DITCM

Dutch Integrated Testsite Cooperative Mobility

Data fusion Pilot - English summary of main results

October 2014

Smart Mobility Solutions

Data fusion Pilot – English Summary of main results

Date	October 2014
Author(s)	Isabel Wilmink, Diana Vonk Noordegraaf, Tom Thomas, Victor Westerwoudt and Manus Barten
Number of pages	11
Number of appendices	-

Table of contents

Summary

The Dutch National Data Warehouse for Traffic Information (NDW) has, together with DITCM Innovations, organised a precompetitive pilot to investigate whether making more extensive use of floating car data (FCD) and data fusion is a good approach to enable a reduction of the number of fixed location sensors (mainly loop detectors), and to assess the added value of FCD and data fusion for NDW's end users.
Three teams participated in the pilot (in-kind participation):

- Team **ARS T&TT** and HERE;
- Team **Be-Mobile**, Goudappel Coffeng, MAPtm and VORtech;
- Team **CGI**, SAS, Grontmij, Accenture, HP, TU Delft and TomTom.

A selection of main roads in and around the city of Delft (motorways and provincial road) served as test site. For several scenarios it was assessed which level of quality of road traffic data could be achieved using FCD, data from fixed location loop detectors and data fusion. The fused data was compared to the loop detector data. The focus was on speed and traffic volume indicators.

In an atmosphere of cooperation and openness, NDW/DITCM together with the teams have gained much insight into the application of FCD and data fusion. The main conclusions are (see chapter 2 for a more extensive set of conclusions and recommendations):

- The pilot showed that FCD and data fusion can be applied effectively to generate data about speeds. The use of FCD and data fusion for generating traffic volume data is less straightforward and we will continue to need high quality fixed location sensors for this, albeit (far) fewer than we use in the current situation.
- It appears that traffic patterns occurring during incidents and shock waves can be visualized quite well with fused data.
- It is recommended to judge the quality of road traffic data based on FCD and data fusion in a different way than with the current NDW quality criteria, because these are minute-based, and small deviations in time can cause large differences between the fused data and loop detector data (caused by the stochastic nature of traffic).
- The added value of FCD and data fusion is that the visualisation of these data give a more complete picture of traffic, one that also gives insight into the quality and plausibility of the data.
- The teams are confident that FCD and data fusion can be deployed soon.
- The precompetitive collaboration is worth repeating. Joint development by public and private partners in the pilot phase lays a good foundation for future collaborations. Industry partners will be more likely to participate actively if the future business case and associated vision of the future of NDW are clear. It is, for instance, important that NDW and its end users indicate for which purposes they will use certain sets of data and how they would like to deploy FCD and data fusion for this.

All involved have gained much insight in a short period of time in this pilot. The teams can use these insights to determine their future market position. Note that their participation in the pilot has no impact on their position in a future tender

process (level playing field). NDW will use the insights in shaping the new tender for contracts with its data suppliers.

Terms and abbreviations

Term	Beschrijving
DITCM	Dutch ITS Test site for Cooperative Mobility
FCD	Floating Car Data
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HMI	Human Machine Interface
ITS	Intelligent Transport System
NDW	Nationale Databank Wegverkeersgegevens (National Data Warehouse for Traffic Information)
OBU	On-Board Unit
RWS	Rijkswaterstaat (Dutch national road authority)

1 Introduction

1.1 Motivation for the data fusion pilot

The Dutch National Data Warehouse for Traffic Information (NDW) brings together traffic data from various sources, in one database, and delivers these data to road authorities and service providers (see www.ndw.nu). NDW expects that Floating Car Data (FCD) will play an important role in this process in the future. NDW wishes to use more FCD and plans a tender for a data fusion application. A number of the contracts NDW currently has with data suppliers will end in the near future. NDW has, together with DITCM Innovations (see www.ditcm.eu), organised a precompetitive pilot as part of the preparation of a new tender. Several organisations expressed their interest and have formed teams to participate in this pilot.

NDW wants to use the pilot to investigate whether making more extensive use of floating car data (FCD) and data fusion is a good approach to enable a reduction of the number of fixed location sensors (mainly loop detectors), and to assess the added value of FCD and data fusion for NDW's end users. Several indicators were to be assessed to see what level of quality can be achieved.

1.2 Goal of the data fusion pilot

The goal of the pilot was:
To gauge which level of quality of road traffic data (spot speed, volume, and travel time) is feasible when using FCD, data from fixed location sensors and data fusion, assuming that the spacing of fixed location sensors is increased.

Ideally, NDW would like to obtain data of the same quality as is delivered by the current data collection system. Only after completion of the pilot can NDW indicate what consequences insights from the pilots will have for the way future data collection will be tendered.

1.3 Research questions and scope

1.3.1 Scope

Given the short time period in which the pilot had to be carried out, and the in-kind contribution of the teams, the scope of the pilot was limited and a small set of scenarios covering relevant situations was defined.
The following points of departure were formulated:

- Motorways as well as the underlying road network were considered.
- The focus is on speed and traffic volume, as data fusion was thought to have the most added value for those data when considering a reduction of the number of fixed location sensors. More is known already about the use of FCD to generate travel time data.
- Several teams participated, in kind, in the pilot, showing what they can do with FCD and data fusion. In an atmosphere of cooperation and openness many new insights were gained.

- A comparison was made of fused data with loop detector data. The loop detector data were the reference case, but we are aware that loop detectors do not provide the ground truth.

1.3.2 Research questions

The pilot aimed to answer the following main research question:

Which level of quality of road traffic data (spot speed, volume, and travel time) can be achieved by using FCD, data from fixed location sensors and data fusion, assuming that the spacing of fixed location sensors is increased?

Accompanying research questions were:

- What is the added value of FCD and/or data fusion?
- Can the solutions investigated be scaled up (widely deployed) easily?
- Which are the lessons learned from the pilot?

1.4 Approach

In the pilot, a collaborative approach was used with respect to both content and process. The main components were:

- A project kick-off with many partners at Intertraffic.
- A more content oriented kick-off with the teams that had registered to participate in the pilot.
- Telephone interviews and site visits conducted by impartial NDW/DITCM experts to discuss the teams' results.
- A meeting with the teams to discuss the final results.
- Data analyses and reporting by the teams and NDW/DITCM.
- An assessment of the process with the teams.
- A symposium during which the results of the pilots were presented to a wide audience. This covered the results made public; the teams decided which of the results they wanted to share.

1.5 Collaboration between NDW, DITCM and the teams

Three teams participated in the pilot (in brackets the name of the leading partner):

- Team **ARS T&TT** and HERE;
- Team **Be-Mobile**, Goudappel Coffeng, MAPtm and VORtech;
- Team **CGI**, SAS, Grontmij, Accenture, HP, TU Delft and TomTom.

Three other teams have participated in the pilot, but pulled out before the end of the pilot and did not provide results:

- Team SWARCO and TomTom;
- Team VID;
- Team Imtech and TomTom.

NDW and DITCM supervised the pilot and provided experts that the teams could spar with.

1.6 Scenarios

De basic scenarios given to the teams were the following:

Traffic volume “any vehicle” and spot speed scenarios:

- Regular traffic (working days 11:00 – 13:00h)
  - Scenario 2.1: motorway, remove 1 loop detector location without exchange.
  - Scenario 2.2: motorway, remove all loop detector locations minus 1 location between two junctions.
  - Scenario 2.3: motorway, remove all loop detector locations between two junctions.
  - Scenario 2.4: motorway, minimal scenario.
  - Scenario 2.5: provincial road, remove 1 of 2 loop detector locations.
- Congested traffic (working days 16:00 – 18:00h)
  - Scenario 2.6: motorway, remove 1 loop detector location without exchange.
  - Scenario 2.7: motorway, remove 50% of loop detector locations.
  - Scenario 2.8 :motorway, minimal scenario
  - Scenario 2.9: provincial road, remove 1 of 2 loop detector locations.

NDW has provided the teams with a dataset and instructions as to which loop detector locations had to be removed.

Teams could, in addition to these basic scenarios, evaluate several other predefined scenarios:

Traffic volume “any vehicle” and spot speed scenarios:

- Low traffic volumes (working days 24:00 – 2:00h)
  - Scenario 2.10 : motorway, remove 1 loop detector location without exchange.
  - Scenario 2.11 : motorway, remove 50% of loop detector locations.
  - Scenario 2.12 : motorway, minimal scenario
  - Scenario 2.13 : provincial road, remove 1 of 2 loop detector locations.

Travel time scenarios:

- Regular traffic (working days 11:00 – 13:00h)
  - Scenario 1.1: motorway, remove 50% of measurement sections.
- Congested traffic (working days 16:00 – 18:00)
  - Scenario 1.2: motorway, remove 50% of measurement sections.
- Low traffic volumes (working days 24:00 – 2:00)
  - Scenario 1.3: motorway, remove 50% of measurement sections.

Vehicle category (light, medium and heavy) traffic volume scenarios:

- Regular traffic (working days 11:00 – 13:00h)
  - Scenario 3.1: motorway, remove all loop detector locations minus 1 location between two junctions.
  - Scenario 3.2: motorway, minimal scenario.

2 Conclusions and recommendations

The insights that resulted from carrying out the data fusion pilot have been summarised in a set of conclusions and recommendations, as given below.

Market readiness of FCD and data fusion
One of the questions with which we started the data fusion pilot was whether the industry is ready to deliver floating car data and data fusion to NDW. This question has two parts:

1. Can FCD and data fusion be delivered with the desired level of quality, and
2. What do the participating teams consider feasible in the short term.

From the comparison with loop detector data it could not be concluded that replacing loop data with other data (such as FCD/fused data) is feasible while maintaining the level of quality as required by NDW – if the loop data are used as a reference case, the accuracy and reliability of the data obtained with FCD and data fusion are too low in many of the scenarios evaluated. It has to be noted, however, that the comparison with loop data does not give a complete picture, because there is noise in the loop detector data as well (i.e. it is not the “ground truth”). The participating teams are confident that FCD and data fusion can be deployed soon. It seems therefore useful to adapt the quality criteria currently used by NDW, in order to (a) enable quality checks of FCD/fused data and (b) have more diversified quality criteria that better match the goal for which the data are eventually used (policy monitoring, traffic information, evaluation of measures etc.). This can have consequences for the temporal resolution of the data storage (not necessarily per minute anymore).

FCD and data fusion appear to have added value in specific situations (such as incidents and shock waves), and can probably also deliver good information for policy monitoring purposes, but it is not a fully formed product yet. To actually implement it, a period of further development of about six months is likely needed to meet all wishes of NDW’s users. Whether the fused data are good enough, cannot be concluded from the results of the pilot alone. The decision about this depends on what exactly will be tendered (in terms of data supply contracts).

The pilot provides little insight into whether the use of FCD and data fusion can be done with little delay – or: whether data fusion (with FCD) can be deployed for real-time traffic information applications. There need to be requirements for data latency for this type of applications.

Quality of traffic data generated with FCD and data fusion
When the data generated with FCD and data fusion are compared to loop detector data, the NDW quality criteria can be met for spot speed data outside peak hours. Traffic patterns can generally be reproduced well, as visualisations show. It is recommended to judge the quality of traffic volume and speed data generated with FCD and data fusion in a different way than with the current NDW quality criteria, by, for instance, looking at distributions of speeds and traffic volume (and to compare those to distributions derived from loop detector data). A quality indicator expressed in just one number does not do justice to data generated with FCD and data fusion.

Factors influencing data quality
From visualisations of fused data it has become clear that the granularity of FCD is important in specific situations (such as congestion caused by incidents, and shockwaves) – a higher resolution in space and time gives better insight into what exactly is happening in such situations.

An important issue for the tender is how to determine the quality of FCD. The pilot showed that this entails more than just penetration rate – the source of the data, and certainly also how the data are processed, are important.

Added value of FCD and data fusion
Visualisations of fused data clearly show the added value of (disaggregated) FCD and data fusion, where it concerns for example incident congestion and shockwaves. The visualisations give a more complete picture of traffic flow conditions (there is also information for the sections between the loop detectors). This type of visualisations gives more insight into the quality and plausibility of the data. There also seem to be opportunities for the underlying road network (urban and rural roads), for which there are much less loop detector data.
For judging the potential of data fusion (also in the context of the tender), it is important to not just check the data quality at cross sections. Data for road sections or routes should be considered too – see also the remark above about generating a more complete image of traffic[1].

Data fusion also has potential as a tool for continuous plausibility checks, because a coherent picture is formed from the various data sources. This offers, on the one hand, the possibility to give a higher quality score to estimates based on data sources all pointing in the same direction; on the other hand there is the possibility to identify a data source that systematically deviates from the 'complete' (fused) picture.

Scalability
The methods applied appear to be suitable for large scale deployment. Note that this may not be the case when traffic models are used, as this makes the methods more complex. It is expected, however, that with the current capabilities of computers this complexity can be handled.

Process
All parties involved have gained many insights, both content-wise and process-wise. All involved in the pilot worked well together. The interactions with and between the teams were pleasant and there was a large amount of goodwill. Participants indicated that working together in a precompetitive setting is something they would like to do again. Also, the teams see opportunities for further collaborations in the future.

[1] This brings with it that another way of storing the data is needed.