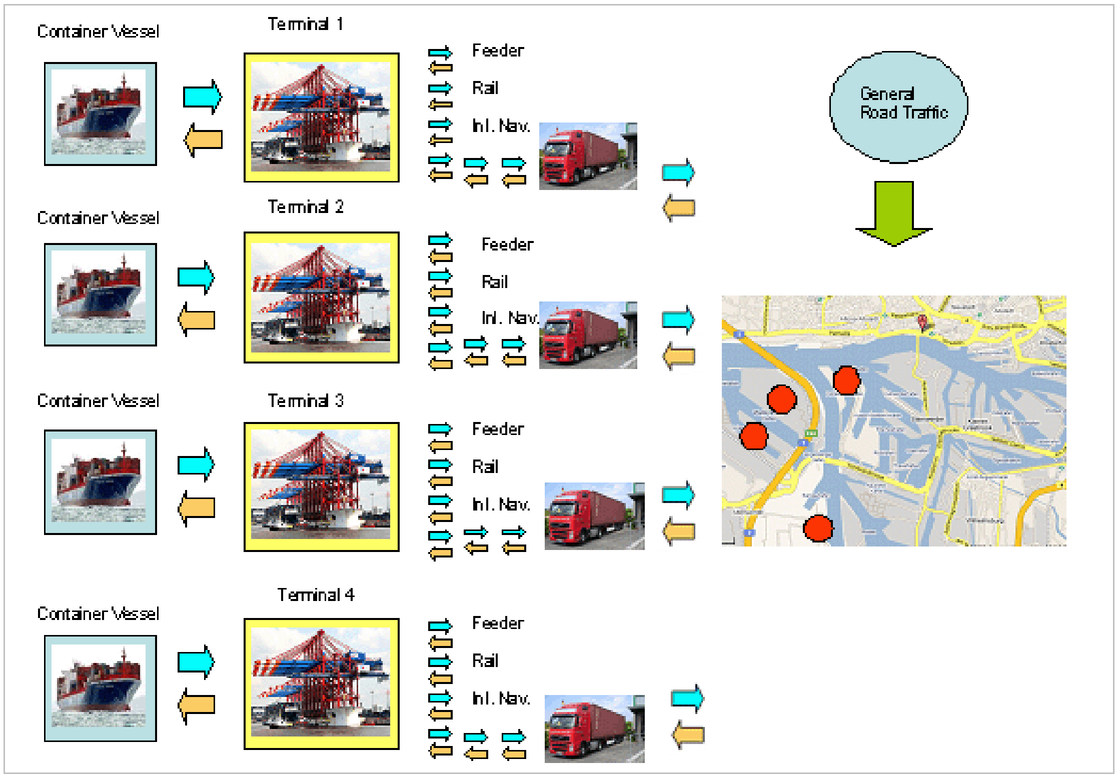


Figure 12: The Modelling Problem

1. Prediction of the generated road container traffic generated by the container terminals

In the Port of Hamburg, there are currently four important container terminals, which work 24 hours a day and 7 days a week. It has to be analysed how the correlation is between the arrivals/departures of vessels and the generation of road container traffic at the container terminal gates. The relationship between the number of incoming and outgoing containers on the side of the vessel and the amount of trucks coming to and leaving the terminal shall be determined. This function is influence by a number of factors related to the operation of the container terminal.

Today, it can be determined in advance from available data how many containers a vessel will bring and pick up. Also the destination of the container is known in advance. The interesting question is how many containers will appear when at the terminal gate as road traffic and where will they go by road/ where do they come from.

This shall allow a prognosis of the road container traffic at a certain time depending on the knowledge of ship arrivals/ departures.

2. Patterns of the general road traffic on the relevant network

In order to estimate the impact of the container traffic on the overall road traffic on the relevant network in the port, the general pattern of the traffic development must be determined.

3. Prediction of the road traffic pattern depending on the ship arrivals/departures

If the predicted amount of road traffic generated by the container terminals is know, conclusions shall be drawn on the impact of the dynamics of the road container traffic on the overall road traffic on the network. The modelling shall deliver a tool, which predicts the road traffic situation at a certain time based on the knowledge of the ship movements.

Geographical scope

The application is foreseen for the Port of Hamburg. The road network concerned is the road network in the port and the one of the motorways and long-distance roads in Northern Germany.

The work on this idea has also significance for the transport capacity planning in relation to several larger ports in Europe.

2.11.2 Data Needs

Currently there are three identified suitable data sources:

- DAKOSY Datenkommunikationssystem AG is the transport information service provider for the Port of Hamburg. The company operates a number applications (SHIPS, TRUCKSTATION, ZAPP) which contain interesting data sets about:
  - all planned **ship arrival and departure times**,
  - all **container** numbers of arriving and departing container on board the ship together with the ship and the destination and transhipment terminal,
  - **historical data** about container truck movement at terminals
- The database of HPA contains historical **traffic counts** on all important road network parts.
- The HPA maintains a model for container traffic including ship moves, berth times and amounts of cargo to be loaded and unloaded, including simulations of ship moves.

These sources would need further detailed investigation. Furthermore, it needs to be investigation where further suitable information can be obtained.

Main data sources and necessary background information may be found with the container terminals and the container transport companies.

Online data in real time is not yet available.

2.12 Free Data

2.12.1 Description

The idea

Free Data is the first and according the seminar participants, most popular of the three general ideas that were short-listed for further study after the First ROADIDEA Innovations Seminar. This idea is not suggesting a new service as such, but a new general data policy that would affect many present and coming transport services in Europe. In principle, Free Data indicates that key data sources for transport services - i.e. weather observations and models, road weather observations and models, traffic volume data, car measurements and other geospatial data - should be accessible and available free of charge (or with minimum copying costs) and in a convenient manner for any service provider for further utilisation. These various data sources may be from public or private sources. In the following, these both alternatives are considered.

In the US and Japan, open and unrestricted data policy is a part of everyday life, resulting in very fast product development and much wider variety of information services and number of companies providing the services. In Europe, investments in production of Public Sector Information (PSI) are about half of what they are in the US, but the economic value using this information is only 20% compared to the US market. Several studies have indicated that the restrictive data policy in Europe is the main barrier and reason for this huge gap. European Commission has reacted to this fact by developing several Directives, which in theory support open data policy. Directive on the public access to Environmental information (1990, revised 2003), Directive on the re-use of Public Sector Information (2002, just now under revision), and the INSPIRE Directive (2006) are examples of Directives giving legal guidelines on the re-use of e.g. weather information.

However, in practice nothing much has happened. Weather information is still in most EU countries very expensive and for small SME weather service providers a strong barrier to entry to the market.

Road weather information is in some countries (like in Finland) available free of charge, but in most countries it is governed by several public or private bodies and completely inaccessible.

Map information is also still very expensive in Europe, though recently due to the American sources such as Google, free alternatives are now available and useful in some applications.

There are also many private data sources that would be very useful in developing new services, but that cannot be accessed even if one is willing to pay. Good examples are CAN-bus output which various car manufacturers still keep as their confidential asset, and available to their own customers only.

The availability and pricing of other transport information varies considerable from country to country and in all cases requires contacts and contracts with the source. There is now centralised transport information server or portal in Europe.

Technical approach

The data that is needed and that is useful in service innovations is the same, whether free or with a price. Thus the first step is to achieve the decrease in data prices through the advent of open data policy in all EU countries. There are some positive signs and decisions in some of the countries (the Netherlands, the UK, Spain and the non-EU country Norway), but the progress has been extremely slow in the past decade. In the new East-European EU countries, the effect to the overall situation has been unfortunately very negative.

But nevertheless, let us be optimistic. Sometimes in the future, Europe will also adopt open and unrestricted data policy. Transport service developers can afford to use several data sources and combine data in new innovative ways, and integrate new data to their existing services. Second major step will be implementing a European-wide data server for easy and efficient access to data. With present web technologies, it is a trivial task technically, but major effort politically and administratively. Here a concerted action from the European Commission is definitively needed.

Geographical scope

Free Data covers all geographical scales and entire Europe. Hopefully also the surrounding countries gradually adopt open data policy, such as in the US and Japan today, to allow global service solutions.

2.12.2 Data Needs

Free Data concept covers all data that is nowadays used and that will be used in the future. Realistically, some data, e.g. from private sources, may still have a price tag and licensing conditions. However, the necessary combination of reliable and comprehensive data set for a transport data provide should be available for a reasonable price, just like the PSI Directive already indicates.

2.13 RTFM, service usability

2.13.1 Description

<u>The idea and background</u>

RTFM (= Read That Formidable Manual) is the second general idea, dealing with the usability of services. The goal is to get better and tailored user-interfaces of text, image, audio, and considering personal characteristics of users, such as language, disabilities, age, health, and other personal needs.

It is not uncommon that an initially brilliant new service idea will never turn out to be a success among users, because it is implemented in a way, which is incomprehensible to begin with, and needs thorough training (or at least reading that thick, depressing manual) before the first use. Also the usage of the service once learned may be cumbersome. Very simple matter and small failure may turn the potential user to give up and even try to use the new service. That is why thorough user tests are of utmost importance before any serious service launch.

Investments and time used for development of better and more comprehensible user interfaces result inevitable into better and also economically more profitable services. This general principle concerns practically all transport services, and more generally, all information services. User groups that are more vulnerable than on the average are those with deficiencies to use common user interfaces (mobile phones, PCs, PDAs, navigators etc), such as the elderly or disabled.

Disabilities and other functional impairments may hinder the usage of ICT devices. According to a study, on average 17% of all Europeans had somewhat limited and 6% of all Europeans had severely limited capabilities to carry out normal activities because of physical or mental health conditions during past six months. Limitations were common in particular amongst the older people. Of the studied respondents aged 85 and over, only quarter had not been limited at all in their activities past six months. [EBS 283, 2007]

There are several EU projects that study the different aspects of user interfaces in changing societies. FP7 Project ICESTARS (Integrated circuit/EM simulation and design technologies for advanced radio systems-on-chip) is a collaborative project that aims to design technologies for advanced radio systems-on-chip. The project has technologically oriented perspective.

FP7 Collaborative project NAPA-WINE (Integrated circuit/EM simulation and design technologies for advanced radio systems-on-chip) aims to analyze the massive deployment of IPTV platforms that will facilitate the change of paradigm of current TV broadcasting from mass TV towards personalized TV.

FP7 Intra-European Fellowships program INFERENCEHCI (Inference and machine learning methods in human-computer interaction) plans to construct human-computer interaction interfaces that are more efficient, less fatiguing and more interesting to use.

FP7 Collaborative project 3DPHONE (All 3D Imaging phone) tries to develop technologies and core applications enabling a new level of mobile 3D experience. The project includes interface and application development.

FP6 Specific Targeted Research Project EIAO (European Internet accessibility observatory) wanted to improve access to Internet content for people with special need such as visual impairment or users accessing web content by mobile phones.

FP6 Specific Targeted Research Project ENABLE (A wearable system supporting services to enable elderly people to live well, independently and at ease) developed a personal, user-centred enabling system, with services, for use by an elderly person in or out of the house, to mitigate the effects of any disability and to increase the quality of life.

FP6 Integrated Project WINNER (Wireless World Initiative New Radio) worked towards enhancing the performance of mobile communication systems. The purpose was to make mobile communication systems more adaptable to user needs.

FP6 specific Targeted Research Project IM@GINE IT (Intelligent mobility agents, advanced positioning and mapping technologies integration interoperable multimodal, location based services).

Technical approach

The necessity of having better and more convenient user interfaces concerns practically all technical solutions used for transport services now and in the future. Thus each developed service must be tested and screened from the specific technology point of view, be it for example based on web browsers, small mobile phone screens or electronic variable signs along roads. Those services targeted for professional users such as road maintenance personnel may be designed to require more complexity and comprehension compared to services targeted for any road user. Some general rules apply, such as textual information must be big enough to read, and that colour blindness must be taken into account when designing warning services with colour codes.

Language is a common barrier for efficient use of services. Within European single market area people, goods, services and money move as freely as within one country. Over 15 million EU citizens have moved to another EU country. Trade within the single market area has risen by 30% since 1992. All and all the single Market has increased the EU's prosperity by 2,15% of GDP. [EU-SM-Benefits, 2008] Besides the 23 official languages in Europe, there are over 200 different languages spoken in Europe. From 400 million living in Europe 60 million speaks other than the official language of the resident country. The multitude of different languages and the level of language proficiency is a real problem with user-interfaces. Many of the ICT device manuals are proven to be too long, poorly translated and impractical. The multitude of functions may stay unused because of the time and effort it takes to familiarize to an ICT device properly. Many difficulties can be avoided if simple graphics is used instead of or supporting the text.

Technological innovations have progressed towards digitalization, miniaturization and convergence of multifunctional and multimodal products. New innovations in user-interfaces are gaining market awareness. GIS positioning based devices, motion sensor games and e-paper are just a few examples.

The problems that may arise from new innovations are much the same as in the ICT field in general such as security issues, user-interface problems, reliability and validity of the content and accessibility problems. Finally, it is necessary to develop dedicated, special services for those with serious deficiencies with their senses, such as the blind and the deaf.

Geographical and demographic scope

RTFM philosophy for developing better user interfaces cover all geographical and demographic scales. It is important to secure the availability of services also for small minority languages and disabled user groups.

2.13.2 Data Needs

There is probably very little difference in input data as such if we compare a transport service with poorly designed user interface to another where good service usability has been considered carefully and used as a guiding line when implementing the final product. There may be cases though when better temporal and spatial resolution of input data – if available and feasible from **economic point of view** - results in better end products. Poor quality of map information may lower user satisfaction and willingness to use the service. Again, better resolution in map data and design of more appealing and clear colour scales may increase the user satisfaction and potential to become a regular user of the service considerably.

In some special services completely new input data must be ingested with existing data to achieve the desired alternative service concept. Examples are integrating Braille text or voice synthesising in services for the blind.

2.14 Stay Home

2.14.1 Description

The idea and background

In the first ROADIDEA innovation seminar in Prague, one recurring idea and question was what can be done to limit transport needs altogether. Is it possible to just stay home?

Transport policies involve a wide range of - and often contradictory - requirements and variables such as safety, energy consumption, greenhouse gas emissions, air pollution, noise, infrastructure provision and maintenance, to mention just a few. At the same time there is a growing need for better transport solutions caused by urban economic growth and the need for reducing greenhouse-gas emissions. ROADIDEA project's main aim is to provide clean, safe and efficient mobility for people and goods using a user-rather than a technology-centered approach.

EU and North America dominate transport-sector emissions in the world and EU alone represents 19,2% of global transport emissions. While environment policies have managed to reduce overall emissions in EU, transport sector emissions keeps rising in the EU area, especially in the new EU states. [GHG, 2008] Both in EU and US the legislative framework has reacted to rising emissions by contributing tougher vehicle emission limits. It has been estimated that traffic related pollution costs Europe 1,7% of its GDP, approximately 160 billion Euros per year. Besides contributing to greenhouse-gas emissions and climate change, aerosol particles from the traffic have several other negative impacts. They may hinder the visibility in urban areas (smog), and also increase the overall level of airborne pollutants causing health problems. Several studies have shown that fine particles from the traffic correlates with mortality. [Tekes, 2006]

The most obvious and efficient way to reduce overall traffic is teleworking and working from home. However if the aim is to study the effects of staying home we must also consider all the different motivators and reason for traveling such as social, economic, political, cultural, religious and other possible reasons. Our daily routines may include numerous separate transport activities that vary according to personal needs.

In a study about attitudes of European citizens towards the environment it was shown that EU citizens attach great value to the environment and are increasingly aware of the impact of its role in their everyday lives. 96% of Europeans stated that protecting the environment was important for them personally. How ever economic factors were seen also important. The study showed that environmental and economic factors influenced their quality of life to a large and nearly equal extent. Socio-demographic factors affected the concept of environment to climate change. The younger the respondents are and the longer they have spent in full-time education, the more likely they are to concept of environment to climate change. [EBS-295, 2008]

European mobility week is an example of the current trend towards more environmentally friendly forms of transport. Growing concern over climate change and the oil price fluctuations have drawn much public attention towards reducing traffic related negative impacts.

Urban living compared to rural living has more impact on environment quantitatively. How ever in general, cities use natural resources more efficient. Concentration of people results more infrastructure, more services and more efficient use of resources. Car use is considerably lower in cities compared to rural areas. Also per capita land and energy consumption are lower. Even if urban living causes many environmental problems, the logical answer to many traffic related problems still lie in urban living. [LIFE, 2006]

Technical approach

Digital technology has changed the rules of communication and information distribution. According to Internet World Stats almost 1,5 billion people use internet with 21,9% population penetration. In Europe the usage is even bigger where over 384 million people have access to the Internet, which means 48,1% population penetration. [IWS, 2008] The majority of European households have a computer and nearly half of the household population has now access to the Internet. [EBS-293, 2008] Digitalization has already enabled several transport needs to be obsolete. It is possible to order food, pay bills, buy tickets and get education via internet just to mention a few of the numerous possibilities that modern technology allows us to do. New technologies improve the quality of life especially for elderly and those who are physically disabled.

In a modern digital society staying home is already a possibility to a group of people. All human needs can be fulfilled thru different modes of digital technology. While restricting all travelling may be economically and environmentally reasoned, the process of socialization and culturally related reasons still dominate our behavior.

Internet usage does not correlate with working from home or teleworking. It has been said that outside the assembly line, almost any work can be done by teleworking. How ever in reality teleworking is still a very marginal phenomenon. Less than 3% of the total working population in EU is working from home. Some 4 % works from home occasionally.

According to European Working Conditions Survey report 2005 employers in European countries use average of around 40 minutes a day to travelling to and from work. Almost 60% of EU workers work all or almost all the time at company premises. On the other hand, although teleworking or working from home is not yet a real alternative to working on company premises, it is used by a substantial proportion of people as a complement to their normal working arrangements. [EWCS, 2007]

The obvious advantages of teleworking include reduction of traffic related negative impacts such as pollution, noise, congestions etc. Recent studies show that overall productivity increases, job satisfaction rises and recruitment and retaining the employees gets easier. Teleworking also reduces company overheads and lowers operational expenses. Teleworking may even persuade population to locate closer to the central business districts and increase centralization of cities. Which means that telecommunications and technology are not decreasing the value of face-to-face interaction and the need for cities and their services. [Sridhar et al., 2003]

There are already many services that help teleworking and reduce the need for travelling. Tele and web conferencing is becoming more and more common, as well as virtual office services. If a person can really individually choose whether going to his work space or staying home, there may be room for new services to help his or her decision making:

- What are the weather risks today along my normal route? The service may calculate a risk factor automatically and warn if a certain level is exceeded.
- What is the saving of CO_2 release if I stay home?
- Maybe the person can earn a bonus from the company if saving a certain total amount a year?

Geographical scope

Staying home philosophy covers Europe as a whole, and the need for transport reduction is actually global.

2.14.2 Data Needs

Transport services that are developed for transport needs, such as route weather forecasting, can be used as such in Stay Home services. Adding extra services such as release calculations, links to virtual office and web conferencing services and other social media services would result in a Stay Home portal. User should define the profile in advance (the daily routes, the car type etc.), but otherwise the service would mainly operate with ordinary **road transport services data**.

2.15 ROADIDEA Data Classification Consideration

As described in the report D2.1 on present data availability of data for transport systems and services a data classification has been designed, which fits best to the ROADIDEA data needs. It comprises all necessary categories to describe all relevant data which have been taken into a closer consideration during the first data source investigation period. These data have been assigned to the data classification. Figure 13 shows the homepage of the ROADIDEA data collection archive where the data classification and the number of dataset for each main category are presented.

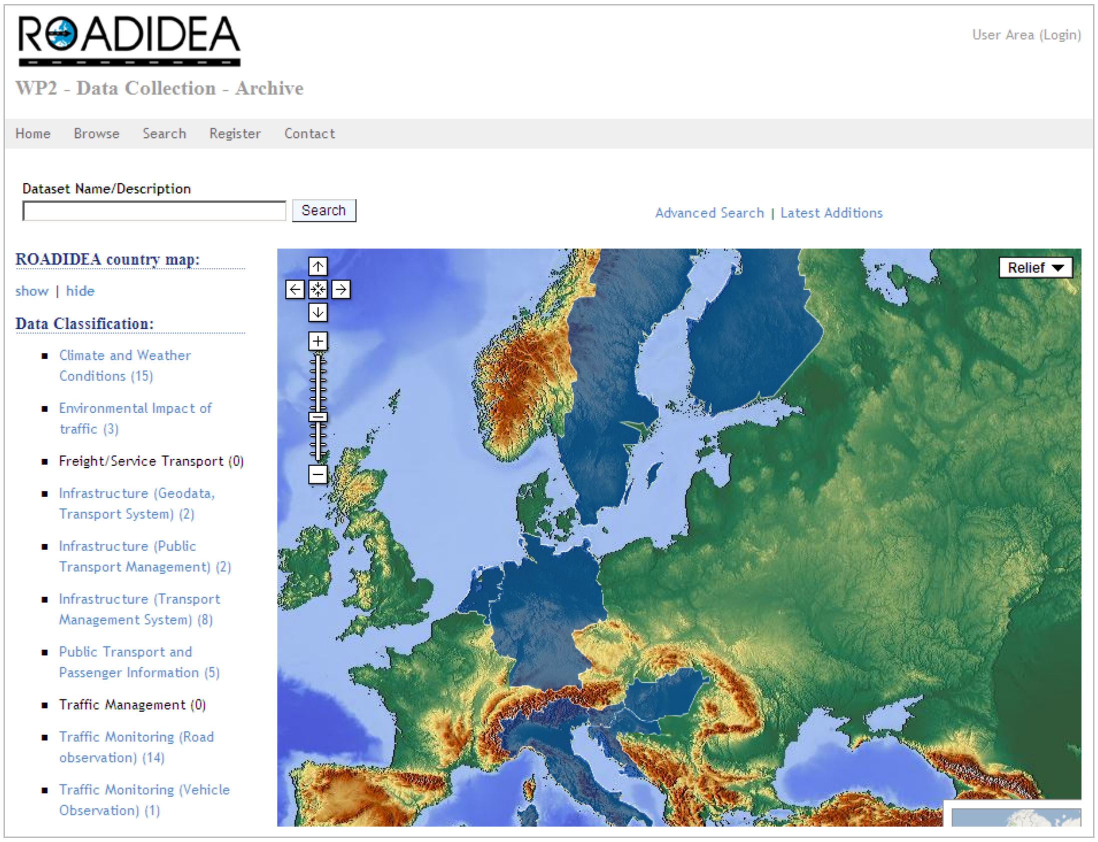


*Figure 13: Homepage of the ROADIDEA data archive (**http://roadidea.dlr.de**)*

As it can be seen, not all main categories could be filled with at least one dataset. For example there are no datasets available for "Freight/ Service Transport" and "Traffic Management". There will be a closer look to all empty or sparsely filled categories.

The main categories are subdivided into further data classes. For instance the main category "Traffic Monitoring (Road Observation)" contains 7 subclasses. Table 1 shows the subclasses and the number of datasets which could be assigned to each subclass of this category.

Traffic Monitoring (Road Observation)		Number of datasets
subclasses	road condition	2
	road works	**0**
	speed/ average speed, travel times	5
	sudden obstructions to road traffic (e.g. incidents, accidents)	6
	traffic density	3

	traffic flow, congestion, level of service	2
	traffic volume	4
	total	**13**

Table 1: Traffic Monitoring (Road Observation)

It is apparent that not all subclasses could be filled with one data set. For example there are no data sets available for "road works" so far. Some of the data sets belong to more than one subclass. That's why there are 14 data sets in total.

As another example Table 2 shows the data assignment for the main category "Climate and Weather Conditions" and there subclasses.

Climate and Weather Conditions		**Number of datasets**
subclasses	air pressure	5
	air temperature	6
	black ice	**0**
	climate and weather monitoring	4
	depth of snow	1
	intensity of precipitation (rain, snow, hail)	6
	relative humidity	5
	remaining salt on road surface	**0**
	road condition	1
	road humidity	**0**
	road surface temperature	**0**
	road weather forecasts and warnings	2
	road weather monitoring	3
	visibility, brightness, fog	5
	weather forecasts and warnings	4

	wind direction	6
	wind speed	7
	total	**15**

Table 2: Climate and Weather Conditions

It is also obvious that not all subclasses of the main category "Climate and Weather Conditions" could be filled. For example there are no data sets available for "black ice", "remaining salt on road surface", "road humidity" and "road surface temperature" so far. Some of the data sets belong to more than one subclass. That's why there are 15 data sets in total.

The following list comprised all main categories and there number of data sets. The list shows also all empty or sparsely filled subclasses and main categories.

- **Climate and Weather Conditions (15)**
    - black ice
    - remaining salt on road surface
    - road humidity
    - road surface temperature
- **Environmental Impact of Traffic (3)**
    - air quality
    - emissions
    - noise
    - pollution reports
    - radiation
- **Freight / Service Transport (0)**
    - freight transportation routes
    - freight transportation schedules
- **Infrastructure (Geodata, Transport System) (2)**
    - address databases (POI, car parks, event locations etc.)
    - camera networks
    - emergency response centres
    - inter-/ innerurban traffic control centres (motorway sections, intersections, networks)
    - permanent counting points (long-distance road network)
    - police departments
    - radar detector networks
    - TIC (police, road administrations)
    - TMC location codes
- **Infrastructure (Public Transport Management) (2)**
    - timetable/ ticketing systems
- **Infrastructure (Transport Management System) (8)**
    - automatic vehicle location systems (AVL)

  - electronic parking guidance systems
  - phone fleets (Floating Phone Data)
  - traffic-adaptive/ traffic-dependent signal control systems
  - truck tolling systems
  - tunnel control systems
- **Public Transport and Passenger Information (5)**
  - disturbances on public transportation
- **Traffic Management (0)**
  - car park occupancy, availability of roadside parking places
  - messages about current traffic situation
  - priority requests at traffic lights
  - routing/ evacuation/ parking management strategies
- **Traffic Monitoring (Road observation) (13)**
  - road works
- **Traffic Monitoring (Vehicle Observation) (1)**
  - vehicle activity status
  - vehicle detection (type)

With respect to the pilots and ideas described in chapters 2.2 to 2.14 some major conclusions are made in the results chapter 3.3.

2.16 Data Innovations from ROADIDEA Work Packages

The chapter describes **efforts** which have been achieved in the project work packages in order to find ways to tap or utilize so far unused data sources – beyond the idea/pilot descriptions. For this chapter I've suggested a structure similar to the project WP structure. At least all work package leader are asked for contribution in this chapter.

2.16.1 WP 2 "in-vehicle data collection"

As stated already in ROADIDEA deliverable 2.1, vehicle can bus data is an interesting data source. It is actually not a new source, but it has tremendous potential which is not been effectively utilised so far. The major truck manufacturers have agreed to give vehicle data access to third parties via open de facto FMS-standard, which provides quite large data set e.g. for fleet management purposes. This has enabled the development of variety of devices, applications and services and therefore created new business opportunities. The passenger and light vehicles have only standardised On-Board Diagnostics (OBD) connector access to vehicle data.

OBD was designed for vehicle diagnostics and inspections purposes and it monitors the vehicle's emissions-control devices and other engine-related components to ensure that they are functioning properly. Currently OBD-II protocol includes engine-related live data and thousands of diagnostic trouble codes, which offers a valuable source of information when troubleshooting problems inside the vehicle. Most of the trouble codes are manufacturer-specific. The generic emissions-related trouble codes are used e.g. by the vehicle inspection sites. There are several aftermarket OBD scan tools (OBD-II connector, cable and PC software) available which are usually can read and clear trouble codes and display live data such as throttle position, engine rpm, air and coolant temperature, etc. This information can be used e.g. in a car repair shop. But data available via OBD-connector does not provide all the information required for more advanced real-time applications such as vehicle data collection for road weather or fuel consumption monitoring applications. This is why some third party vehicle electronics manufacturers have started to develop devices that are connected to vehicle CAN-bus. Even though, digging up on the needed information from CAN-bus is laborious work as there is no standard for the CAN-bus connector and the message configuration in passenger cars.

Methodology

In order to find out the capabilities of a state-of-the-art vehicle CAN-bus based fuel consumption logger and to get hands-on experience, a logger device was tested. The measurements were done within another project (but assisted by ROADIDEA), in which VTT analysed mail delivery driving using a data logger installed to a CAN-bus connector in two vehicles. Only ROADIDEA relevant results (i.e. data acquisition related issues) are reported here.

The logger was ECONEN II from Paetronics, which is a Finnish company that designs and manufactures equipment and software involving vehicle electronics. ECONEN II reads 4 channels from the CAN-bus and outputs 18 features including driving time, driving distance, driving speed, tachometer and instantaneous consumption (litres/hour), etc. It is designed for efficient monitoring of fuel consumption and economical driving education in driving schools and other organisations for both heavy-duty vehicles and passenger cars. The equipment consists of a data logging unit that is installed to the vehicle CAN-bus connector and a display unit for instantaneous analysis of measured data. In this project only the data logging unit with GPS was used and the data was analysed afterwards.

Figure 14: Paetronics ECONEN II data logger and the display unit.

The vehicles chosen for the measurements were a Volkswagen Caddy 1,9TDI car and a Volkswagen Transporter 2,5TDI van, both with automatic transmission. Precise measurement of fuel consumption presumes calibration of the data logger. The logger was calibrated by simulating different driving cycles (Euro test) using VTT's chassis dynamometer. Fuel consumption was measured using an external fuel tank and scales. Distance data was calibrated as well. The actual measurements on the road were performed in normal mail delivery driving during ca. 10 days with both vehicles.

Results

All measurements were carried out successfully. The CAN-bus data was exact and there were no problems in the analysis. This shows the potential of vehicle data usage if the data is accessible. However, the installation of the data logger was challenging. The CAN-bus connectors are not really designed for third party use and thus the placement in the vehicle varies. The expertise of the Paetronics Company helped to find the place but installation was still laborious because of the upholstery and limited space. Even if the CAN-bus connectors could be discovered with reasonable effort the remaining problem is that all car manufacturers have a different set of messages moving in the CAN-bus. Even different models and year models may have a different configuration. As neither the CAN-bus connector nor the message configuration in a car has been standardised it is a demanding task to solve these issues for the large amount of makes and models individually. Currently, data logger manufacturers have to solve this problem 'manually' and consequently this knowledge belongs to their intellectual business capital.

Discussions

Vehicle telematics service market is still immature and most solutions rely on proprietary and closed OEM-platforms, preventing third party developers from creating value added services. Telematics market could grow rapidly if *de facto* open platform and standards are developed, which also foster the development of innovative services [Yunfeng et al.

2007]. One forum which is working with this issue is the Open Services Gateway initiative (OSGi). OSGi technology is universal middleware based on Java. OSGi technology has been adopted in variety industries such as consumer electronics, mobile, communications, software, vehicle and telematics. OSGi defines open specifications and standards e.g. for the delivery of services and applications from different providers. These services are offered via the Internet to the home, car, mobile device environments, and other computing systems. OSGi has been used in European automotive research projects including GST (http://www.gstforum.org/) and CVIS (http://www.cvisproject.org/).

Utilisation of vehicle sensor data has had a great interest in many research projects. GST project defined OSGi based interface to vehicle sensors [GST, 2005]. This work has been continued by OSGi vehicle expert working group which is working on a vehicle Application Programming Interface (API). Vehicle API would provide an uniform access, cross manufacturer, to the static and dynamic vehicle information (e.g. Location, speed, temperature, heading, wind speed, traction, abs activation, airbag deployment, brake lights, technically any sensor, etc.). Currently this information is secret and guarded. The OSGi vehicle group expects the car manufacturers to open up this information in a controlled environment. Moreover, it has been stated that it is not clear if the car manufacturers allow the on board equipment to be used for third party applications, but the vehicle expert group sees that it is possible in the secure software environment. (http://www.osgi.org/Vehicle/HomePage)

Utilisation of the vehicle data has numerable applications such as collection of road weather information with vehicles (Extended Floating Car Data), economic driving support, etc. But these applications will not realize is large scale before a standardised access for third party developers to the vehicle data is available. On the other hand, business case of car manufacturers (current owners of vehicle data) should be taken into account and liability issues must be tackled. This needs cooperation between software and telematics industry, car manufacturers and standardization organisations.

Regulators are not isolated from the above described problem. Unless there is political will to facilitate standardised access to CAN-bus, the automotive industry's motivation is lagged perhaps by many years and the deployment of advanced ITS is reduced in scope and delayed.

2.16.2 WP 3 "methods and models"

Methodology

The main objective of work package 3 "WP3 - Method and model development" is converting data into information; to use data from work package two as input and by models/data fusion produce valuable and correct information as an output.

In order to achieve this, methods for filtering incoming data from errors have to be developed and implemented, as well as models to extract information from the data. Methods to quantify and estimate their performance also play an important role.

Whenever data is used to reflect information, rules describing the data must be used to interpret it. The most obvious one describes the meaning of the numbers, letters etc. It answers the questions: What was measured, where, when and on what scale? This kind of information is essential if you want to make use of the data.

A less obvious kind of information describing the data collected is its accuracy or validity. One reason for this is that can be that the specification/legend of the data is made before the process of collection. The errors in the dataset will be generated unintentionally during the process.

- Categorize different types of errors occurring in input data.
- Find and describe existing general methods for detecting the errors.
- Find/develop algorithms to correct the errors.
- Test the algorithms on existing data for evaluation of performance.

The purpose of the data collected is to reflect a part of reality. It would be too optimistic to expect the data to have a "one to one" correlation with this part of reality. When it is copied from reality via a sensor/observer, through transmission and finally stored, it could be corrupted for many reasons. We are not always sure if a suspected error really is an error and if so where the error has occurred and how should we succeed in recreating the lost information.

Results

Data filtering

The data is treated in series, so the filters are distributed and have intermediate steps where information is passed from one filter to another. At Semcon MATLAB language was chosen for the implementation of the algorithms. Starting with raw data, the format of the raw data can differ; in our case the data came in ASCII format (.txt) and excel format (.xls). In order to be able to filter the data, the raw data files where red by an extractor program and stored in MATLAB-matrix format. The information content is the same; it is just the format that is changed. Both the metadata and the actual data pass this process. No filters are applied to the metadata; the datasets are often small and can be checked manually. The product of the large actual datasets we call extracted data.

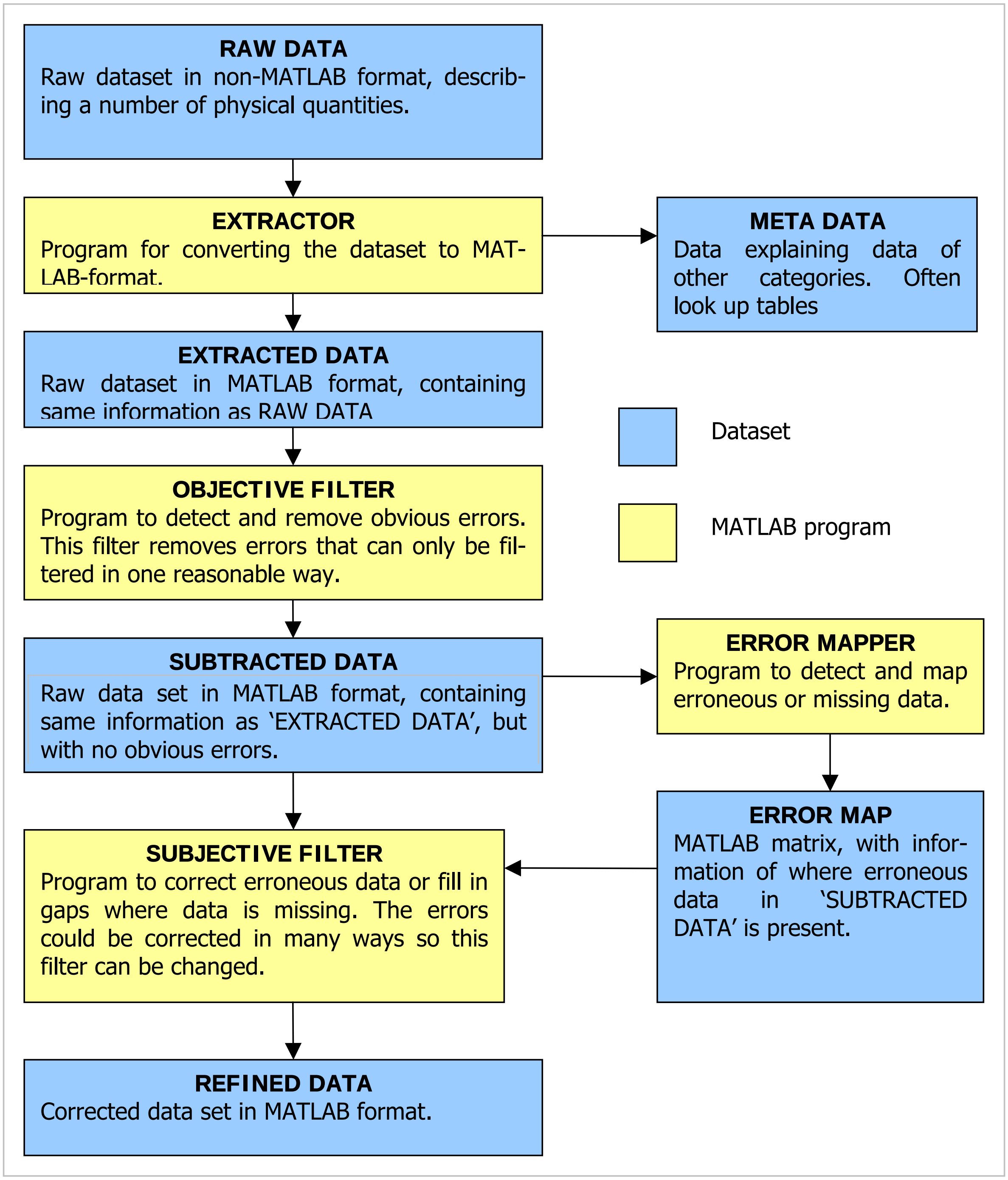


Figure 15: Structure for filtering the traffic data.

Models

Road traffic is a consequence of population activity to which many agents (~100 M) participate. In spite of this, the traffic flow does not exhibit completely random character because the population activity is synchronized to a high degree. The synchronization is stimulated externally by changing properties of the environment, as well as internally by social agreements about working days and holidays. The external stimulation can be physically described by the time and weather variables, while the internal one has to be modeled by some specific dynamic law. In agreement with these properties we treat road traffic as an example of very complex deterministic chaos. Due to external influences we further treat the road traffic flow as a stochastic non-autonomous dynamic phenomenon and try to model its generating equation statistically. The basic information for the creation of the model can be extracted from records of traffic flow rate and related environmental variables.

For this purpose a general non-parametric statistical method is proposed and the applicability of the method is demonstrated by examples of traffic flow prediction in various countries. The performance of the proposed method is quantitatively estimated by a cross-validation method. It is based on estimation of the correlation coefficient between predicted and actually measured traffic data. Examination of the correlation coefficient dependence on the structure of the condition indicates how the method can be tuned to a specific case of modeling. The final goal of our approach is to provide for a quantitative forecasting of traffic flow rate that is needed for an efficient information support to participants in road traffic and an advanced approach to its intelligent control.

3. Results

3.1 The pilot ideas

Besides generating new innovations for the European transport sector another main objective of ROADIDEA is the implementation of some pilot services which are able to demonstrate the potential of these innovations in terms of business applicability, reliability, utilization, and user acceptance.

As already stated in the chapter 2.1 of this report a couple of ideas were born at the ROADIDEA innovation seminar in Prague. Some of those ideas were suggested as pilots during group suggestions, evaluations and basketing, and finally they were pointed out in the last plenary session where all preferred ideas and pilots were considered and discussed by all participants. The fog pilot was not coming from a ROADIDEA innovation idea. As early as the proposal stage of the project ROADIDEA it was planned to tackle the fog problem in the Veneto region. With the project partner ARPAV the fog detection and forecasting was implemented in the project proposal strategically. Nevertheless, the idea to engage the Italian fog problem was discussed, evaluated, and basketed during the innovation seminar in Prague as well.

In a later technical meeting the top list from the Prague seminar was discussed there and the friction model and the Gothenburg case were decided as the most feasible for actual piloting. Fog case has not yet been in a state for concrete pilot. Finally, the Hamburg pilot was taken because of its general interest and its model-like combination of traffic modeling and weather forecasting. So, it was decided to take those four ideas into further pilot consideration. Besides that, a general distinction between "practical"

- Friction pilot,
- Gothenburg pilot

and "theoretical" pilots

- Port pilot and
- Fog pilot

has been made. The implementation of the practical pilots and the preparation and feasibility studies of the theoretical pilots are topics of the work package 6 "Creation of pilot services". The work of WP6 is closely connected to the data and its availability considerations in WP2. In order to provide the necessary data and supplementary information for the pilots WP2 is involved in the pilot service implementation work package intensively.

The four pilots are described in the following more detailed.

3.1.1 The Pulp friction pilot

For the description of the "Pulp Friction" pilot see chapter 2.2 of this report.

3.1.2 The Gothenburg pilot

The main purposes of the pilot are:

- Study the data in a research perspective
- Investigate what kind of algorithms that is possible to use for this specific pilot.
- Investigate key factors that might be of value and have impact to the Gothenburg pilot.

The roads in many cities have not grown in as rapidly as the traffic flow has increased and this requires a big effort to meet the traffic needs now and in the near future.

To test the ideas, Gothenburg is proposed to a Road Idea Pilot, due to data access from several sources in the area and also that both Klimator and Semcon; ROADIDEA partners are located there.

Gothenburg's traffic situation is characteristic due to that the city is separated by a river and there are only three connections over the river close to city, the last one built forty years ago. Since then, more roads have been built in Gothenburg but no more connections over the river. This causes a lot of traffic problems and might be typical for other cities with growing transport needs.

The challenge in this project is not to look on separate sources. The challenge is to combine them together and use different algorithms connecting the sources in a new way.

3.1.4 The Port Pilot

Modelling the Multimodal Traffic Situation in the Port of Hamburg

<u>Situation at the Site</u>

In a modern mainport like Hamburg, container transport is the dominating branch of the port business. So, road container transport will have an effect on the general traffic in the road network which provides access to the container transhipment terminals and the other port facilities which are involved in the container handling process (e.g. the customs offices, the container depots, the veterinary office, logistics companies).

Hamburg has currently five significant container terminals in the port. All of them generate road freight traffic.

By far not all containers which are transported by the ships end up on the roads as road traffic. They are treated in different ways:

- Most containers are transported from and to the hinterland by road transport.
- A considerable part of containers is transhipped to another vessel.
- A considerable part goes by rail transport.

- There is also a need to transfer containers from one terminal to another. This is mainly done by road transport. However, alternatively also barges are used.
- a few containers go by inland waterways.

The decision on the transport mode is taken by a huge range of forwarding agents, and it is unlikely to identify all decision makers here. So the container terminals have no influence over the chosen transport mode.

Furthermore, container moves from and to container depots may be triggered which are done by road transport.

The containers are all treated differently on the container terminals. The time spans how long the containers stay on the terminals vary around a statistical curve. So, the prediction of the exact time when a single container will appear as a container move on the road is not trivial.

Also, procedures at the customs office and other locations in the port may influence the vehicle moves.

The Hamburg Port Authority (HPA) is responsible for planning und operating the road network in the port area. This means, HPA is responsible for the road construction and the traffic planning for all roads in the port area. It HPA`s interest to keep the traffic on the roads fluent and to provide access to the container terminals and all other port facilities.

For the future, HPA consider a complete re-organisation of the container hinterland transport on the roads, introducing an inland hub. For the justification and planning of these measures, it would be helpful to model the generation of container flows.

HPA does yearly traffic counts and has a statistical overview of the general development of road traffic in the relevant network. Also, the share between normal road traffic and HGVs can is determined. The counts cover selected locations in the net a single day of the year. The counting takes place regularly once a year.

Currently, there is no online traffic data available.

The Hamburg Port Authority is currently planning the establishment of a port-related traffic management system. Part of this system will be a road traffic model for the port which is now built up step by step. An important feature of this model should be a prediction of road freight traffic depending in the maritime transport activity.

Purpose of the Pilot

The purpose of the pilot is providing a prediction tool for the generation and distribution of road traffic depending on the ship arrivals and departures.

The problem consists of two parts:

- the determination of a correlation between ship arrivals and road traffic generated by the container terminals

- the distribution of the generated road traffic on the road network (building on the available VISUM approach).

There are two levels of pilot development. A first and slightly simpler problem is to provide a model for traffic planning. This model would then predict the road HGV traffic volumes based on planned ship arrivals/ departures and based on statistical data on road traffic. This approach would support road traffic planning purposes.

The second and more complex problem would be to make the process dynamic based in real ship arrivals and the online traffic situation based on measured road traffic data. This approach would more support traffic management purposes.

The pilot could be planned within ROADIDEA. It is not feasible to implement and demonstrate the pilot with the time frame and resources of ROADIDEA.

<u>The Challenge for Innovation</u>

The determination of the correlation between ship departures/arrivals and the generated road traffic is most likely the most demanding part of this investigation.

The process is not so simple that a direct correlation between ship movements and the activity on the road network is obvious. For example, a ship arrival would not cause tangibly a larger amount of trucks coming to the port. The problem is more complex and has to pay attention to the following factors. More likely is a correlation where the overlapping processes of container handling at all terminals lead to and overall logistic situation which correlates with the overall road traffic situation. Factors to be considered may be:

- the fact that the Port of Hamburg has five container terminals working in parallel all generating road traffic,
- the fact that only a certain share of handled containers reappear as road traffic,
- the amount of containers loaded and unload by a ship,
- the fact that several ships are loaded and unloaded at the same time,
- the fact that the size of ships varies,
- the fact that are 20 ft and 40 ft containers,
- the final destination of the containers in the hinterland,
- the varying times for a container to stay on the terminal,
- road container traffic which is generated by transfer between the terminals,
- road container traffic in relation to container depots,
- opening times of the different organisations,
- generated trips with empty containers and chassis

Most likely, the solution cannot be found by tracing single container moves. A statistical model needs to be generated which defines input and output parameters for the calculation of the generated HGV volumes at the terminal gates together with a distribution of their likely destinations/ origins.

In a second step, these HGV moves must go into an O/D calculation, and it must be calculated how the generated HGV traffic flow is distributed in the road network and how it mixes with the general traffic flow which has nothing to do with road container traffic.

Here, delays at customs gates and other service stations need to be considered as well.

As a bonus, these processes could be made dynamic based on online measured data such as

- real ship arrivals (interface to AIS system for ship tracking)
- measured road traffic data.

Necessary Local Partners

The pilot needs the co-operation with HPA, DAKOSY Datenkommunikationssystem, HHLA (largest terminal operator) and maybe the container transport companies.

Necessary ROADIDEA Partners

The pilot needs workforce for data analysis and concept development.

First of all, a model specialist needs to find a suitable mathematical model for the processes. The correlation between ship arrivals/departures and generation of road container traffic needs to be but into a mathematical set of formulae. Furthermore, the relationship to existing modelling tools for the distribution of road traffic like VISUM needs to be defined.

- AMANOVA for support in modelling
- VTT for concept building and evaluation
- DLR for concept building

Next Steps and Outlook

The planning process needs to be started up together with HPA. HPA has declared an interested in this pilot under ROADIDEA. DAKOSY Datenkommunikationssystem has announced the willingness to give support with data.

Next steps should be:

- Setting up a local team consisting of ROADIDEA members and local partner and pilot kick-off
- Collection and Analysis of identified data sources with DAKOSY and HPA
- Clarification of conditions for the receipt of data
- Further conceptual ideas on the model and the modelling concept
- Identification of the need for further research into the process.

3.1.3 The Fog pilot

The idea

Fog is a relatively frequent phenomenon in the Po Valley and constitutes a major issue for all road traffic. Fog monitoring is particularly difficult because the phenomenon exhibits large horizontal variability and is measured at only very few sites as such sensors are not part of the standard equipment of the principal meteorological surface monitoring network. Satellite observations can partially contribute to observing fog and low visibility but they need to be suitably combined with surface observations.

The idea is to develop an areal monitoring system of the fog presence on Venice Region territory, combining in a novel way ground based observations and satellite imagery, and develop suitable products for disseminating to the end users, where various end users groups with potentially different needs will have to be identified. These fog monitoring products have to be clear and simple to understand and suitable for broadcast via Internet and mobile phones.

The main result expected is to set up a system which allows possibly large end user groups to access visibility information in real time as an important element for their decision making process. This process can lead to pursue several options:

- Change of route in view of thick fog. For professional end users this can mean taking a detour but saving travel time and lowering the risk to incur in an accident.
- Break the trip until visibility conditions improve. For professional truck drives who are subject to systematic rests could optimize their planning.
- Privates could decide to modify the route, take the train instead of the car, or simply postpone the trip.

There are numerous scenarios which can illustrate the usefulness of such a service in making travelling and transport more efficient and less dangerous.

Technical approach

The pilot idea is relatively simple in that visibility related products will be developed and disseminated to end users in real or near-real time. For the sake of modularity, the pilot is conceived in three distinct parts, i.e. the:

1. production layer (→ identify, acquire, and integrate suitable data into visibility related products);
2. dissemination layer (→ identify and implement suitable dissemination channels, which probably will be web and mobile phones),
3. end user layer (→ identify end users and their specific needs).

Geographical scope: 'regional' for several regions in Europe

The topography of the political region Veneto, located in north-eastern Italy, is structured in a plain (2/3 of the area) bordering the Adriatic coast and the pre-alpine chain, the pre-alpine chain and the Alps (Dolomites). These areas are characterized by rather distinct fog climatologies, where the plain exhibits by far the largest fog occurrence, along with a number of alpine valleys.

As the pilot will rely on in situ data as complement to the areal satellite observations, the homogeneity of area of interest is of primary importance to allow for horizontal extrapolation of the in situ observation. Valley climates can be quite distinct from the plain and from valley to valley. The pilot will therefore focus on the plain and its applicability in other regions of Europe will be limited by (complex) topography.

Data needs

Fog is a relatively 'slow' phenomenon on the time scale of minutes to fraction of an hour. It does however evolve on the timescale of one to several hours. From the point of view of temporal coverage the 15-minute frequency of satellite data seems appropriate, as well as the 10- and 5-minute frequency of meteorological data measured by ARPAV.

Satellite data can provide a good starting point for fog and low cloud conditions. They provide areal coverage with horizontal resolution of 1 to 3 km in longitude and 1.5 to 4.5 km in latitude. The limitations to the 'satellite-only' approach to monitor fog are given by medium and high level clouds which can cover the fog and low cloud areas, and an inherent uncertainty which arises from the difficulty to judge whether the low clouds extend to the surface or not with very different effects on the visibility.

Thus it becomes important to have direct **in situ observations** of the quantity of interest, i.e. the visibility. For this purpose, ARPAV has identified so called visibilimeter which are capable to yield a quantitative measure of the visibility in meters. For Roadidea a network of 10 such devices has been designed on the basis of a rough fog climatology which was compiled from four years worth of observations paired with satellite imagery for this purpose. The reason for recurring to such a 'homemade' climatology is that systematic fog climatologies are virtually inexistent.

This compilation identifies areas with coherent fog behaviour, which can be summarized as in Figure 16, and the following list:

- A: western Alpine foothills (highway stretch Brescia-Padova and Brennero-Modena);
- B: central western plain (highway stretch Brennero-Modena);
- C: southern plain, Po river surroundings (highway stretch Padova-Bologna);
- D: central northern plain (highway stretch Brescia-Padova and Padova-Venezia)
- E: central eastern plain (highway stretch Padova-Bologna);
- F: coastal region and Adriatic Sea (highway stretch Autovie Venete).

In addition, a distinct seasonal behaviour emerges from the compilation:

- Sep-Nov: most frequently observed fog is advection fog and radiation fog, in particular in areas A, B, C, and D;
- Dec-Feb: radiation fog is predominant, frequent, and occasionally persistent, particularly in areas A, B, C, D, and E;
- Mar-May: mostly radiation and advection fog, more frequent close to the relatively cold Sea (around 10°C), mostly in areas E and F;
- Jun-Aug: rare occurrence of radiation fog and low level clouds caused by occasional advection of very humid air from the south, mostly in areas E and C.

Another criterion for positioning the visibilimeter is of technical nature related to the operation of the instruments and the data transmission. The sites of the automatic surface weather stations operated by ARPAV are a natural choice. As not all the sites are equally equipped, only those with sufficient electrical power supply, data port, and hourly data transmission are candidate sites. Of the around 40 sites, about 20 are located in the areas with significant fog occurrence.

For the remaining candidate sites the following criteria were applied in order to maximize the representativeness of the visibilimeter network:

- The station height should be representative (no peaks);
- There should not be a humidity bias (e.g. close to rivers, creeks);
- Avoid excessively shaded areas;
- Avoid orographic circulation patterns (e.g. valley outflows);
- Avoid urban areas, local turbulence, and building tops.

Given all these criteria and the number of available stations the identified areas were covered on an approximated grid of 25 km x 25 km.

The result is displayed in Figure 16.

It’s possible to see that the aims to install at least one visibilimeter for each area have been achieved. The northern plain and eastern foothills present a lower incidence of fog in comparison with the other areas, so we excluded, during this project, the extension of the visibility network to that areas. Also we excluded to monitor visibility reductions in mountain areas, because of the irregular phenomenon distribution in complex topography (too much visibilimeter should be necessary!).

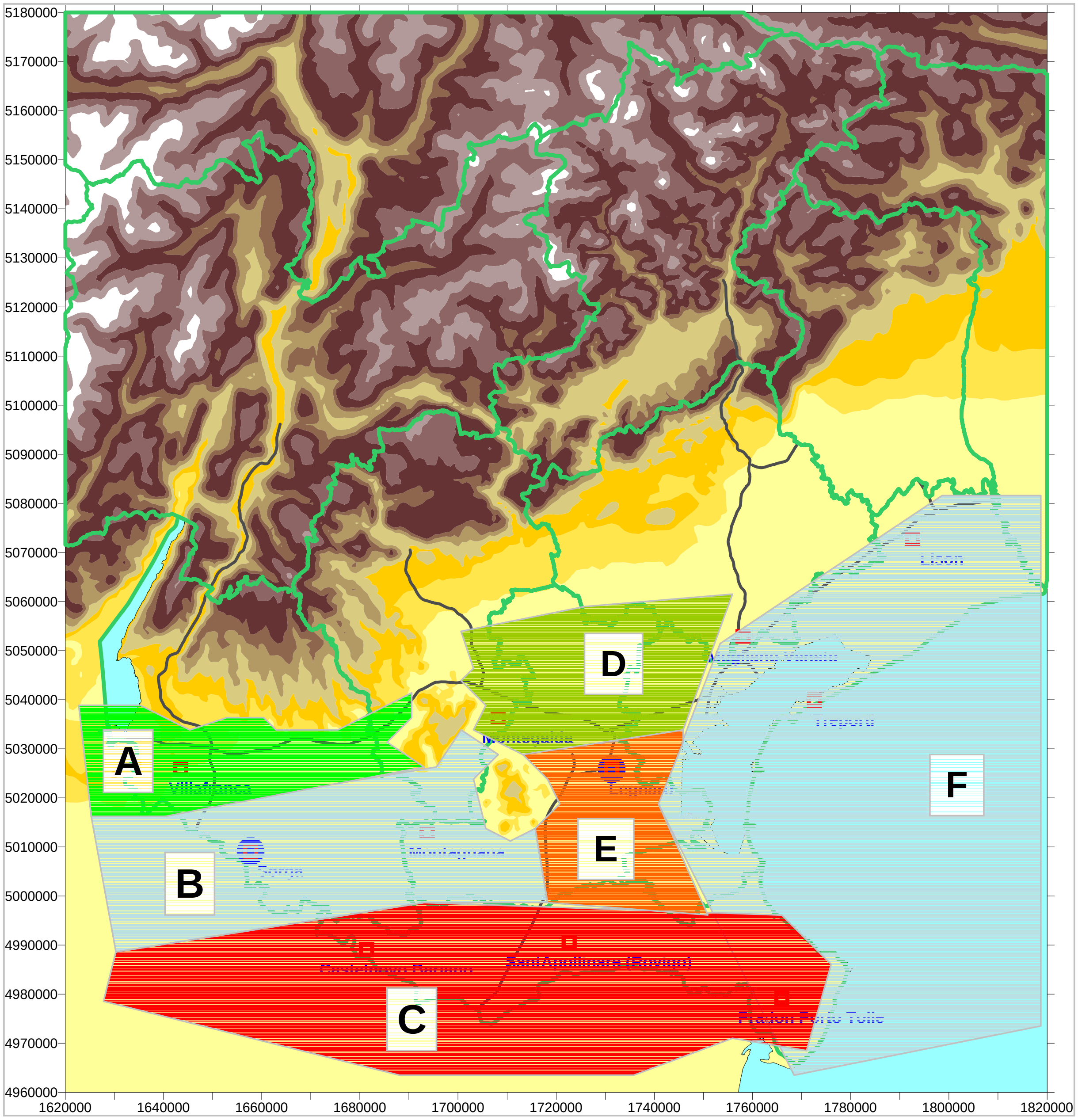


Figure 16: Outline of the visibilimeter network (red squares) as a result of the fog climatology. The coherent fog areas are denoted by the letters A-F and are described in the text.

Data processing

The combination of the satellite and in situ observations constitutes the major scientific challenge for the ARPAV pilot as no proven methodology can be found in the literature. In principle, the combination of satellite data with visibility observations is not limited to quantitative measurements, such as derived by the visibilimeter. It is conceivable that semi-quantitative and qualitative information be included into the procedure. Such information could be provided from webcams, car-mounted visibility sensors, or driver reports. In this case, the data integration system needs to be able to account for the relative quality/uncertainty of the information source.

The data management for the ARPAV pilot includes the following steps:

- Acquisition of satellite data on a predefined grid covering the area of interest;
- Acquisition of visibilimeter data of the ARPAV Roadidea network and storage in the ARPAV central data base;
- First data quality control procedure (gross error and plausibility checks);
- Transmission of visibilimeter data to the Roadidea data platform.

3.2 Overview of needed and available data

The following table gives an overview of ROADIDEA pilots and ideas and the corresponding data needs. The table also tries to state the general data availability. The first column specifies whether it is a ROADIDEA pilot or an idea.

<table>
<tr><th></th><th>Name</th><th>Needed data</th><th>Data availability</th></tr>
<tr><td rowspan="7">Pilot</td><td>Pulp Friction</td><td>Online and historical Road weather Data</td><td>Available and utilised at FMI for Scandinavian countries</td></tr>
<tr><td rowspan="3">Gothenburg</td><td>Various road traffic counting’s</td><td>Available, access currently still unclear, discussion ongoing process</td></tr>
<tr><td>Road weather</td><td>Free available, road authority permission needed</td></tr>
<tr><td>Airport weather</td><td>Free available</td></tr>
<tr><td rowspan="2">Port</td><td>Ship schedules and port logistics</td><td rowspan="2">Given and utilised for modelling purposes by the Hamburg Port Authority (HPA), free availability to be clarified</td></tr>
<tr><td>Port traffic counts</td></tr>
</table>

	Fog	Satellite data	Available and utilised by ARPAV
		In-Situ observations	Visibilimeter network designed and in implementation
Idea	**Cross Border Weather Alerts**	Weather alerts, road weather alerts and observations	www.meteoalarm.eu for weather alerts. No centralised service for road weather observations
	Mobile Phones as Sensors	Cell phone tracking	Not available
	MyRoute Mobile Pocket Guide	Road traffic observations	Available for some regions with different coverages
		Weather (Current & forecasted)	Available for some regions with different coverages
	MyTravel Toilet TomTom	Facilities and points of interest	Commercial available via Map Service Provider
	In-Vehicle Information	Can bus vehicle data	Historical data from SRIS project (Slippery Road Information System) available by Semcon
		Road traffic counting	Access via the Swedish road administration to both historical and live data of traffic counting.
		Road weather	Free available for some countries, general availability to be investigated
		Airport weather	Free available, Historical weather: Available for Swedish airports.

	EUROADMAP	Road/Grid weather	Heterogeneous, automated access possible in some European countries, according data source investigation needed for some countries
		Airport weather	Free available
	EYEAR	Friction data/ CAN bus data	Only proprietary data samples used by car manufacturers for internal usage and research purposes
		Vehicle positions	Some vehicle fleets available, however CAN-bus access is very limited
	Traffic Forecast Models	Road traffic statistics	Available for some regions
		Road weather	Available for some regions
	FREEDATA	General	Comprehensive data source identification and description needed
	RTFM	Acceptance, usability, economics	Not to be covered
	StayHome	Road transport service data	Generally available

Table 3: Data availability listing

Conclusion:

It has become obviously that most ideas are tailor-made for specific given both traffic and weather problems in specific European countries. Having this in mind, it does not surprise that the data availabilities vary country-wise with the degree of necessity to tackle those specific problems. Scandinavian countries are often faced with slippery road conditions, while this problem might never appear in Adriatic countries. Hence, the northern European countries spend big efforts monitoring road weather conditions for reliable road condition monitoring and forecasting and, as a consequence, they have obtained a big expertise to handle this issues. Also the degree of the development of the entire national transport systems and its adjustment to European regularities and policies varies country-wise in Europe.

In order to solve urgent problems, in some model regions needed data for specific ideas are available with a very good coverage and reliability - or those ideas are even ready to be implemented. In other regions traffic and road weather data availabilities and responsibilities still have to be clarified or decided.

Thus, it has to be pointed out that all described ideas and pilots are not yet ready to be implemented in all European (member) states. In order to pilot the ideas specific regions (countries, international highways, geological regions, etc.) have to be assessed in terms of pilot applicability and decided accordingly. This assessment has to comprise at least the following topics:

- Necessity to tackle the corresponding issue,
- Data availability (charges, coverage, accessibility, etc.),
- Public authorities support and involvement,
- Reliable business model,
- Expected affect.

WP2 may support all idea teams performing this analysis.

Another crucial obstruction to implement some ideas is the diversity of data formats for similar contents. Grown historically, there are many data formats for both traffic and weather observations, whereas the traffic data sector is obviously more complex. In order to make data sets comparable, which is a prerequisite for implementing pilots in cross-border regions covered by different data sources, first of all a data format standardisation is needed. This can hardly be achieved in the ROADIDEA project completely – if ever only for a few examples. Several European projects and initiatives are about to standardize data formats containing similar contents. All projects are confronted with the huge complexity of this issue.

As for the CAN-bus data, which are needed for some ideas (e.g. In-Vehicle Information and EYEAR), only a very few proprietary data samples are given. A direct access to a fleet of vehicles providing CAN-bus data is not available by now. On this specific field no change can be expected within the near future. Car manufacturers are not willing to open their CAN-busses for free access due to secrecy policies. It has to be investigated whether CAN-bus data from truck manufacturers are accessible more easily. If so, according ideas have to be adjusted to truck costumers needs. Truck manufacturers have to be contacted and invited to be involved into the implementation of those pilots, which might be applicable for truck purposes. Nevertheless, further investigations of CAN-bus data sources from cars have to be performed.

As stated in the table above, for some ideas/pilots as the Gothenburg pilot the data availability and accessibility discussion with the according road authorities and data providers is a running or ongoing process, which has not been finished by now. As for the practical pilots "Gothenburg" and "Pulp Friction", it can be proceeded from the assumption that all needed data are generally available and ready to be accessed online if needed. WP2 is involved in the implementation of the practical pilots and will support the establishing of online data feeds. These data streams are to be implemented into the planned WP2 data mediation architecture as well. The development of the WP2 data mediator will be performed in parallel with the pilot implementation. The mediator is to be set up with the task 2.4.

For all European countries reliable weather data and weather forecasts are given. Very often these data sets are free of charge and may be implemented in desired applications easily. With partners from the weather monitoring and forecasting sector a very good experience in terms of acquisition and utilisation of weather data can be applied immediately.

3.3 Data classification

While some data classes are filled properly, other data classes are empty or filled with a very few data sets. Generally spoken, it has to be determined whether empty data classes are not needed for the ROADIDEA approach. The data classification has been made at the beginning of the project, where the question, which pilots are to be developed further, has not yet been answered. From this point of view, it could be possible to have unnecessary data classes in the ROADIDEA data classification.

This might be true for instance for the class "traffic management", which is empty. There are no pilots or ideas dealing with traffic management means by now. However, it seems to be wrong for the example data class "Freight/ Service Transport", with the subclasses "freight transportation routes" and "freight transportation schedules". Data from these data classes are needed for instance for the Hamburg port pilot.
On the other hand, also at first sight well filled top-level classes as "Climate and Weather conditions" as shown in Table 2 of chapter 2.15 may comprise empty sub-classes which are quiet relevant for some pilots or ideas (e.g. subclasses "black ice", "road humidity" and "road surface temperature" for friction or weather alert pilots).
The mentioned examples emphasize both

- the need of a re-consideration of already described data sets and its assignment to given data classes and
- the need to continue the data source identification process during the entire project runtime.

A closer look into the data base shows, that project partners have mainly investigated data sources for those classes which were most relevant for their everyday work. Comprehensive data availability analyses have only been made for Finland and Germany (which does not mean that all data sets have been described with the WP2 data questionnaire). In order to fill the major data classes for all ideas/pilots data needs and for all member countries a multi-disciplinary approach is needed, which forces all partners to investigate data from all relevant fields (general data catalogue approach emphasized in the conclusion chapter).

Besides that, it has to be clarified that data mostly come from running services, applications or use cases. In almost all cases the data covers not only one data class; complex aggregated data can comprise content for different data classes. Aggregated data often can not be disaggregated to get easy-to-assign raw data sets. Thus, it was hard to make a right data assignment to the classification. To use the same example as above: There are data about black ice and road humidity, but, it is included in data bundles from weather measurement stations, which were assigned to another data class (here: "relative humidity or road weather monitoring).

With respect to the given data needs, at least for the following data classes the data source investigation has to be strengthened:

- traffic monitoring (vehicle observation) data, with special respect to CAN-bus data collected in the vehicle and sent to a central server instance
- freight/service transport data, for the theoretical Hamburg port pilot
- infrastructure (geodata, transport systems) data, for supplementary data sources and traffic management authorities
- environmental impact of traffic, for addressing future interrogating

4. Conclusions

The practical pilots "Gothenburg" and "Pulp Friction" have been decided to be implemented as ROADIDEA pilot services. The pilots are well described – beyond the description given in this report - and the implementation work for it is in progress. In several project and sub-group meetings relevant piloting attributes as relevant data sources, access possibilities, charges, contracts, and work loads for contributing ROADIDEA partners have been discussed in the preceding month – or will be discussed in the coming weeks. **In the according regions (Gothenburg for the "Gothenburg pilot" and Finland for "Pulp Friction") all necessary data are available in general for the mentioned pilots.** The Friction approach can easily be adapted to other regions, where sufficient road weather observations are given. This is not valid for the Gothenburg pilot, which needs the set-up of more elaborate prerequisites.

With this report all pilots and ideas from the ROADIDEA innovation process are considered concerning their data needs and given data availabilities. The general rule is - as already stated above - that almost all **pilots and ideas (except "Mobile phones as sensors") can be implemented in at least one specific region, where the necessary data are available and ready to access immediately**. To enable cross-border or multi-regional consideration of traffic, road or weather conditions, suitable regions or use cases have to be found and assessed concerning pilot applicability.

Some ideas as "StayHome" or "FreeData" are not such to be implemented as services. They rather consider a general handling with data, traffic and mobility behaviour.

However, further efforts concerning data source identification are needed urgently. From WP2 point of view **it would be reasonable to decide a minimal ROADIDEA catalogue for data, for which the availability has to be investigated for every participating country.** This approach is to be followed in the idea "Euroadmap" described in chapter 2.8 for road weather data. This approach could be enhanced easily to other data classes. A general data catalogue might increase the possibility to enable pilot applicability for more than only one piloting region. It also might enable diversity consideration as well as necessary data standardisation and comparison achievements, which are needed for cross-border or multi-regional approaches.

The latter recommendations are to be decided in the ROADIDEA technical coordination group. But, this might also be a subject to be dealt with in another research project.

As shown in the chapter 2.16 some achievements concerning new data sources and data merging and filtering methods were obtained in the ROADIDEA work packages. Especially the **CAN-bus data** consideration in 2.16.1 emphasized the general shortage of according data sources. CAN-bus data acquisition and provision for both private cars and trucks in combination with car-to-car-communication might be very powerful for generating new and innovative services for the public and private transport service sector. The same is valid for **mobile phone tracking data**, which might also be a very valuable data source for end user information services, if security and privacy issues were solved. **Additional warning information** from various sources should be integrated to a central information database. To overcome these shortages and barriers for data usage, respective data investigation efforts should be strengthened as soon as possible.

References

[EBS-283, 2007] Health and long-term care in the European Union. Special Eurobarometer, Report 295, European Commission, 2007.
http://ec.europa.eu/public_opinion/archives/ebs/ebs_283_en.pdf

[EBS-293, 2008] E-Communications Houshold Survey. Special Eurobarometer, Report 293, European Commission, 2008.
http://ec.europa.eu/public_opinion/archives/ebs/ebs_293_full_en.pdf

[EBS-295, 2008] Attitudes of European citizens towards the environment. Special Eurobarometer, Report 295, European Commission, 2008.
http://ec.europa.eu/public_opinion/archives/ebs/ebs_295_en.pdf

[EU-SM-Benefits, 2008] How the EU Single Market benefits you.
http://ec.europa.eu/internal_market/benefits_en.htm

[EWCS, 2007] European Foundation for the Improvement of Living and Working Conditions. Fourth European Working Conditions Survey. 2007.
http://www.eurofound.europa.eu/ewco/surveys/ewcs2005/results.htm

[GHG, 2008] Greenhouse Gas Reduction Strategies in the Transport Sector. Preliminary Report. OECD/ITF, 2008.
http://www.internationaltransportforum.org/Pub/pdf/08GHG.pdf

[I. Grabec et al., 1997] I. Grabec, W. Sachse, *Synergetics of Measurements, Prediction, and Control,* Springer, Berlin. 1997

[I. Grabec et al., 2008-04] I. Grabec, K. Kalcher, F. Švegl, *Modelling and Forecasting of Traffic Flow on Slovenian High-Ways*, Proc. of the conf.: Transport Research Arena Europe 2008", Ljubljana, SI, April 21-24, 2008, ISBN 978-961-6527-12-5 (DRC)

[I. Grabec et al., 2008-05] I. Grabec, K. Kalcher, F. Švegl, *Statistical Forecasting of Traffic Flow Rate*, Proc. of the 14th Int. Road Weather Conf. 2008", Prague, CzR, May 14-16, 2008, ID 50, ISBN 978-80-87205-01-3, disc published by Wirelesscom, www.sirwec.cz

[I. Grabec et al, 2008-10] Igor Grabec, Kurt Kalcher, Franc Švegl, Statistical forecasting of traffic activity, 9. SLOVENSKI KONGRES O CESTAH IN PROMETU, Portorož, 22.-24. oktobra 2008, Mobilnost, P-8.

[I. Grabec, Website] Several references of interest for non-parametric statistical modelling of complex physical laws are available at web page of I. Grabec: http://www.fs.uni-lj.si.

[GST, 2005] GST – Global System for Telematics (2005). GST Open Systems Architecture and Interface Specifications. DEL_SP_3_1. Available at
http://www.gstforum.org/en/downloads/deliverables/sp-level_deliverables/

[Helbing, 1997] D. Helbing, *Verkehrsdynamik*, Springer, Berlin, 1997.

[IWS, 2008] Internet World Stats, 2008. http://www.internetworldstats.com/stats.htm

[LIFE, 2006] LIFE in the city - Innovative solutions for Europe's urban environment *LIFE ("The Financial Instrument for the Environment")* is a programme launched by the European Commission and coordinated by the Environment Directorate-General (LIFE Unit - BU-9 02/1). 2006 http://ec.europa.eu/environment/life/publications/lifepublications/lifefocus/documents/urban_lr.pdf

[Porenta et al., 2008-09] Tine Porenta, Kurt Kalcher, Franc Švegl, Igor Grabec, Statistical Modeling and Forecasting Activity of Traffic Folws, Int. Conf. »Applied Statistics«, Ribno-Bled, Slovenia, Sept- 21-24, 2008, Program and Abstracts, published by Satiatistical Society of Slovenia, Eds.: Lara Lusa, Janez Stare, http://www.statistics.com, p.66.

[Porenta et al., 2008] Tine Porenta, Statistično modeliranje in napovedovanje aktivnosti prometnih tokov, Diplomsko delo, Fakulteta za matematiko in fiziko, Oddelek za fiziko, Matematično fizikalna smer, Univerza v Ljubljani, 2008, (Statistical Modeling and Forecasting Activity of Traffic Folws, Dissertation of Dipl. Eng., Faculty of Mathematics and Physics, Dept. of Physics, University of Ljubljana, 2008).

[ROADIDEA 2008a] D2.1 Present availablity of data. Public report available from www.roadidea.eu.

[ROADIDEA 2008b] D5.1 Innovation plan. Public report available from www.roadidea.eu.

[ROADIDEA 2008c] D5.2 Results of the First Innovation Seminar. Public report available from www.roadidea.eu.

[Sridhar et al., 2003] Kala Seetharam Sridhar, Varadharajan Sridhar. The Effect of Telecommuting on Suburbanization: Empirical Evidence. The Journal of Analysis & Policy. 2003 33:1 http://www.jrap-journal.org/pastvolumes/2000/v33/33-1-1.pdf

[Tekes, 2006] Reducing Particulate Emissions in Traffic and Transport. Views and Conclusions from the FINE Particles – Technology, Environment and Health Technology Programme. Tekes. http://www.tekes.fi/julkaisut/Fine_Liikenne.pdf

[TIS, 2007] Traffic Information Services on the Trans-European Road Network. Success stories, current practice and perspectives. Traveller Information Services (TIS) Expert Group. May 16, 2007, version 5.3.

[Yunfeng et al., 2007] Yunfeng Ai; Yuan Sun; Wuling Huang; Xin Qiao, "OSGi based integrated service platform for automotive telematics," *Vehicular Electronics and Safety, 2007. ICVES. IEEE International Conference on* , vol., no., pp.1-6, 13-15 Dec. 2007
URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=4456369&isnumber=4456353

[i] Health and long-term care in the European Union. Special Eurobarometer. Report. 2007.

[ii] http://ec.europa.eu/internal_market/benefits_en.htm

[iii] Greenhouse Gas Reduction Strategies in the Transport Sector. Preliminary Report. OECD/ITF, 2008

[iv] Reducing Particulate Emissions in Traffic and Transport.

Views and Conclusions from the FINE Particles – Technology, Environment and Health Technology Programme. Tekes.

[v] Attitudes of European citizens towards the environment. Report 2008. European Commission.

[vi] LIFE in the city - Innovative solutions for Europe's urban environment *LIFE ("The Financial Instrument for the Environment")* is a programme launched by the European Commission and coordinated by the Environment Directorate-General (LIFE Unit - BU-9 02/1). 2006

[vii] Internet World Stats. http://www.internetworldstats.com/stats.htm

[viii] E-Communications Houshold Survey 2008. European Commission. Publication 2008.

[ix] European Foundation for the Improvement of Living and Working Conditions. Fourth European Working Conditions Survey. 2007. http://www.eurofound.europa.eu/ewco/surveys/ewcs2005/results.htm.

[x] The Effect of Telecommuting on Suburbanization: Empirical Evidence. The Journal of Analysis & Policy. Kala Seetharam Sridhar and Varadharajan Sridhar. 2003 33:1