FHWA Traffic Monitoring Approach. Innovative procedures in a study case

Massimiliano Gastaldi
Gregorio Gecchele
Riccardo Rossi

University of Padova
Department of Structural and Transportation Engineering – Transportation Laboratory
Introduction to FHWA Monitoring Factor Approach
- Main concepts
- Focus of the paper

Clustering Methods for Road Grouping
- Characteristics of Clustering Methods
- Methodology of Study
- Analysis of Case Study Results

Conclusions
- Concluding Remarks
- Future Developments
Introduction to FHWA Monitoring Factor Approach

- Main concepts
- Focus of the paper
• **FHWA Procedure. Main Concepts**
  – Monitoring Guide periodically updated by FHWA
    • Traffic Monitoring Guide (2001)
    • Supplement (April 2008)
  – Sample strategy with the objectives of:
    • Obtaining the maximum rate between coverage of estimates on network links and resources needed
    • Giving a prefixed level of confidence in the parameters’ estimates
  – Main Output:
    • Annual Average Daily Traffic (AADT)
• FHWA Procedure. Main Concepts
  – Why using a Monitoring System?
    • Mandatory activity for roads managing authorities
    • Knowledge of temporal evolution of road traffic flows
  – Main advantages of using Seasonal Factors
    • Knowledge of road traffic flows in periods not directly monitored with field observations
    • Combined use of a small number of Permanent Traffic Counters and Short-Period Traffic Counts
• **FHWA Procedure. Some Details**

  – **3 Monitoring Strategies**

    - Permanent Traffic Counts (PTCs)
    - Seasonal Traffic Counts (STCs)
    - Short-Period Traffic Counts (SPTCs)

  - PTCs with Automatic Traffic Recorders (ATRs, e.g. inductive loops)
  - STC e SPTC with Portable Counters (e.g. radar)
• FHWA Procedure. Step 1

• Creation of Road Groups based on Seasonal Factors Similarity
• FHWA Procedure. Step 1

• Section Factor
  – $k$ is the index of the section
  – $i$ is the index for the day, with $i = 1, \ldots, 7$
  – $j$ is the index for the month, with $j = 1, \ldots, 12$
  – $AADT_k$ is the Annual Average Daily Traffic for section $k$
  – $DT_{ijk}$ is the Daily Traffic for section $k$ for a day $i$ of month $j$
  – $ADT_{ijk}$ is the Average Daily Traffic for section $k$ for a day $i$ of month $j$

\[
 f_{ijk} = \frac{AADT_k}{\frac{1}{n} \sum_{i=1}^{n} DT_{ijk}} = \frac{AADT_k}{ADT_{ijk}}
\]

– Considering daily (7) and monthly (12) traffic patterns:
  • For each section 84 factors
• **FHWA Procedure. Step 1**
  
  • **Group Factor**
    
    – $k$ is the index of the section
    – $i$ is the index for the day, with $i = 1, \ldots, 7$
    – $j$ is the index for the month, with $j = 1, \ldots, 12$
    – $n_k$ is the number of permanent counters of group $k$

\[
f_{ij} = \frac{1}{n_k} \sum_{k=1}^{n_k} \frac{AADT_k}{ADT_{ijk}}
\]
Introduction to
FHWA Traffic Monitoring Factor Approach

• FHWA Procedure. Step 2

• Assignment of roads monitored for short periods to the most similar Road Groups
• FHWA Procedure. Step 2

• Assignment of roads monitored for short periods to the most similar Road Groups
Introduction to FHWA Traffic Monitoring Factor Approach

• FHWA Procedure. Step 3
  • Application of Group Factors
  • Estimation of AADT
FHWA Procedure. Step 3

\[ AADT_{\text{Estimate}} = DT_{\text{Sunday, March}} \cdot f_{\text{Sunday, March}} \]

\[ AADT_{\text{Sunday, March}} + DT_{\text{Sunday, March}} \]

Known

Measured

Introduction to FHWA Traffic Monitoring Factor Approach
• FHWA Procedure. Main Issues

• Main sources of errors in the procedure (Sharma et al., 2001)
  1. Sampling error in Daily Traffic choice:
     • Due to day-to-day traffic volumes variability.
  2. Error using group factors:
     • Instead of using the adjustment factors for the sample count site, groups factors are used.
     • This error is made also in the case of perfect assignment.
  3. Error in the assignment:
     • Due to the incorrect factor group assignment.
     • Errors in assignment could lead to a tripling of the estimation error (Davis, 1996).
• Focus of the paper

2. Error using group factors:
   • Instead of using the adjustment factors for the sample count site, groups factors are used.
   • This error is made also in the case of perfect assignment

• Comparative analysis of clustering methods
  – Definition of road groups with similar traffic patterns
  – Highlighting of differences among vehicle types

• First step of a more comprehensive analysis
  – Measure of the baseline level of errors in the procedure
Clustering Methods for Road Grouping

- Characteristics of Clustering Methods
- Methodology of study
- Analysis of Case Study Results
• TMG suggestions for the definition of road groups:
  – Geographical/Functional classification
  – “Same road” application of the factors
  – Cluster analysis
    • Least-squares minimum distance algorithm

• Data Mining Clustering Methods
  – Statistical and Data Mining softwares
    • R (R Development Core Team, 2010), Rapid Miner (Mierswa et al., 2006)
  – New tools available for road groups definition
    • Non-parametric
      – Hierarchical
      – Partitioning
    • Parametric (Model-based)
Agglomerative Hierarchical Clustering

- Hierarchical grouping of objects by gradually agglomerating them.

- Main steps:
  - At the beginning each object is assigned to a group
  - At each step the two most similar groups are merged in a new group that substitutes the previous ones
  - Merging is repeated until only one cluster exists

- Measures of similarity used:
  - Various choices, the most common is the p-dimensional Euclidean distance, where p is the number of attributes used for the clustering

- Similarity calculations tested:
  - Single link, Average link, Centroid link, Ward, Energy
Agglomerative Hierarchical Clustering

Step 1 \rightarrow Step 2 \rightarrow Step n \rightarrow Final Step
Main difficulty for agglomerative hierarchical clustering

- Choice of the optimal number of groups
  - Some performance indices were defined and can help to determine the optimal number of clusters
- In this paper a combination of criteria was adopted for this purpose:
  - Pseudo F Statistic
  - Analysis of the variance of clusters
  - Davies-Bouldin Index
  - Practical considerations
    - The groups must have a reasonable number of ATRs, avoiding to have groups with only one ATR.
Partitioning Clustering

- Each object of the dataset is placed in only one group.
- Main steps:
  - Usually the starting point is a random assignment of each object into one of the $k$ groups;
  - At each step the algorithm reassigns the objects to the nearest cluster on the basis of a certain measure of proximity.
  - The value $k$ can be defined a priori at the beginning of the clustering, or determined by the algorithm itself.
- Algorithms tested:
  - K-means (Hartigan&Wong, 1979)
  - Partitioning Around Medoids (Kaufman and Rousseeuw, 1990)
  - X-means (Pelleg & Moore, 2000)
Clustering Methods for Road Grouping: Partitioning Clustering

Partitioning Clustering

Step 1

Step i

Final Step
Model-Based Clustering (McLachlan & Peel, 2000)

• It assumes that:
  – each cluster could be represented by a density function belonging to a certain parametric family (e.g. the multivariate normal)
  – the associated parameters could be estimated from observations

• Main steps:
  – It determines the probabilistic density function for the $k$-th group estimating the $p$ dimensional mean vector and the $p \times p$ covariance matrix $\Sigma_k$, where $p$ is the number of attributes used for the clustering.
  – The merging process of clusters is controlled by a maximum-likelihood criterion which aims at maximizing the likelihood over the mixture parameters and identifying the group to which each object in the dataset belongs
  – The optimal number of groups is found employing the Bayesian Information Criteria (BIC).
Model-Based Clustering

• The covariance matrix can be decomposed as

\[ \Sigma_k = \lambda_k D_k A_k D_k^T \]

where \( \lambda_k \), \( D_k \), and \( A_k \) control the volume, the orientation and the shape of the \( k \)-th group (Banfield & Raftery, 1993)

• Different models can be built and analyzed considering these geometric features as separate and independent parameters.

• Identifiers codes can be used to describe the geometric characteristics of the model (volume, orientation and shape): E for equal, V for variable and I for Identity. For example a model classified as VEI means that the clusters are assumed to have a variable volume, equal shape and orientation on the coordinate axes.
Clustering Methods for Road Grouping

- Characteristics of Clustering Methods
- Study Methodology
- Analysis of Case Study Results
General Procedure

1. **Input data**
   a) Vehicle Classification (5.0 m length threshold)
      • Passenger cars PV
      • Trucks TV
   b) Directional Analysis (Tsapakis et al., 2011)
      • Traffic data from each section separated by direction

2. **Clustering Analysis of ATRs with the various techniques**
   a) Using PV and TV factors together (Gecchele et al., 2011)
   b) Separating PV and TV factors

3. **Calculation of Estimation errors** (Sharma et al., 1999) (Sharma et al., 2000)
   - 24-h sample counts generated from the main dataset
   - Comparison with actual AADT
Sample Generation and AADT Estimate

\[ DT_{i,j} \times f_{i,j} \]

24-h sample counts
adjustment factors

Estimated AADT
Calculation of Estimation Errors

- Estimation error (Percent Error, PE):

\[
PE = \frac{|\text{Estimated AADT} - \text{Actual AADT}|}{\text{Actual AADT}} \times 100
\]

where

\[
\text{Actual AADT} = \frac{1}{7} \sum_{i=1}^{7} \left[ \frac{1}{12} \sum_{j=1}^{12} \left( \frac{1}{n} \sum_{k=1}^{n} DT_{ijk} \right) \right]
\]

(AASHTO, 1992)

- Mean of the Estimation Error (MAE)
- Standard Deviation of the Estimation Error (SDAE)
- Percent Cumulative Frequency Distribution of the Estimation Error
Clustering Methods for Road Grouping

- Characteristics of Clustering Methods
- Study Methodology
- Analysis of Case Study Results
SITRA Program
Information Transportation System
- Monitoring Program based on FHWA Procedure
- Established in 2000
- 51 Permanent Traffic Counters
Temporal Traffic Pattern – PV class

Clustering Methods for Road Grouping: Analysis of Results
Temporal Traffic Pattern – TV class

Clustering Methods for Road Grouping: Analysis of Results
Mean of the Estimation Error – PV case (Total + Weekdays)

![Graph showing Mean of the Estimation Error for different clustering methods: Centroid, Energy, EEI, EII, EII, VII, VI, VI, K-Means, PAM, Ward, X-Means, Single Group, PV, PV_Weekdays, PV_Sep, PV_Weekdays_Sep. The x-axis represents different clustering methods, and the y-axis shows the mean estimation error.]
Mean of the Estimation Error – PV case (Saturdays + Sundays)
Mean of the Estimation Error – TV case (Total + Weekdays)
Mean of the Estimation Error – TV case (Saturdays + Sundays)
Clustering Methods for Road Grouping: Analysis of Results

Mean of the Estimation Error – PV case by periods

- PV Jan-Feb
- PV Mar-Apr
- PV May-Jun
- PV Jul-Aug
- PV Sep-Oct
- PV Nov-Dec

Mean values for different clustering methods and periods.
Mean of the Estimation Error – PV case by periods

Clustering Methods for Road Grouping: Analysis of Results
Mean of the Estimation Error – TV case by periods
Mean of the Estimation Error – TV case by periods
Mean of the Estimation Error – Error by road groups (EEI Method)
Conclusions

- Concluding Remarks
- Future Developments
Conclusions:
Concluding Remarks

Clustering methods identify a common basic structure of ATR groups which has functional and geographical significance.
The differences concentrate on the attribution of a small number of ATRs.

Clustering methods show common error patterns and give comparable results in term of average percent error distribution.
These results are better compared to a single group attribution.
Errors patterns show worse results in AADT estimation in some specific cases:

- Day-type: Saturdays and Sundays show higher errors than Weekdays
- Period of the year: summer period shows higher errors than winter period
- Vehicle types: the errors are higher for truck vehicles

Conclusions:

Concluding Remarks
Conclusions:
Research ideas/1

• **Observations:**
  • Road groups with crisp boundaries
  • In practice uncertainty about number and characteristics of groups

• **Ideas:**
  • Introducing *fuzzy set theory* to deal with the vagueness of boundaries between individual road groups

• **Implementation:**
  • Using Fuzzy C-means algorithm in order to:
    • Identify and separate ATR sites that clearly belong to just one group for their high membership grade to the group.
    • Identify ATR sites with intermediate membership grades as “uncertain groups,” meaning “I don’t know”, because they may belong to more than one group
Conclusions: Research ideas/2

• **Observations:**
  - Errors in assignment could lead to a tripling of the estimation error (Davis, 1996)

• **Ideas:**
  - Introducing the **measures of uncertainty** (non-specificity and discord) to deal with the difficulties of identifying the group that matches the given road section:
    - Help to interpret the quality of the estimates in an objective manner
    - Indicate the need for additional data collection

• **Implementation:**
  - Artificial Neural Network (ANN) in the assignment step
  - Calculation of non-specificity and discord measures
Conclusions:

Future Developments

Test the use of 48-h sample counts instead of 24-h counts

Test the influence of socio-economic and land-use characteristics in grouping process, in order to understand if this additional information could help to find more reliable and significant groups
Thank you for your attention...