QUATRA

State-of-the-art analysis

Deliverable Nr D 2.1

January 2012

nast consulting

TRANSVER

...

This project was initiated by ERA-NET ROAD.

Project Nr. 832570

Project acronym: QUATRA

Project title:

Software and Services for the Quality Management of Traffic Data

Deliverable Nr D 2.1 – State-of-the-art analysis

Due date of deliverable: 01.02.2012

Actual submission date: 23.01.2012

Start date of project: 01.10.2011

End date of project: 30.09.2013

Author(s) this deliverable:

DI Daniel Elias, nast consulting ZT GmbH, Austria

DI Birgit Nadler, nast consulting ZT GmbH, Austria

DI Andrea Schieferstein, TRANSVER GmbH, Germany

Version: draft 01

Executive summary

Traffic management systems are using traffic data from several different data collection sources for various purposes. Therefore the quality of the detected traffic data needs to be ensured and erroneous data needs to be identified. The reason laid within the fact that effective control decisions are strongly dependent on the correctness of the data.

Main objective of the work package "State-of-the-art analysis" is the gathering of knowledge about existing data quality evaluation systems and approaches as well as the collection of research concepts whose contents could potentially be integrated into the QUATRA model and software.

Four analysis categories include details of technical approaches (based on specific data features), traffic engineering approaches (based on logic combinations of traffic flow fundamentals and neighbouring traffic detection sites), guidelines and standard procedures (setting out specifications for the assurance of the data quality of incoming traffic data) and standardised systems and tools (system that are being developed/used at present).

Several interesting technical and traffic engineering related approaches are identified in the final conclusions. The studies contained vital information for the freeway and urban model.

Furthermore the guidelines, standard procedures and standardised systems and tools provide numerous quality conditions that will be assessed during the development of the QUATRA model and software.

Nevertheless the integration of different tests and methods will also depend on the data that will be made available by the Austrian, Swiss and German infrastructure operators.

List of Tables

List of Figures

Table of content

1 Introduction

“ERA-NET ROAD – Coordination and Implementation of Road Research in Europe” was a Coordination Action funded by the 6th Framework Programme of the EC. The partners in ERA-NET ROAD (ENR) were United Kingdom, Finland, Netherlands, Sweden, Germany, Norway, Switzerland, Austria, Poland, Slovenia and Denmark (www.road-era.net). Within the framework of ENR this joint research project was initiated. The funding National Road Administrations (NRA) in this joint research project are Belgium, Switzerland, Germany, Netherlands, Norway and United Kingdom.

Traffic management systems are using traffic data from several different data collection sources for purposes such as visualisation of traffic situation, detection of abnormal road conditions or the generation of appropriate traffic control decisions. Therefore the reliability and plausibility of traffic data needs to be identified and confirmed, faulty data needs to be detected immediately. The reason for these quality assurance measures laid within the quality of traffic control systems itself. Effective control decisions are strongly dependent on the correctness of the underlying traffic data collection.

2 Objective of the state-of-the-art analysis

Main objective of the work package is the gathering of knowledge about already existing data evaluation systems and approaches as well as the collection of research concepts whose contents could potentially be integrated into the proposed model and software approach of QUATRA. Therefore the literature review has been divided into four main categories:

- Technical approaches for identification of faulty detection data based on specific data features
- Traffic engineering approaches for identification of faulty detection data based on logic combinations of traffic flow fundamentals and neighbouring traffic detection sites
- Guidelines and standard procedures setting out specifications for the assurance of the data quality of incoming traffic data
- Standardised systems and tools for the analysis of traffic data quality that are being developed/used at present

3 State-of-the-art analysis

For all of the four categories *Technical approaches*, *Traffic engineering approaches*, *Guidelines and standard procedures* and *Standardised systems and tools* several interesting studies and other contents were found that could be used for the model and software development of QUATRA.

3.1 Technical approaches

Technical approaches focus on the detection process itself (threshold value tests) and the analysis of certain data attributes as well as the detector communication and the combination of specific attributes.

In the US multiple detection algorithms have been developed on behalf of the Washington State Department of Transportation that detect erroneous detection data that is caused by short and long pulses from detectors, chattering detectors and intermittent malfunctions. These errors allow the identification of detector malfunctions and the existence of "abnormal" road conditions and incidents. The algorithms compare volumes and volume/occupancy ratios with threshold values to identify the reliability of detector data (based on a 20-second interval). Most of these issues are identified easily through data analysis whereas the problem of long pulses has been solved through the use of an automated video detection.[a]

Two other reasons for the wrong detection of vehicles have been identified in another research project: in case vehicles change lanes physically at the location of the detection field (this is called splashover) or in same vehicles are detected twice at the same site (e.g. heavy vehicles with trailers). To decide whether a vehicle is detected falsely at two detectors algorithms were implemented, which tested suspicious data to find out, whether there were actually two cars at two detectors at the same time, or whether it was a splashover effect. An approach was defined to correct these wrongly identified vehicles. Furthermore the study concluded the swapping of the road detectors with roadside units for a unique detection. [b]

Other algorithms that have been developed analyze the footprint of single vehicles on detector sites. On the one hand average vehicle lengths have been calculated from observed traffic data. Consequently vehicle distributions were estimated and compared with historical data.[c] On the other hand on-times (the time one vehicle is over the detector) were assessed. They represent the ratio of vehicle lengths over speed. As a result the estimated vehicle distributions were logically compared with similar vehicle distributions [d] and the average on-time for a time interval [e] respectively to identify faulty data.

Another research project was conducted on behalf of the Washington State Department of Transportation which compared storage rates for time intervals with the sum of storage rates over time. Consequently time variation curves were estimated and compared to identify potential faults in the detection process. The upstream ramp metering could be improved as well as loop "chattering" errors identified. [f]

---

[a] cf. Nihan / Wang (1995)

[b] cf. Coifman / Lee (2011)

[c] cf. Turochy / Smith (2000)

[d] cf. Coifman / Lee (2006)

[e] cf. Chen / May (1987)

[f] cf. Nihan (1997)

Sensitivity problems can be a further source for identification problems. In a research study a Gaussian Mixture Model (GMM) was used to identify locations with sensitivity problems. Correction factors were calculated to improve the quality of occupancy measurements.[g]

Apart from approaches that mainly focus on single vehicle based analysis other studies also engage errors on aggregated data levels. One of these studies engaged eight detector validation tests (unfortunately no further information could be retrieved so far for the state-of-the-art analysis). These tests were used to contrast the performance of different sensor models and were based on 24 hour data intervals. According to the study results the tests could also be modified to be applied to hourly traffic volumes.[h]

At last common problem that has been identified during field tests were the switching of directions of loop detectors that are connected in a joint cabinet. Although this issue should not occur due to common sense it still seems to be a realistic problem which has been solved with semiautomatic and automatic methods on the basis of correlations between measurements made by spatially close sensors. The proposed methods included a multi-dimensional scaling map of sensors that displayed similarities between different detection sites.[i]

3.2 Traffic engineering approaches

Apart from the technical tests a lot of engineering approaches have been published in the last years that deal with the quality analysis of traffic data. These tests mainly focus on traffic flow theory with the use of the inherent relationships among speed, volume, density and occupancy to assess data validity. Furthermore the principle "conservation of vehicles" is used for assessment which states that upstream incoming traffic volumes have to be identified further down the track again and that the traffic volumes can not be higher than the road capacity.

In a macroscopic point of view over a series of detectors the data can be adjusted using a constrained nonlinear optimization approach whenever the conservation principle is violated.[j]

In the US (California) an algorithm has been developed that detects erroneous traffic detection data by modelling the relationship between neighbouring detection sites. Linear regression is used to estimate the value of missing traffic data. Furthermore historical data is used for the evaluation. [k]

One paper specifically dealt with the detection of erroneous traffic data at signalised intersections in the US. Quality checks were carried out that were based on minimum and maximum flow thresholds as well as the principle of conservation of vehicles. In urban areas the conservation of flow proved to be difficult because of numerous intersections and side roads. Therefore additional field observations had to be carried out.[l]

A dissertation at the Technical University of Munich analyzed and assessed existing methods for traffic data detection on motorways. The proposed methods were divided into microscopic methods (which regard only the quality of local detectors) and macroscopic methods (which regard aggregated detector data of a whole network).

---

[g] cf. Corey et al. (2011)

[h] cf. Coifman / Dhoorjaty (2004)

[i] cf. Kwon et al. (2004)

[j] cf. Vanajakshi / Rilett (2006)

[k] cf. Chen et al. (2003)

[l] cf. Weijermars / Berkum (2006)

The method for detection of moderate coarse flow errors was based on the comparison of the measured flow values (over a certain measurement horizon) of three successive measurement cross sections (considered cross section and the one before and after). Hence the balance between the measurement cross section was compared and different balances were marked as faulty.

To detect the systematic measurement errors for flow values a similar approach like for detection of moderate coarse flow errors was applied. Systematic errors are generally only small errors. Therefore a bigger sample of traffic data has to be assessed to find the error and to achieve a higher sensitivity, at the same time a smaller amount of false alerts.

The following table provides an overview of the four methods which were developed and assessed in the work.

Table 1: Classification and properties of the developed methods for improving data quality

Type of error/ Measurement value	Moderate coarse measurement errors	Systematic measurement errors
Flow	Section related method Vehicle conservation and balance logical combination with statistic decision logic	Section related method Vehicle conservation and balance logical combination with statistic decision logic
Speed	Rule-based local methods speed level and speed jumps further estimations	Section-related method Travel times from v and q logical combination with statistic decision logic

Source: HOOPS, M. (2002): Methodology for quality management of aggregated data of a measurement system when operating traffic management systems, Dissertation on the chair of traffic engineering and control of Technical University of Munich (FGV-TUM)

The method for detection of moderate coarse speed errors was a local method that considered the speed of vehicles of the detection unit. The method contained three components

- observation of unexpected high jumps in sequent speed values
- observation of unexpected high relative deviation compared to the moving average of the standard deviation
- If the speed is unexpected small, it is tested whether the speed is due to a traffic incident or measurement error occurred. The assumption was made that traffic incident in contrast to a measurement error is usually observed on all lanes.

For the detection of systematic speed measurement errors a method was developed which is based on a correlation analysis between load curves of two consecutive detectors. The correlation analyses checks the similarity of flow data of the two consecutive detectors and tries to identify the time difference between similar flow pattern of the two detectors. The so found time difference between the two detectors could be interpreted as the travel time between the detectors. Under the assumption that the travel time of a considered and specified driver group stays stable, the travel time, calculated from the distance between the detectors and the measured speed, has to be similar to the travel time found by correlation analysis. Otherwise a measurement error could have occurred.

The method shows strong boundary conditions and was therefore quiet prone to errors. Additionally the methods worked better with a smaller measurement interval of the detectors. The different developed methods were tested on their sensitivity and reaction time. [m]

On behalf of the ASFINAG the project consortium's participant nast consulting carried out a study in 2008 that focused on the development of an automated statistical approach to evaluate erroneous traffic detection data. The first work package dealt with the analysis of historical traffic data - time variation curves and hourly traffic volumes - on designated motorways in Austria. In this regard total vehicles counts as well as separate vehicle categories were considered. Furthermore different detection systems (loop detectors, HV toll gantries, multifunctional scanners) have been compared. The outcome showed severe differences in terms of total number of registered vehicles for some of the locations that were the consequence of erroneous detection.

The second work package specified the automated mathematic-statistical approach which consisted of tools for the assessment of single detection sites as well as neighbouring sites. Decision rules were developed to determine the bandwidth of correct and erroneous data. This was done through estimation of the mathematical distribution for typical (historical) hourly traffic volumes from Monday to Friday (data sets). The statistical analysis estimated a poisson distribution for most of the hourly volumes. Based on a level of significance for each data set the analysis of detected traffic volumes allowed the differentiation between correct and erroneous data. For neighbouring sites the ratio of the data sets of all sites was used to define the mathematical distribution and their level of significance.

Main output of the study was that the approach can be used for an automated, simultaneous evaluation of the whole network as long as historical data is available for the estimation of distributions. The analysis of neighbouring sites could be identified as more powerful tool in case non-aggregated traffic detection volumes are chosen and typical information about the sites e.g. exit and entry points of side roads and speed limits are defined.[n]

A diploma thesis was carried out at the Technical University of Munich that tried to detect different malfunctions of detector loops which could occur especially for urban detectors. For that reason different plausibility tests were developed and tested. The urban detector mainly contained the traffic value occupancy (o) and flow (q). The following tests were applied:

- For the value occupancy and flow minimum and maximum thresholds were specified which could vary in different times of the day.
- The speed was calculated from q and o due to the correlation of the fundamental diagram. If the speed exceeded the speed limit at the specific detector, the dataset was marked as implausible
- The data was checked regarding recurrence
- A new fundamental diagram test was developed. A regression curve second order was placed in a fundamental diagram. Measurement data with a significant high distance to the regression curve were marked as suspect
- The coefficient of determination described how good the scatter plot was represented by the regression curve. If the coefficient of determination was small, there was only a small correlation between o and q und the detector was treated is possibly defect.[o]

[m] cf. Hoops (2002)

[n] cf. NAST CONSULTING, TU WIEN (2008)

[o] cf. FREUDENBERGER (2001)

3.3 Guidelines and standard procedures

Due to the importance of quality assured traffic detection data, guidelines and procedures have been developed worldwide. For the purposes of QUATRA selected documents were analysed in terms quality assurance. These guidelines and procedures focus on basic parameters that represent the quality level of traffic data: accuracy, completeness, validity, timeliness, coverage and accessibility. [p]

MCH 1529 - Guidance Notes for Assessment of Detector Technology/Systems [q]

The guideline MCH 1529 deals with the assessment of detectors and their corresponding technology prior to the use on UK highways. Objective of the guideline is the definition of a common method for performance based assessment and the possibility to compare the quality of different detector systems. The guideline however does not provide specific information how to conduct assessments.

The general process for certification of suppliers is based on a four stage approach. The assessment requirements for each stage need to be achieved prior to the next assessment stage. Successful suppliers are certified through a statement of performance that shows that specific targets, functions and accuracy were met. Specific requirements for a particular application and/or a particular class of vehicle or an individual function are included in the definition (new and existing requirements are defined by the Assessment Agency). The expected performance, installation and maintenance requirements are also part of the statement. The four certification stages are as followed:

a) Initial Feasibility Study: the Highway Agencies' requirements for assessment are analysed by the supplier who outlines how the product specifications (e.g. detection principles) will meet those

b) Off Road Assessment: in this stage the technology performance (e.g. error rate and confidence levels) is tested against an existing "Reference Detector" in a controlled environment to include different test conditions (based on the appropriate event tables from the European Standard ENV 13563) that would be difficult to create during an "On-Road" Assessment (e.g. dangerous tests, false detection of specific vehicles, power and communications failures).

c) On Road Assessment: the detector performance is examined under a full range of traffic (Free Flowing, Heavy traffic with flow breakdowns, Stop-start conditions, Varied classifications of vehicles) and environmental conditions likely to be experienced on the network. The performance of the Detector Technology is assessed based on a comparative analysis with the Reference Detectors and Ground Truth.

d) Operational Testing: in the final evaluation stage the integration with operational systems is tested: e.g. fault monitoring/reporting, identification of potential issues regarding system integration, detector performance in a multi carriageway

The description of the detector function comes in three parts: the target or targets whose parameter is to be detected, the parameter of the targets to be detected, the degree of accuracy that the supplier wishes to claim for that detection.

---

[p] cf. VERSAVEL (2007)

[q] cf. HIGHWAYS AGENCY (2011)

Each detector needs to be classified according to four main components:

- Detection Target: e.g. Cars, HGVs, Motorcycles, All road traffic
- Detection Function: e.g. Speed, Count, Length, Occupancy, Queuing, Incident Detection, Presence of Vehicles
- Detection Accuracy: individual requirements need to be identified for different functions and classes
- zones of Detection: the zones are differentiated between must/may detect and must not detect

Furthermore the supplier needs to present details of the range of detection, tolerance and exceptions.

Reference notes for quality requirements and safety of local traffic data collection (FGSV AK 3.5.20) [r]

In this paper the requirements for traffic data for the management of traffic control systems are stipulated and the requirements regarding scope of data, exactness of the measurements and completeness are specified. Furthermore, general fault scenarios for all types of data collecting are specified and the sources of error are listed separately. Besides, recommendations for fault finding and its correction are given. After the requirements for planning of measurement sites have been defined, control procedures for a stable and continuous quality control of traffic data are suggested. In conclusion, it is described which organizational course of action and management measurements have to be taken to ensure continuous quality control and fault finding.

Possible fault scenarios in all types of data collection:

- Wrong number of vehicles or no vehicle has been recorded
- Wrong classification of vehicles
- Wrong speed
- Repetition of interval data
- Static „getting stuck“
- „Fluttering“

Likewise chapter 6 „Definition of control methods to verify requirements“ is of special interest. Chapter 6.3.2 „Plausibility check in the sub centre“ describes the currently already applied plausibility checks in the sub centers. They include for example

- Plausibility check of short-term data
- Plausibility check by means of differentiating between inflowing and outflowing traffic
- Plausibility on the basis of checking the trusted value areas
- Analysis of systematic detector faults

[r] cf. FORSCHUNGSGESELLSCHAFT FÜR STRASSEN- UND VERKEHRSWESEN (2005)

QUANTIS - Quality Assessment and Assurance methodology for Traffic data and Information Services [s]

The aim of the Quantis project is to investigate the relationship between ITS service quality and benefits/costs, determine the optimum service quality in four European service cases (Bavarian Traffic Information Agency, road weather information related services in Finland, RDS - TMC Plus service in Austria, and dynamic journey time information and event management in the UK), identify levels of data quality providing optimal service quality and to give a recommendation for European guidelines for quality assurance of traffic data.

Furthermore the following contents are being addressed:

- key European ITS services
- data types (traffic data, traffic control data, weather data, event data and co-modal data) and sources required by the services
- guidelines specifying data quality requirements for the services in different operating environments
- quality assessment scheme including a definition of quality levels, methodology for the quality assessment and objects and the quantified quality parameters for services based on the data types from deliverable one which serve for the assessment
- exploration of the current data quality of the services in European countries
- Quality Assessment and Assurance methodology

The section "Guidelines for Data Quality of the Key European ITS Services" within the Quantis report provides guidelines for the development of quality assessment schemes, a summary of measurable quality parameters (quality indicators) and a comparison of already established quality aims and thresholds of different organisations. Quality parameters are:

- completeness
- availability
- veracity
- precision
- timeliness

Deliverable two of Quantis deals with the identification of quality requirements from users as well as the qualitative evaluation of benefits and costs in relation to service quality. Quality and quantity of information are key requirements for the effective implementation of robust local, national and especially cross-border ITS information and traffic management. Cross-border covers here both geographical and organisational borders. Any quality imbalance between parties undermines the basis for cross-border cooperation. The Quantis methodology is meant to evaluate and optimise data and services in terms of quality and costs in order to promote cross-border data exchange and service provision to support and deliver policy objectives.

[s] cf. AUSTRIATECH et al (2010)

The Quantis methodology can provide the end user with:

- A standardised way to evaluate the data quality provided by a service. The result is a grading of the service within one of four quality levels
- The ability to help understand how a service's quality compares with equivalent services in other European countries
- Guidance on the areas that could be improved to optimise the benefits of a service to a higher quality level
- Extra justification to apply resources to improve a service within a particular direction and a gauge the degree of improvement that will be made by applying those resources

The Quantis methodology was developed following research into ITS data quality, ITS service requirements, ITS user requirements and evaluation methodologies. Large elements of the quality objects and parameters used within the methodology are based on the ISO 21707 standard on "Data quality in ITS systems" which are as follows:

Error Probability (EP), Data Correctness (DC), Standard Deviation (SD), Data Skew (SK), Average Error (AE), Reliability (RL), Cross-Verified (CV), Validation Process (VP), Confidence, Data validity flag and Data validity

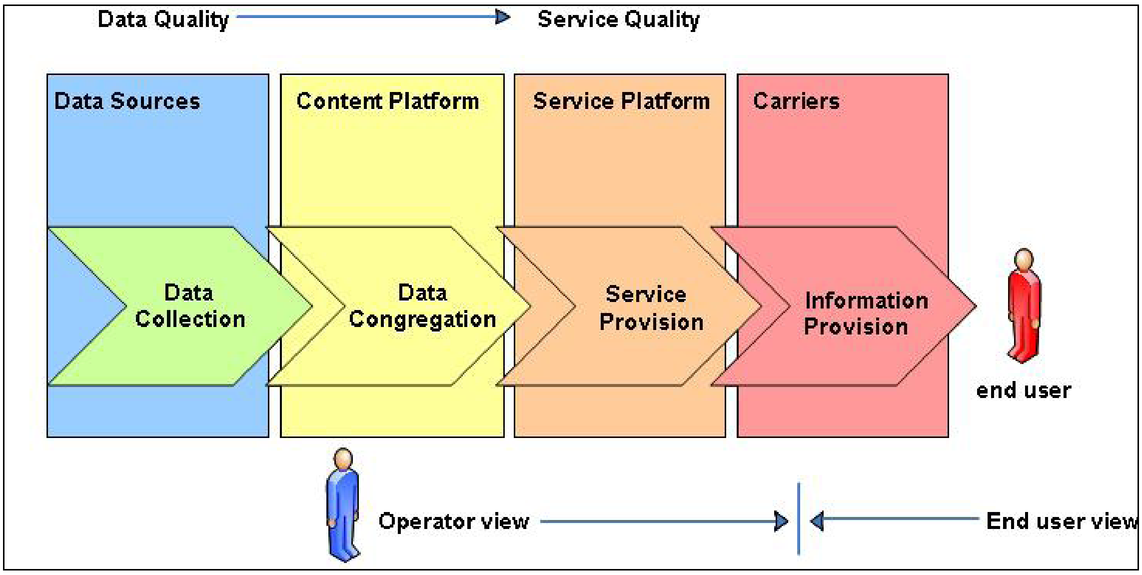

Figure 1: Quantis data processing chain

Source: AUSTRIATECH, PÖYRY INFRA TRAFFIC GMBH, BAVARIAN ROAD ADMINISTRATION, TECHNICAL RESEARCH CENTRE OF FINLAND, WSP GROUP (2010): "QUANTIS - Quality Assessment and Assurance methodology for Traffic data and Information Services"

Quantis recognises the typical data processing chain of a service (refer to Figure 1) and can be applied to the output of any of the three process involved within the chain. The processes can be defined as:

- Raw Data – The data provided directly from a measuring device or individual fields from a database
- Processing – The point at which different data sources may be combined to enhance the benefit of the data
- Service Provision – The point at which an end user is presented with the useful information. Note the end user perspective may also take into account a delivery medium (e.g. internet) which may not be an intended part of the review
- End User – User interprets information

As data progresses through the chain the quality or data is usually amplified. In this way inaccurate sensor readings at the raw data process stage will often have a greater effect on quality seen by the end user than faults after the processing of raw data. Quantis develops a methodology to evaluate and optimize data and service levels in order to foster cross border data exchange and service provision according to European policy objectives.

Quantis describes the need of good quality data, but not methods for finding errors at detectors. It also describes the importance of precise traffic data for end-users and which type of data they need.

3.4 Standardised systems and tools

In the field of traffic data quality management there are already standardised systems and tools in place in various countries worldwide that provide generous input to the development of the QUATRA system.

LOTRAN DQ [t]

LOTRAN-DQ is a software product of the participant TRANSVER of the project consortium which is an automated tool to control the quality of traffic data for motorways. It has been implemented for the first time in 2004 in the traffic control centre of the Autobahndirektion Südbayern in Munich. In advance, quality objectives, quality indicators and methods have been developed to their automatic calculation and visualization. The processes necessary to control the indicators and launch tool check-ups and corrections have been realized in practice. Since then the system has been applied for daily quality monitoring as well as a quality standard in acceptance of new sites.

Quality in the sense of availability, completeness and correctness of data is measured by means of the number and rate of various kinds of detector and cross section related errors. The following indicators are applied:

Local indicators for each single lane detector:

- Number or rate of missing data sets
- Number or rate of detector-specific error prompts
- Number or rate of data sets with all values equal 0
- Number or rate of implausible data sets

Global indicators for each cross section:

- Indicator for passenger car flow balance with upstream cross section
- Indicator for heavy goods vehicle flow balance with upstream cross section

The detection of implausible data sets is applied by the criteria of MARZ[u] and the supplementing criteria of the publication "Reference notes for quality requirement and quality control of local traffic data collection for traffic control systems" of the FGSV. An indicator for each individual criterion as well as a summing up indicator for implausible data sets is determined.

LOTRAN-DQ consists of a basic program, a data bank and a graphic user interface (refer to Figure 2). The basic system and the data bank are installed on a server. The user surface (GUI) can be installed optionally often by a client in the network of a traffic control centre or motorway management agency.

[t] cf. TRANSVER (2008)

[u] cf. MARZ 1999

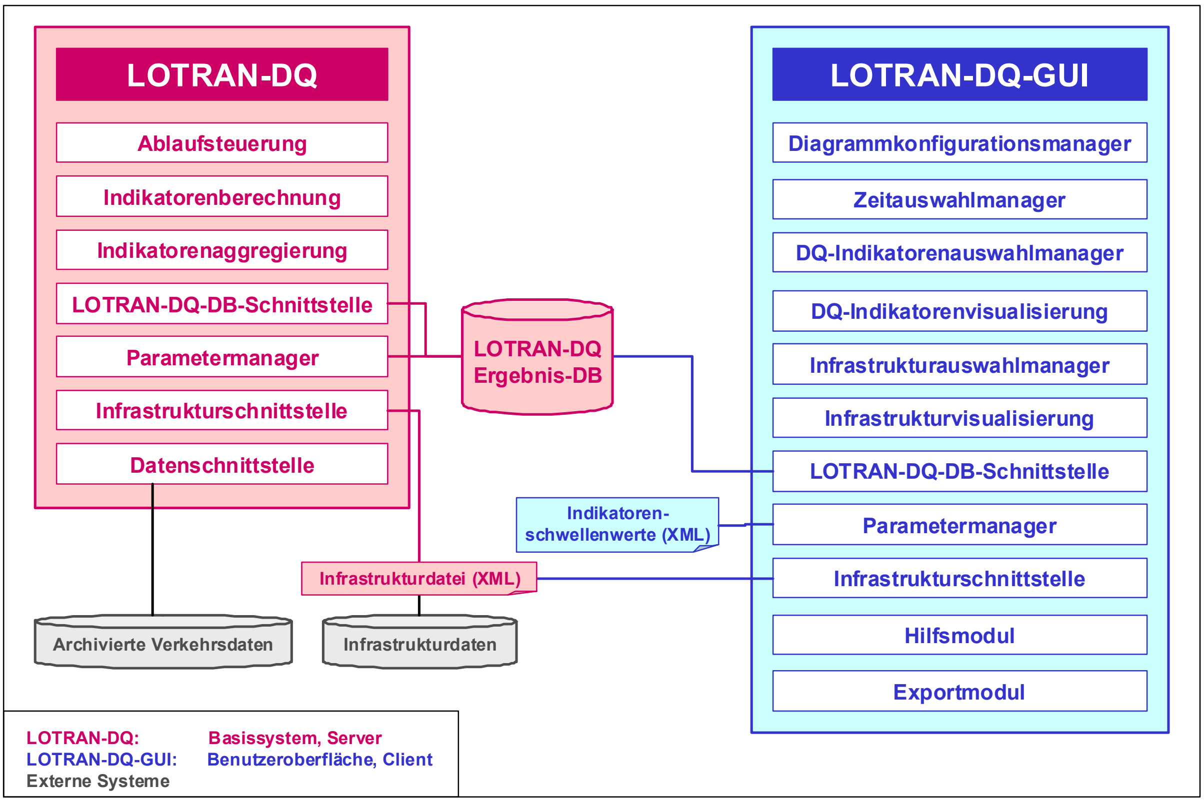

Figure 2: System architecture of the quality control tool LOTRAN-DQ

Source: TRANSVER (2008): "Quality control of traffic data in the federal highway network in Germany", LOTRAN DQ

The information about the measuring tools controlled by the traffic control centre is configurated manually by means of a XML data file or – if an electronic infrastructure data management is available – automatically. This builds the basis for the program to correlate between the cross sections necessary for the balancing. On the other hand LOTRAN-DQ sustains archived traffic data of the traffic control centre by means of a data interface to the archive.

The calculated quality indicators for past periods of time are saved in the LOTRAN-DQ data base. These results can be retrieved and visualized by means of the graphic user surface any time.

The current system is not capable for the online use and the application in urban areas.

TRAFFICIQ [v]

The pilot project Traffic IQ aims at providing an overall evaluation system for the quality of traffic data and information. The project is funded by the Federal Ministry of Economics and Technology (Bundesministerium für Wirtschaft und Technologie).

Various monitors (analysis tools) continuously evaluate different traffic data collection systems (video, mobile or stationary collection systems) according to their position (urban or non-urban). The evaluation is carried out with the different control systems developed in the project. On various levels of the data collection process (e.g. system components, detection process, generation of data) the recorded data is checked for hardware and software errors (e.g. failure of tools or data transfer components).

The detector data monitor checks the merely measured data and the model data monitor the “refined” data generated by the users, such as aggregated or smoothed data which contribute to models for evaluation of the traffic situation or control of traffic technology equipment.

Various key quality figures resulting from the testing procedures of the various monitors are fused for each monitor.

A plausibility monitor checks whether the results of the various collection systems for identical stretches match, or whether there are differences between the various collection systems.

The processing and analysis of the key quality figures is done on a data platform (so called data warehouse) which is also used for data storage. The data platform is able to generate reports based on locations, times, classes and events.

The defined project goals of Traffic IQ are as follows:

- Specification, realization and checking of testing procedures for the collection of stationary, mobile and video-based traffic data
- Specification, realization and checking of testing procedures for data and information which is generated on the basis of stationary, mobile and video-based detection by means of model-based procedures
- Specification, realization and checking of testing procedures for the information quality of stationary, mobile and video-based traffic data
- Definition of key quality figures to describe the traffic data quality of stationary, mobile and video-based traffic data
- Development of a multi-stage system of quality monitors to consistently evaluate and use quality values (especially for the users of traffic management centers)
- Realization of the developed testing procedures in a first implementation phase in chosen pilot areas (see test fields)
- Demonstration of the testing procedures after successful evaluation
- Realization of a “business intelligence component” which, on the basis of “data warehouse traffic data quality”, provides spatial and temporal aggregated information and trend forecasts on data quality (quality charts, reports, charts)

In order to check the testing procedures and the overall system directly regarding its efficiency in practice the system is tested and run in different test fields. Hence the testing procedures are directly checked on the basis of real data from different users.

---

[v] cf. MOMATEC (2012)

Each of the individual test fields are focused on a special testing procedure such as kind of detectors (stationary loop detecting, mobile data assessment, video data recording) and their position (e.g. in town or out of town). The test fields are:

- Autobahndirektion Nordbayern, Landesbetrieb Straßenbau Nordrhein-Westfalen, Baden-Württemberg: focus on stationary data assessment on highways
- Düsseldorf, Leipzig, Frankfurt/M.: focus on stationary data assessment in town or out of town
- Dresden, Nürnberg und Leverkusen (BAB 3): focus on video data assessment
- München BAB: focus on travel time assessment
- Frankreich: focus on mobile assessment

DAVINCI - Data Validation and Inspection for the Corporate Information Chain [w]

The DaVinci application improves the traffic data quality and contributes to reduce the costs of maintaining traffic systems. The approach includes the assessment of different kinds of errors such as:

- technical errors (faults in detectors,…)
- software and configuration errors (software bugs,…)
- errors caused by road works (incorrect installation,…)
- errors caused by road usage (traffic accidents,…)

Within the DaVinci application a method has been developed that amongst other filters false alarms caused by traffic events which involves the following steps:

- Uses validation models to automatically detect large deviations in data
- Conducts computer assisted analyses to find explanations for the deviations. Results of the analyses are stored in a knowledge database in order to help automating the trouble-shooting tasks
- Filters deviations caused by known traffic events; label the data to avoid new false alarms in later validations
- Identifies sources of the errors
- Requests the maintenance departments to remove the sources of errors
- Disqualifies the bad data in the historical database..

The DaVinci application uses three different validation models for detecting corrupt data.

1. The logical model

The logical model includes traffic engineering rules for automatic validation of traffic data. It triggers the validation rules every minute to determine the reliability of traffic sensors through:

- Validation of speed, traffic flows and travel times
- Indication of sources of errors based on previous events of the same type that were saved in the database
- Evaluation of the trustworthiness of the traffic detection data

2. Conservation of Flow model

The DaVinci project also considers the principle of "conservation of vehicles" as mentioned in chapter 3.2. The algorithm checks the consistency of observed flows on the traffic network. For example when compared with neighbouring sensors or data history, obvious deviations in the data can be identified automatically: an example in the paper shows that suspicious detectors in special circumstances (rush hour lanes) can be identified. Detailed analysis shows that the intensity outside rush hours are incorrect due to a software error introduced in a sub-system elsewhere.

---

[w] cf. CHAN (2008)

Rijkswaterstaat, Centre for Data and ICT, department of ICT Project Management

Schoemakerstraat 97b, P.O. Box 5023, 2600 GA, Delft, The Netherlands.

Consequently the Decision Support System would have suggested opening the rush-hour-lane at the wrong moment due to the heavy occupation (which was wrong)

3. OSCAR model

The OSCAR model detects “outliers” and “structural changes” simultaneously (refer to Figure 3). OSCAR is based on the Bayesian forecasting modelling technique. It can be effectively applied to intensity data of different temporal aggregate levels if a historical trend can be observed.

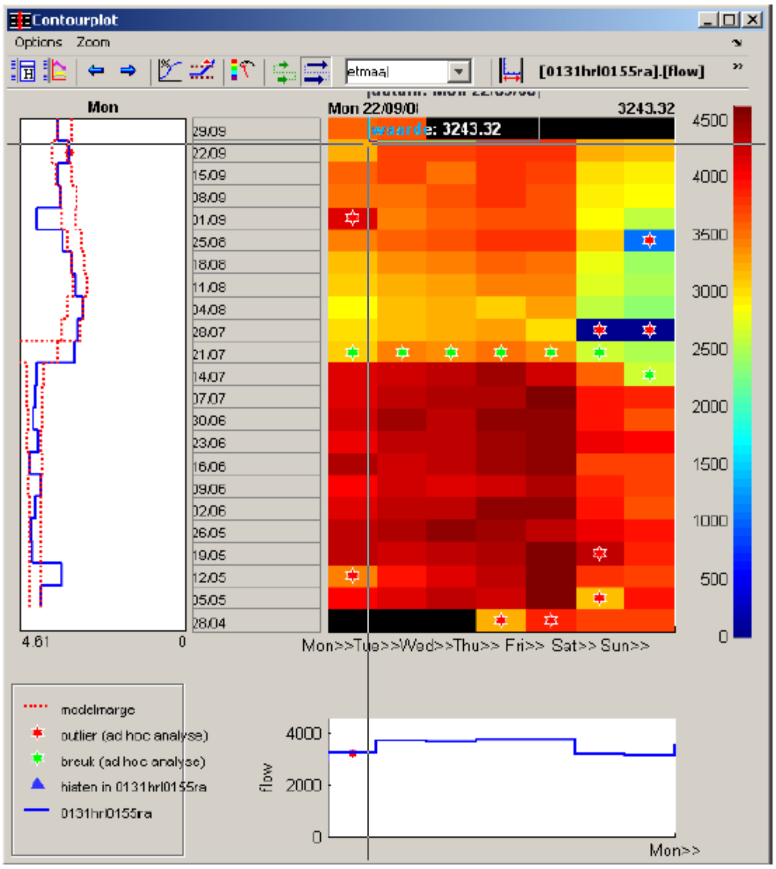

Figure 3: Detection of structural changes and outliers in the OSCAR model

Source: CHAN, K.F. (2008): “Leading traffic data to a high quality standard”, DaVinci, 16th ITS World Congress Final Paper

4. Analysis module

The DaVinci analysis module assists traffic engineers to quickly analyze the detected errors. It supports the user in analyzing the data of the entire information chain. DaVinci can significantly reduce the costs of application maintenance and development and making more valuable traffic engineering services become possible.

Each model has a method to detect a specific type of errors. By combining all three models, the analysis can trace a large range of corrupt data by different sources of errors.

Benchmarking for traffic data acquisition systems and traffic control systems [x]

A comprehensive benchmarking tool for traffic management systems was developed in Germany in 2006. The functions of the tool exceed the currently implemented plausibility checks in the control centre because all relevant system components, such as software and hardware of a traffic control system, are checked.

The developed tool ought to guarantee a permanent observation of the quality of the traffic control system. Therefore different quality checks on different system layers were developed. For each layer a monitor is introduced*:*

- Road unit monitor (checking availability and functional capability of the hardware)
- Data monitor (checking plausibility of the received traffic data)
- Modelling data monitor(checking plausibility of the modelled data which were calculated in the control centre for traffic control systems e.g. substitute values or the quality of a traffic control with false alarm on VMS)
- Acceptance monitor (checking whether the traffic control strategy, which was shown on the road, was accepted by the drivers)
- Effect monitor (checking whether the expected effect of the traffic control was fulfilled)

Up top 45 plausibility checks for the different monitors were implemented and tested. The results of the single plausibility checks are merged into one value per monitor (data fusion). In order to unify the quality standards for all system components, different service levels were introduced, which make the different units and methods comparable.

[x] cf. BUSCH et al. (2006)

4 Conclusions

The advantage of different technical approaches laid within the fact that different attributes and logic combinations are used for the identification of erroneous traffic detection data. Main benefit of the gathered information are numerous potential tests that could be integrated into the model development. Nevertheless the integration of these tests will also depend on the data that will be made available by the Austrian, Swiss and German infrastructure operators. The approaches already contained vital information for the freeway and urban model.

The QUATRA model will also test the incoming data in terms of traffic engineering fundamentals. The participants of the project consortium are experts in the field of traffic engineering fundamentals and therefore will look into converting the internal knowledge as well as the results from the approaches that were mentioned in the state-of-the-art analysis into the model. The approaches of Hoops (2002) and Freudenberger (2001) were very interesting in regards to the plausibility methods and hence the methods that were explained. Furthermore the project consortium will enhance the statistical approach from nast consulting in 2008 which will help to analyse especially the data in urban area networks.

In terms of the analysed guidelines and standard procedures the main interest lied within the bandwidth of the required quality information. In the MCH1529 the different test sites that need to be considered during the application of new traffic detection systems were of great interest. Hence the QUATRA model and software will be tested during different traffic conditions (e.g. "Free Flowing" traffic, "Heavy" traffic up to where flow breakdown occurs, "Queuing" traffic with stop-start conditions). Furthermore the need of reference systems for evaluation purposes was mentioned. The team of QUATRA will therefore try to collect reference data during the fault data detection process to compare the identified errors with the reference systems (to see the benefits in comparison to existing fault detection systems).

The content of the FGSV AK 3.5.20 reference notes is relevant for the urban data quality tool, which has to be developed in this project and hence ought to be integrated in the quality control logic. For the current project the possible fault scenarios and sources of error of the individual ways of data collecting described in the reference note are of interest. The procedures to be developed can be analyzed regarding the fault scenarios and hence the robustness and reliability of the methods can be tested.

The project Quantis has covered a wide area of interest. General quality objectives, parameters and thresholds are identified in the first deliverable. For QUATRA though these objectives and thresholds are not detailed enough. Nevertheless they help to define quality areas. Quantis recommends to fuse of multiple data sources to mitigate failed sensors. Consequently the quality parameters of Quantis and the ISO 21707 will be discussed in the model development phase and will also be used for the description of erroneous data. The paper also states that errors can occur within different data collection and gathering stages. The team of QUATRA will therefore look into assessing different levels of data collection.

For the current project the software LOTRAN-DQ will be taken into account. Modification of the initial quality indicators for motorways will be carried out. The current system is not capable for the online use and urban areas and needs to be extended. The software will be extended by further quality control procedures which are to be developed in the project.

In addition to the conclusions above the additional testing procedures in TrafficIQ (2012) and DaVinci (2010) are of interest. In TrafficIQ they are already structured according to the detector position within or outside of urban areas. Furthermore the different stages of data evaluation will be assessed for QUATRA. The self learning algorithm for incident detection within DaVinci (the indication of sources of errors based on previous events of the same type) is a very interesting approach which could maybe considered in a future stage of QUATRA. In general the approach of different monitor systems will be assessed for the model development.

Sources

AUSTRIATECH, PÖYRY INFRA TRAFFIC GMBH, BAVARIAN ROAD ADMINISTRATION, TECHNICAL RESEARCH CENTRE OF FINLAND, WSP GROUP (2010): "QUANTIS - Quality Assessment and Assurance methodology for Traffic data and Information Services"

BALMBERGER, M. / BALZ, W. / SCHEER, T. (2000): „Überprüfung von Dauerzählstellen im Autobahnnetz der Autobahndirektion Nordbayern“, Straßenverkehrstechnik Heft 12/2000, S. 648ff

BUSCH F. et al. (2006): "Benchmarking for traffic data acquisition systems and traffic control systems", Scientific journal series research road building and traffic engineering, book 949, Bonn

CHAN, K.F. (2008): “Leading traffic data to a high quality standard”, DaVinci, 16th ITS World Congress Final Paper

CHEN / MAY (1987): “Traffic detector errors and diagnostics", Transportation Research Record No. 1132, Freeway Management and Operations. p. 82-93

CHEN, C. ET AL. (2003): “Detecting errors and imputing missing data for single-loop surveillance systems”, in: Transportation Research Record, Issue: 1855, p. 160-167

COIFMAN, B. / LEE, H. (2006): "A Single Loop Detector Diagnostic: Mode on-Time Test”, Applications of Advanced Technology in Transportation, Proceedings of the 9th International Conference, American Society of civil engineers, p.623-628

COIFMAN, B. / DHOORJATY, S. (2004): "Event data-based traffic detector validation test", Journal of Transportation Engineering, 130/3, p.313-321

COIFMAN, B. / LEE, H. (2011): "Diagnosing Chronic Errors in Freeway Loop Detectors from Existing Field Hardware“, UC Berkeley Transportation Library, Retrieved from: http://www.dot.ca.gov/new..._tid_0978_final_report.pdf (01/2012)

COREY ET AL. (2011): "Detection and Correction of Inductive Loop Detector Sensitivity Errors Using Gaussian Mixture Models“, Conference: Transportation Research Board 90th Annual Meeting

FORSCHUNGSGESELLSCHAFT FÜR STRASSEN- UND VERKEHRSWESEN (2005): "AK 3.5.20 - Reference notes for quality requirements and safety of local traffic data collection“, Partly retrieved from: http://www.fgsv-verlag.de/catalog/product_info.php?products_id=2219 (01/2012)

FREUDENBERGER, P. (2001): "Analyses of detector loop data und development of a method for plausibility checks”, Diploma thesis on the chair of traffic engineering and control of Technical University of Munich (FGV-TUM)

HIGHWAYS AGENCY (2011): "MCH1529 - Guidance Notes for Assessment of Detector Technology/Systems"

HOOPS, M. (2002): "Methodology for quality management of aggregated data of a measurement system when operating traffic management systems", Dissertation on the chair of traffic engineering and control of Technical University of Munich (FGV-TUM)

KWON, J. ET AL. (2004): "Statistical methods for detecting spatial configuration errors in traffic surveillance sensors", Transportation Research Board, Issue: 1870, p. 124-132

MARZ 99 - Bundesanstalt für Straßenwesen (1999): "Information sheet for the equipment of traffic control centres and sub centres"

MOMATEC GMBH (2012): "Traffic IQ - Pilot project Informationsqualität im Verkehrswesen"

NAST CONSULTING, TU WIEN - Institut für Wahrscheinlichtkeitstheorie (2008): "Machbarkeitsstudie Qualitätskontrolle für Verkehrsdetektoren im ASFINAG-Netz", on behalf of the ASFINAG

NIHAN, N. L. ET AL. (1990): "Detector Data Validity", Report No.: WA-RD 208.1

NIHAN, N.L. (1997): "Aid to Determining Freeway Metering Rates and Detecting Loop Errors", Journal of Transportation Engineering Vol. 123 pp. 454-458; Washington State

NIHAN / WANG (1995): "Improved error detection using prediction techniques and video imaging - final report", Report No: WA-RD 386.1,TNW 95-01

NIHAN / WANG (2000): „Freeway traffic speed estimation using single loop outputs", Department of Civil Engineering University of Washington, Retrieved from: http://faculty.washington.edu/yinhai/wangpublication_files/TRB_00_SP.pdf# (01/2012)

TRANSVER (2008): "Quality control of traffic data in the federal highway network in Germany", LOTRAN DQ

TUROCHY / SMITH (2000): "New procedure for detector data screening in traffic management systems", Transportation Research Record, Issue Number: 1727

VANAJAKSHI, L.D. / RILETT, L.D. (2006): "System-Wide Data Quality Control of Inductance Loop Data using Nonlinear Optimization", Journal of Computing in Civil Engineering, 2006/5, p.187-196

VERSAVEL, J. (2007): "Traffic Data Collection: Quality Aspects of Video Detection", Transportation Research Board 86th Annual Meeting, TRB 86th Annual Meeting Compendium of Papers, Retrieved from: http://trid.trb.org/view/2007/C/801361 (01/2012)

WEIJERMARS, W. A. M. / VAN BERKUM, E.C. (2006): "Detection of Invalid Loop Detector Data in Urban Areas", Transportation Research Record: Journal of the Transportation Research Board, Issue: 1945, p. 82-88