Data Fusion

A simple and robust method for fusing heterogeneous data from different traffic sensors

December 16, 2008 – 30th CAITR, Perth
A/Prof Dr. Hans van Lint
ITS & Dynamic Traffic Management
The Netherlands ...

- Delft University of Technology
- Faculty of Civil Engineering
- Transport & Planning Department
- Dynamic Traffic Management Group
Mobility in The Netherlands

- 16.2 million inhabitants
- 3.1 trip per person per day
- 32 km per person per day
- 1 hour per person per day
- 6.9 million cars
- 16 million trips per car
- 250 million car km per day
- 120 congestion locations
- Average length 3.2 km
- Average duration 65 minutes
Congestion in The Netherlands

Dec, 10 2008 - 30th CAITR, Perth
Congestion in The Netherlands

% increase congestion (duration x length)

- 2003
- 2004
- 2005
- 2006
- 2007

0 2 4 6 8 10 12 14 16 18

TU Delft
Traffic Management begins and ends with good data.
Contents

- Some characteristics of (freeway) traffic data
- Main concept data fusion approach
- Examples and results
- Conclusions and applications
Some empirical facts about (freeway) traffic data

Direction of “information flow” in free conditions

Direction of “information flow” in congested conditions
Different traffic sensors measure different things

Local detection, e.g. loops
Different traffic sensors measure different things

‘floating car data’
Different traffic sensors measure different things

Non-Local detection, e.g. travel time
So ...

- Can you combine all these data ...
  - Which have different semantics
  - Different aggregations
  - Different quality, reliability, etc

- ... Into one coherent picture of the traffic conditions?
Approach: influence of measurement

\[ t_{\text{cong}} = t - \frac{x}{c_{\text{cong}}} \]

\[ t_{\text{free}} = t - \frac{x}{c_{\text{free}}} \]

\[ \varphi_{\text{cong}}(x - x_0, t - t_0) \]

\[ \varphi_{\text{free}}(x - x_0, t - t_0) \]
Multiple measurements

So how to distinguish between free and congested conditions?
Use measurement itself (e.g. speed/density)

Around 65-75 km/h
Overview approach

- Estimate of some quantity $z(t,x)$ (could be speed, flow, density) by
  - linear combination of estimates $z^{(j)}(t,x)$ on the basis of datasource $j$ with weights $\alpha^{(j)}(t,x)$ ...

$$z(t, x) = \frac{\sum_j \alpha^{(j)}(t, x) \phi^{(j)}(t, x) z^{(j)}(t, x)}{\sum_j \alpha^{(j)}(t, x) \phi^{(j)}(t, x)}$$

- ... and weights $\phi^{(j)}(t,x)$ which differentiate between free and congested

$$\phi^{(j)}(t, x) = w^{(j)}(t, x) \cdot \phi^{(j)}_{cong}(t, x) + (1 - w^{(j)}(t, x)) \cdot \phi^{(j)}_{free}(t, x)$$
Experiment with synthetic data

- A13 freeway (18 km)

- 3 types of data:
  - Induction loops every → 500m, 1000m, 1500m
  - Floating car data (TomTom, GPS) → 2, 5, 10% of all vehicles
  - AVI (realized travel times) every → 1500m, 3000m
Experiment with synthetic data

- A13 freeway (18 km)

Dec, 10 2008 - 30th CAITR, Perth
Results

data induction loops randomly remove 50% every 500 meter of the data.
Results

data induction loops every 1500 meter

Combine with 2% FCD (e.g. TomTom units)
Results

AVI stations every 3000m

Combine with loops every 1500m
Results – a few numbers

- Individual datasources
- Multiple datasources
Conclusions

- The filter data fusion approach is beneficial
  - Reduces both Bias and variance
  - Few loops and few % FCD is nearly as good as high % FCD alone, or densely spaced loops
  - In many cases: Data fusion > sum of parts

- Approach can be used for
  - Enhancing large historical data archives
  - Input for real time ITS applications
    - Traffic and travel time information
    - Ramp metering, intersection control, lane control, etc
    - Dynamic route navigation (in-car)
Thanks !

Dr. Ir. Hans van Lint

j.w.c.vanlint@tudelft.nl
www.transport.citg.tudelft.nl

Thanks to Prof Serge Hoogendoorn