Large scale estimation of arterial traffic and structural analysis of traffic patterns using probe vehicles

A. Hofleitner, R. Herring, A. Bayen, Y. Han, F. Moutarde and A. de La Fortelle

1) Traffic estimation from sparse probe vehicles

Fleet of **500 vehicles**, reporting their location every minute (San Francisco, CA).
Black dots: Cumulative measurements received between 12am and 7am.
Red circles: Location of the probe vehicles at 7am on that day.
A *map-matching and path inference* algorithm filters out GPS noise and reconstructs the path of probe vehicles between successive locations.

Statistical model of arterial traffic on large scale networks:

• Decomposition of path travel times into (partial) link travel times, maximizing the likelihood of the decomposition (small scale optimization with analytical algorithm).
• Historical learning of the distribution of travel times per time of day (TOD) and day of week (DOW).
• In real time, the parameters learned by the historic model are used as prior information to estimate current traffic conditions from streaming data via a *Bayesian* update.
• Traffic estimates produced by the *Mobile Millennium* system are used for large scale structural analysis of traffic patterns.

2) Non-negative Matrix Factorization (NMF)

What is NMF?

• Matrix factorization used for dimensionality reduction, in the same category as *Principal Component Analysis* (PCA)
• Factorize **X** ≈ **M V**, where **M** and **V** are matrices with ***non-negative*** elements which minimize the Froebenius distance between **X** and **MV**

$$\arg\min_{(\mathbf{M},\mathbf{V})} \|\mathbf{X}-\mathbf{MV}\|_F \text{ s. t. } \quad M \geq 0,\ V \geq 0$$

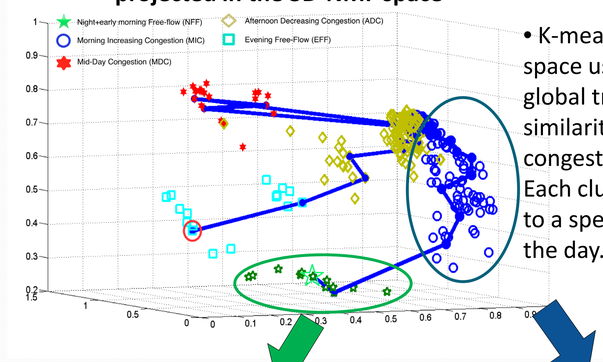
Why use NMF to analyze structural traffic patterns on large networks?

• Dimensionality reduction to analyze the evolution and clustering of traffic states.
• NMF, as opposed to PCA or other dimension reduction technique, can also provide a ***part-based decomposition***, *i.e.* each NMF component corresponds to a localized or at least synchronized subpart of the road network

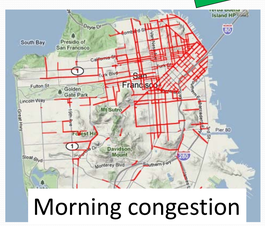
Example of localized NMF components obtained on face images

3) Temporal clustering of daily traffic states

• **Distributions of the global traffic states projected in the 3D-NMF space**



• K-means in the 15-D NMF space used to cluster global traffic states by similarity of geographical congestion pattern.
Each cluster corresponds to a specific moment of the day.

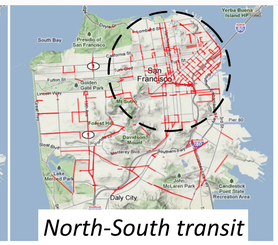


Morning congestion

Mid-day congestion

4) Spatial decomposition of the road network

East-West transit



North-South transit

Spatial segmentation of the network
For each NMF basis, we display the links with the 20% largest contribution to the component.

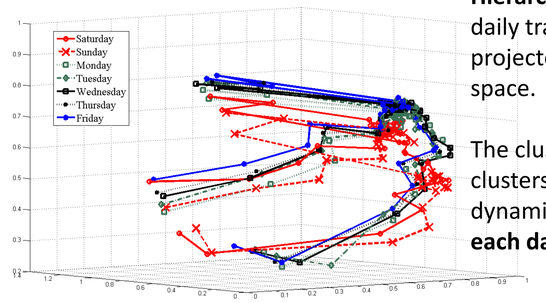
Links with similar orientations are part of the same basis (in particular in the downtown area), indicating that they are likely to have correlated dynamical behaviors. Links with orthogonal orientations are less correlated.

West component

Downtown component

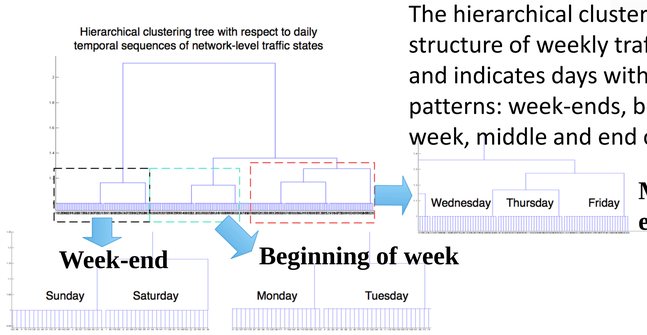
NMF basis exhibits regions of the city in which the correlation between links is high.

5) Temporal analysis



Hierarchical clustering of the daily traffic state trajectories projected on the 15-D NMF space.

The clustering exhibits **7** clusters representing typical dynamical traffic patterns for **each day of the week**.



The hierarchical clustering unveils the structure of weekly traffic dynamics and indicates days with similar patterns: week-ends, beginning of the week, middle and end of the week.

6) Conclusion and discussions

• **Real-time estimation** of travel time distribution from streaming probe vehicle data. Algorithm is based on historical learning and real time Bayesian update.
• **Data mining analysis**: NMF based manifold projection of large amounts of estimates (several months of estimates on a large urban network). Low dimension representation to provide a structural analysis of large scale traffic patterns.
• Typical **spatial configuration patterns** of traffic congestion over the entire transportation network underline links with similar dynamics and exhibits spatial dependencies in the network. This is important to understand how congestion spreads over a network, predict traffic and develop appropriate control mechanisms.
• **Hierarchical clustering structure of daily traffic dynamics** in the low-dimensional NMF representation of the global traffic states unveils the daily and weekly dynamics of traffic. It can also be used to rapidly detect unusual traffic conditions.