An integrated algorithm for fusing travel times, local speed and flow

Qing Ou
Transport & Planning
Delft University of Technology
Delft, Netherlands.
Q.Ou@tudelft.nl

J.W.C. van Lint
Transport & Planning
Delft University of Technology
Delft, Netherlands.
J.W.C.vanLint@tudelft.nl

S.P. Hoogendoorn
Transport & Planning
Delft University of Technology
Delft, Netherlands.
S.P.Hoogendoorn@tudelft.nl

Abstract – Traffic information plays an important role in traffic management and control. In order to better estimate traffic states from varieties of traffic sensors, traffic data fusion techniques are employed. This paper proposes an algorithm which can fuse travel time information and local loop data. Travel times from in-car GPS or cameras can provide average journey speeds over a few road segments but fails to provide the traffic details on each segment. Loop detectors can provide local traffic information, but there are biased errors in speed measures and the error in vehicle counts is accumulated over the time. The proposed algorithm exploits the strength of each type of data and avoids their weakness. In contrast to the previously often-used traffic fusion techniques, this algorithm needs very few assumptions on traffic behavior but can fuse more types of data. The validation shows this fusion algorithm can improve the estimation accuracy by up to 10 times.

Keywords: traffic data fusion, loop data, travel time.

1 Introduction

Effective (dynamic) traffic management (DTM) strongly depends on the reliability and accuracy with which one can deduce the current state of the traffic system from different types of sensors that are installed to monitor the network. However, data from a multitude of different sensors do not necessarily mount up to coherent and meaningful information on the state in a traffic network. Data from different sensors, such as cameras, induction loops or in-car GPS/GSM devices are typically characterized by different formats, semantics, temporal and spatial resolution and accuracy. They also greatly differ in availability and reliability both as a function of location, time and circumstances [1, 2]. In many fields of science, such as robotics, medical diagnosis, image processing, air traffic control, remote sensing and ocean surveillance (see e.g. [1, 2, 3, 4, 5, 6, 7, 8, 9]), the most common method for state-estimation is multi-sensor data fusion. Data fusion is a technique by which data from several sensors are combined to provide comprehensive and accurate information. The available data fusion algorithms (see [10] for an overview) range from low level (sensor level) to high level inference (at network level). They provide users with increasingly aggregated and comprehensive information on the state of a system. Using similar arguments as in [11, 12]), data fusion is generally able to increase confidence and accuracy, reduce ambiguity, increase robustness, enhance spatial-temporal coverage, decrease costs, etc.

With these arguments in mind, data fusion is an obvious solution for dealing with data from different traffic sensors and translating these into meaningful information. In [14] the authors conclude among other things that fusing floating car data with loop data considerably improves both performance and robustness to missing data as compared to using loop data only. In [15] the authors fuse loop data (speeds and flow) by using macroscopic traffic model and extended kalman filters. In [16] the authors use macroscopic models to fuse vehicle trajectories and vehicle counts by loops. But still there is no effective way to fuse travel time, loop speeds and loop flow altogether and simply uses some simple physics law and assumptions. This paper proposes an algorithm to deal with the above issue. It is able to assimilate more types of data including travel time, loop speeds and flow. It uses simple laws and assumptions such as conservation law and assumption of homogeneous traffic in a small time-space region. These laws and assumptions are independent of data. However, the previously often-used methods usually employ macroscopic traffic models that need to be calibrated by data and has more assumptions. In addition, these methods barely fuse the travel time information with other local information e.g. loop data.

Next some basic concepts about traffic and data characteristics will be briefly given. Following that, the algorithm is presented and validated by the artificial data.
1.1 Basic relationship between traffic variables

There are three major variables in macroscopic traffic. They are density \((\rho)\), speed \((v)\) and flow \((q)\). The evolution of these traffic states can be compactly and efficiently visualized by means of time-space contour plots. These plots are discrete representations of traffic states by using discretized time-space cells and color (e.g. Figure 7). With contour plots the traffic states at the given time and location can be simply shown. In each cell, the traffic is assumed to be homogeneous and stationary. According to Edie’s definition [13], these three traffic variables have such relationship between one another as shown in the below equation:

\[ q = \rho v \tag{1} \]

where \(q\), \(\rho\) and \(v\) are all time-space mean quantities.

In addition, there is a physics law in macroscopic traffic, called Vehicle Conservation Law: “the change in vehicle number on a road segment equals to the net difference between inflowing vehicle number and outflowing vehicle number”. It reads

\[ \rho(i, j) = \rho(i - 1, j) + \frac{\Delta t}{L_j}(q(i, j - 1) - q(i, j)) \tag{2} \]

where \(i = 1, 2, 3...\) represents discrete time, \(j\) indicates the location, \(L_j\) is the length of the road segment \(j\) and \(\Delta t\) is the span of one discrete time.

1.2 traffic data sources

The most common data available for traffic state estimation come from (dual) loop detectors. In the Netherlands for example, the main highways have an inductive loop for about every 500 meters. These loop detectors can provide speed measures and flow measures at the exact locations where they are installed. These data are presented in aggregated values for a time period ranging from 30 seconds to 10 minutes.

These speed measures in a collection system are named time-mean speeds. Another type of speed measures are called space-mean speeds. The latter ones represent the journey speeds with which vehicles cover a certain road stretch. Seen from this, estimates for travel time are better computed from space mean speed. And it is also true for traffic density estimation. The following example shows the difference between time-mean speeds and space-mean ones.

Figure 2a shows a one-kilometer long ring road on which there are three cars running with constant speeds 10km/h, 20km/h and 30km/h respectively. The time-mean speed measure by the loop detectors is 20km/h, which leads to travel time of 3 minutes. However, the ground-truth travel time on average is 3.67 minutes, which can be derived from space-mean speeds.

It can be seen that there is structural bias in the speed measures by the loop detectors. This bias is significant, specifically under congested (low-speed) conditions. This bias has been demonstrated for example by [17] when estimating travel times (errors of over 30%), and by [18] when estimating densities, where the resulting errors can mount up to over 100%.

Apart from speed measures, loop detectors are able to count the number of passing vehicles during a certain interval. These counts lead to estimates of flow \(q\), the number of vehicles passing per unit time. Theoretically, the accumulated vehicle counts may tell travel time and vehicle density (or vehicle number) on a closed road section between two consecutive loop detectors as shown in Figure 2b. In this figure, \(N_a(t)\) and \(N_b(t)\) are the accumulated vehicle counts from loop A and loop B respectively. Given the initial condition that no vehicle is on the road section, \(N_a(t) - N_b(t)\) is the number of vehicle on this road section, and \(TT(t)\) satisfying \(N_a(t) = N_b(t + TT)\) can be taken as the travel time from loop A to B under the assumption of no overtaking. These estimates are reliable and accurate if loop detectors made no errors in counting vehicles.
However, correct estimates cannot be obtained in reality due to error accumulation. In the above example, the errors in \( N_v(t) \) and \( N_v(t) \) are accumulated with the time \( t \). As a result, it may be found that the total number vehicle in a road segment is minus thousands in the end.

In this paper, travel time is given a particular focus. Travel time can be measured by means of for example automated vehicle identification (AVI) systems, which identify vehicles at two consecutive locations A and B at time instants \( t_A \) and \( t_B \) and deduce the realized travel time afterwards with \( TT = t_B - t_A \). AVI systems may employ cameras and license plate detection technology, or may match vehicles through induction footprints, tolling tags or otherwise. It is worth note that this paper uses individual travel times instead of aggregated or average travel times. In contrast to other traffic information such as traffic flow, density and speed, travel times may be regarded as a kind of causal aggregation of traffic history information over realized travel space. Compared with loop data, travel times have a higher order of accuracy without structural bias. Although, travel times can be derived from time-space speed information, the reverse process is impossible. For this reason, it is quite a challenge to use travel times to estimate the local traffic states.

This paper proposed a new algorithm to fuse these three ingredients (flow, biased speed and travel times) to achieve reliable and more accurate estimates of traffic density and speeds without using advanced traffic models such as first-order traffic models, but only using the basic relationship between traffic variables as shown in Equation (1) and (2).

## 2 Methodology

### 2.1 Framework

The whole fusion algorithm consists of two parts. The first part fuses travel time and speed measures by loop detectors. The second part further fuses the flow measures. In the end, the density and speed over the whole time space region are achieved. (See Figure 3)

- In the first part, the vehicle trajectories are reconstructed on a time-space plot by combining individual travel times and the given speed measures. This part of algorithm comes from **PISCIT algorithm** [19]. The individual travel times are obtained from vehicle identification systems e.g. in-car GPS or cameras. The given speeds are measured by loop detectors which cause measurement bias due to time-mean aggregation. The reconstructed trajectories are able to remove the bias effects to some extent by satisfying the given travel times as constraints.

- In the second part, the traffic density and speed in the time-space cells where the trajectories pass are deduced by simply using these trajectories. Next, the flow information is used to further deduce the density in the other time-space cells by employing **Vehicle Conservation Law**. Assuming that the traffic is homogeneous and stationary in each time-space cell, the traffic speed in the all time-space cells also becomes available by applying \( v = q/\rho \) for each cell.

![Figure 3: Fusion framework](image)

### 2.2 Fusion part one

This part of the algorithm is able to reconstruct individual trajectories by combining the given travel times and previously observed (estimated) cell-speeds from loop detectors. Cameras or in-car GPS can provide the entry point for a vehicle, that is where and when the vehicle enters a road stretch. Also the exit point is given about where and when this vehicle leaves the road stretch. Any line which links the two points could be a trajectory for this vehicle. The algorithm presented below is able to find the most 'likely' trajectory with the help of the time-space speed information from loop detectors, even though there is considerable bias in these speed measures. The mechanism behind is quite simple. For a fixed road segment in a road stretch, longer travelled distance, more travel time; higher speed, less travel time.

For simple illustration, it is assumed a probe vehicle \( k \) entered road segment 1 at reporting time \( t_1^{(k)} \) and excited segment 6 at the next reporting time \( t_6^{(k)} \) (Refer to Figure 4). This trajectory reconstruction algorithm is made up of the below steps, the first four of which accomplish reconstruction on segment level while the last two on cell level. Table 1 lists the important symbols used below.

**STEP 1:** Get \( t_j^{(k)} \) and \( t_j^{(k)} \) from this previously-estimated trajectory as shown in Graph (a) in Figure 4...
STEP 2: Based on the given time-space speeds (biased), (re)-calculate the average speed \( \hat{\bar{v}}((\hat{t}_j^{(k)}, \hat{t}_j^{(k+1)}), j) \) over segment \( j \) during the time between \( \hat{t}_j^{(k)} \) and \( \hat{t}_j^{(k+1)} \).

STEP 3: Update \( \hat{t}_j^{(k)} \) and \( \hat{t}_j^{(k+1)} \) based on the average speed \( \hat{\bar{v}}((\hat{t}_j^{(k)}, \hat{t}_j^{(k+1)}), j) \). The updated \( \hat{t}_j^{(k+1)} \) can be obtained under the below restrain equations

\[
\hat{t}_j^{(k)} \propto \frac{L_j}{\hat{\bar{v}}((\hat{t}_j^{(k)}, \hat{t}_j^{(k+1)}), j)}
\]

\[
\sum_j \hat{t}_j^{(k)} = tt
\]

where \( tt \) is the given travel time for a vehicle over the whole stretch. After that, update \( \hat{t}_j^{(k+1)} \) based on \( \hat{t}_j^{(k)} \).

STEP 4: Repeat STEP 2 and STEP 3 until \( \hat{t}_j^{(k)} \) and \( \hat{t}_j^{(k+1)} \) converge to a specific extent. (Refer to Graph (b) in Figure 4)

STEP 5: Deduce \( t\hat{(k)}(i, j) \) from \( \hat{t}_j^{(k)} \) and \( \hat{t}_j^{(k+1)} \) and cell size. (Refer to Graph (c) in Figure 4)

STEP 6: Deduce \( \hat{s}^{(k)}(i, j) \) under the below equations. (Refer to Graph (c) in Figure 4)

\[
\hat{s}^{(k)}(i, j) \propto t\hat{(k)}(i, j) \ast \hat{\bar{v}}(i, j)/\hat{\bar{v}}((\hat{t}_j^{(k)}, \hat{t}_j^{(k+1)}), j)
\]

\[
\sum_i \hat{s}^{(k)}(i, j) = L_j
\]

### 2.3 Fusion part two

In fusion part two, the reconstruction of trajectories returns the traffic speed and density where they pass through, and then traffic density and speed over the whole time space region are deduced by fusing flow information (Refer to Figure 5).

Assuming trajectory \( k \) passes the cell \( (i, j) \), the measured (estimated) density and speed by the trajectory at this cell can be obtained with the below equations:

\[
\hat{v}^{(k)}(i, j) = \frac{\hat{\bar{v}}^{(k)}(i, j)}{t\hat{(k)}(i, j)}
\]

Assuming the traffic is homogeneous in each time-space cell, it reads

\[
\hat{\rho}^{(k)}(i, j) = \frac{\hat{q}^{(k)}(i, j)}{\hat{v}^{(k)}(i, j)}
\]
In order to distinguish the final estimation \( \hat{v}(i,j) \), we simply call \( \hat{v}(k)(i,j) \) ‘measured’ speed by trajectory \( k \), though it is actually deduced from the reconstructed trajectory. Similarly for density and flow.

For a fixed road segment \( j \), Vehicle Conservation Law leads to below equations

\[
0 = \rho(i-1,j) - \rho(i,j) + \Delta t L_j (q(i,j-1) - q(i,j)) \tag{9}
\]

where \( i = 1, 2, 3... \) represents discrete time.

The measured density by trajectories leads to measurement equations for density:

\[
\hat{\rho}(k)(i,j) = \rho(i,j) \quad i = 1, 2, 3... \tag{10}
\]

The measured flow by loop detectors leads to measurement equations for flow:

\[
\hat{q}^- (i,j) = q(i,j) \quad i = 1, 2, 3... \tag{11}
\]

A regression model can be easily established by combining these three sets of formula (9),(10) and (11)

\[
y = Ax \tag{12}
\]

where \( y \) contains measures \( 0, \hat{\rho}(k)(i,j) \) and \( \hat{q}^- (i,j) \) for a fixed \( j \), and \( x \) contains estimated states \( \rho(i,j) \) and \( q(i,j) \). The optimal estimation of \( x \) is

\[
\hat{x} = (A^T A)^{-1} A^T y \tag{13}
\]

Now we have density, flow and speed estimates over the whole time-space region.

3 Validation

In the first part, the synthetic ‘ground-truth’ data are generated by assuming the real loop data are true, and then the observed data are generated by tampering the ‘ground truth’ data. In the second part, the proposed algorithm is applied on the observed data and returns the estimated data. The performance of this algorithm is shown by comparing the ‘ground-truth’ and estimated results.

3.1 Experiment setup & data generation

First of all, a 9.5 kilometer stretch of 3-lane Highway A4 eastbound in Netherlands is considered (Graph (a) in Figure 6), where 18 loop detectors are placed spacing around 500 meters and aggregated traffic speed measures and counts every one minute.

![Figure 6: Illustration of the study road and how the ground-truth data are tampered](image)

![Figure 7: Ground-truth time-space speed plot](image)
• **Ground-truth speed** We assume the loop detectors give the ground-truth speed measures over certain segments. The resulting time-space speed contour plots (Figure 7) shows 5 hour traffic condition on this stretch from 6:00 A.M. till 11:00 A.M. on July 8th in 2008, during which congestions onset and dispersed twice.

• **Ground-truth density & flow** The 'ground-truth' density and flow are generated by using the 'ground-truth' speeds and loop counts as boundary condition. The generated data satisfy Conservation Law and the homogeneous condition $\rho v = q$.

![Figure 7: Ground-truth time-space density plot](image1)

**Ground-truth speed**

![Figure 8: Ground-truth time-space density plot](image2)

![Figure 9: Comparison between observed speeds and estimated speeds after applying the proposed fusion algorithm](image3)

![Figure 10: Comparison between observed density and estimated density after applying the proposed fusion algorithm](image4)

• **Observed speeds** The observed speeds in each time-space cell are assumed by tampering the ground-truth speeds with the below assumption:

\[
v^o = e^{1.1v^g(0.5 - 0.5v^g/120)}
\]

, where $v^o$ is the observed speed and $v^g$ is the ground-truth speed. With this assumption, the observed speed is 10% higher when ground-truth speed is 120km/h, and 70% higher at the speed of 20km/h. The resulting observed time-space speeds are shown in Graph (a) in Figure 9. The relationship between them is show in Graph (b) in Figure 6.

• **Observed flow & density** The observed flow is assumed to equal to the ground-truth ones. The observed density is actually estimated by using $\rho = q/v$ (Refer to Graph (a) in Figure 10).

• **Travel times** The travel times are generated by sampling the 'ground-truth' time-space speed plots. There are three virtual cameras placed at the entry, exit and middle of the whole road stretch. It is assumed that 10% of vehicles are captured by the cameras, giving the travel times from milepost 0km to 4.8km and others from 4.8km to 9.5 km.

![Figure 11: Comparison of travel time](image5)

**Observed speeds**

**3.2 Results**

We use mean absolute relative error (MARE) to evaluate the results. The definition of MARE is shown in Equation 16.

\[
MARE = \frac{1}{M \times N} \sum_{i}^{M} \sum_{j}^{N} \frac{|\hat{x}(i,j) - x(i,j)|}{x(i,j)}
\]

, where $\hat{x}(i,j)$ represents the estimate and $x(i,j)$ represents the ground-truth quantity. The comparison of the results before and after using the algorithm can be seen in Figure 9 and Figure 10.

• **Before** The observed speeds and density have large errors. MARE for the observed speeds is 33.7% in the given scenario. And MARE for observed density is 24.12%.

• **After** After the proposed algorithm is applied to fuse the observed speeds, travel times and flow, the above errors remarkably decrease. MARE for estimated speeds becomes 3.3% and MARE for estimated density becomes 3.46%.

With time-space speed plots, travel times can be easily derived. Figure 11 makes comparison of travel time...
estimates between before and after using this algorithm. Before using it, the travel times based on observed speeds have mean absolute error of 230.4 seconds and MARE 26.89%. After using it, the travel times have a much smaller error of 6.6 seconds and MARE 0.74%. In Figure 11, the thick green line represent the ground-truth travel time, dark dashed line represents the results after using the algorithm and thin red line represents the travel time estimation based on the observed speeds. The former two lines almost overlap with each other.

Further more, the time cost for the computation is low. In the above case, the total time cost for data fusion processing is 18.52 seconds, in which about 11400 travel time records are processed.

4 Conclusion

This paper proposed a new algorithm for fusing speeds and flow from local detectors with individual travel times measured by AVI systems. In contrast to the previously often-used techniques for traffic data fusion, the proposed algorithm does not employ advanced traffic models, e.g. first-order traffic or second-order traffic model which includes strong assumptions on traffic behaviors and needs parameter calibration. It only needs basic and weak assumptions without need for parameter calibration, but it can fuse more types of data. In addition, this algorithm is quite computationally efficient, in which only simple iterations are needed.

On the basis of synthetic ‘ground-truth’ data, we demonstrated how this algorithm is able to successfully correct strongly biased prior speed measurements. It is able to improve the estimation accuracy up to ten times, e.g. decreasing MARE from 33% to 3.3%, decreasing errors in travel time estimation from 230 seconds to 6.6 seconds. The applications for PISCIT are manifold. First of all, in an offline context, it enables a simple but principled method to fuse data from local detectors and travel time data from AVI systems and hence improve the quality of archived datasets significantly. This is beneficial for a multitude of applications which depend on such historical data archives (e.g. simulation studies, performance analysis, policy evaluations, etc). In case the AVI data contains travel times over short distances, the algorithm might also have benefits for real-time ITS applications such as route guidance systems or ATIS.

Further research will focus on incorporating this algorithm in a more real context. We will study at least how the errors in the given travel times and flow impact the speed and density estimation with this algorithm.

5 Acknowledgment

This work is sponsored by the ICIS Data fusion project (http://www.icis.decis.nl/) and partially supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs under project DCB.7814 Reversing the Traffic Jam.

References


