Incorporating Spatio-Temporal Smoothness for Air Quality Inference

Xiangyu Zhao[1,2], Tong Xu[1], Yanjie Fu[3], Enhong Chen[1], Hao Guo[1]

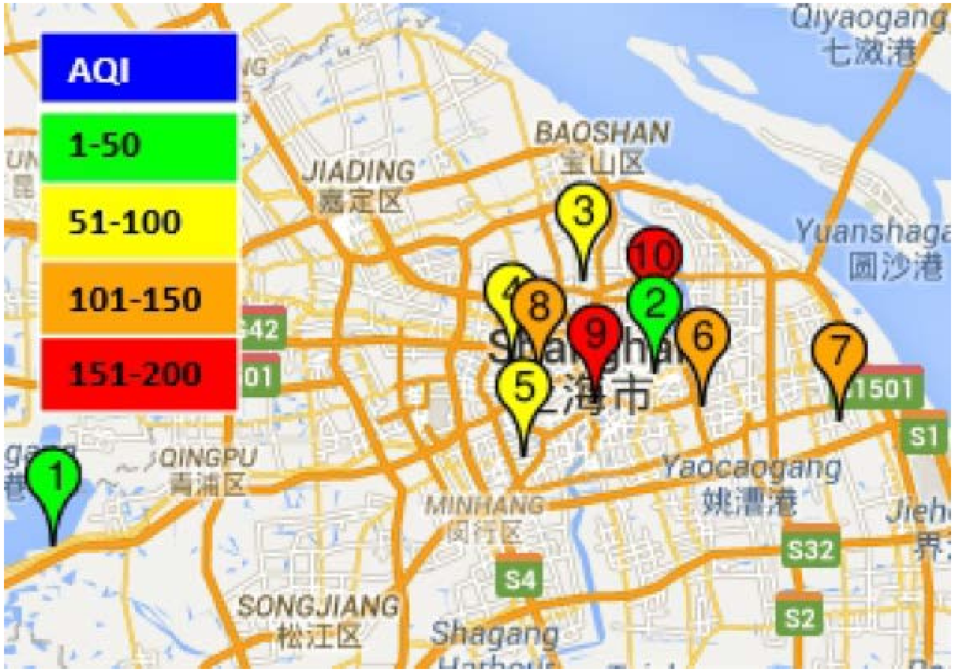
1 University of Science and Technology of China
2 Michigan State University
3 Missouri University of Science and Technology

Motivation

- Increasing concern of urban air quality
  - Life quality of residents
  - Sustainable development of city

Motivation

- Challenge:
  - The number of monitoring stations is limited
  - Monitoring stations are not evenly distributed



Motivation

- Two Intuitive Assumptions:
  - **Temporal dependence**: *intra-station* time dependence within a single monitoring station, as current AQI value won't change a lot compared with air quality in the near future.
  - **Spatial relatedness** : *inter-station* spatial relatedness across all the stations, as two stations which located nearby should have similar AQI.

Urban Air Quality Inference Framework

- Spatio-Temporal Smoothness:
  - Basic model

$$\min_{\mathbf{W}_{:,k}} L_k = \|\mathbf{Y}_{:,k} - \mathbf{X}_{:,k} \odot \mathbf{W}_{:,k}\|_2^2 + \gamma \|\mathbf{W}_{:,k}\|_F^2$$

  - Distance-based Spatial Smoothness

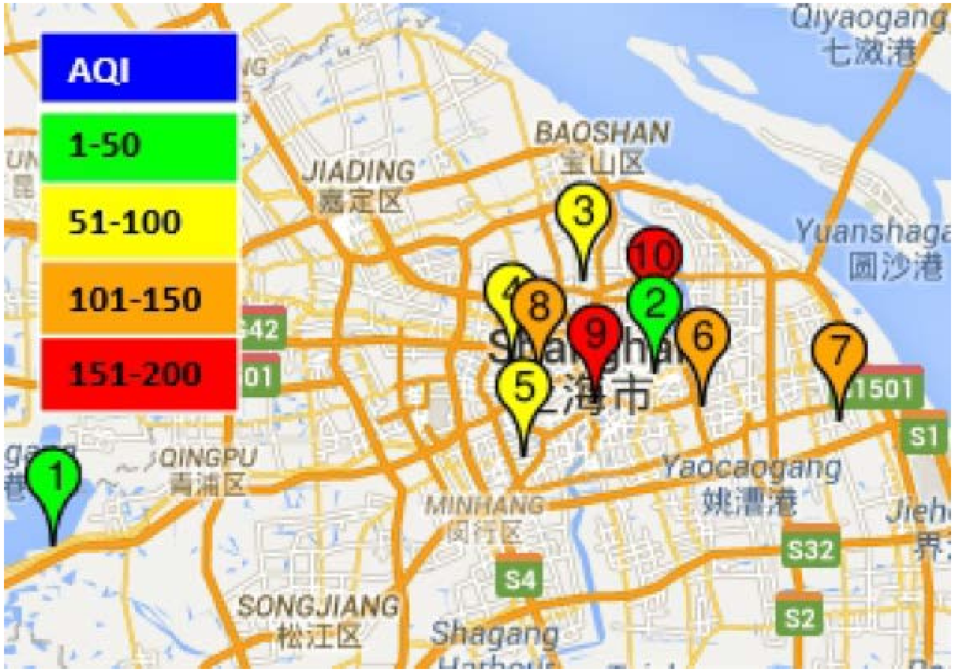
$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} D_{ij}\|\mathbf{W}_{i,k} - \mathbf{W}_{j,k}\|_2^2$$

  - Temporal Smoothness

$$\sum_{n=1}^{N}\left(\sum_{k=2}^{K}\|\mathbf{W}_{n,k} - \mathbf{W}_{n,k-1}\|_2^2\right)$$

Urban Air Quality Inference Framework

- Spatio-Temporal Smoothness:
  - Real-Time Feature-based Smoothness

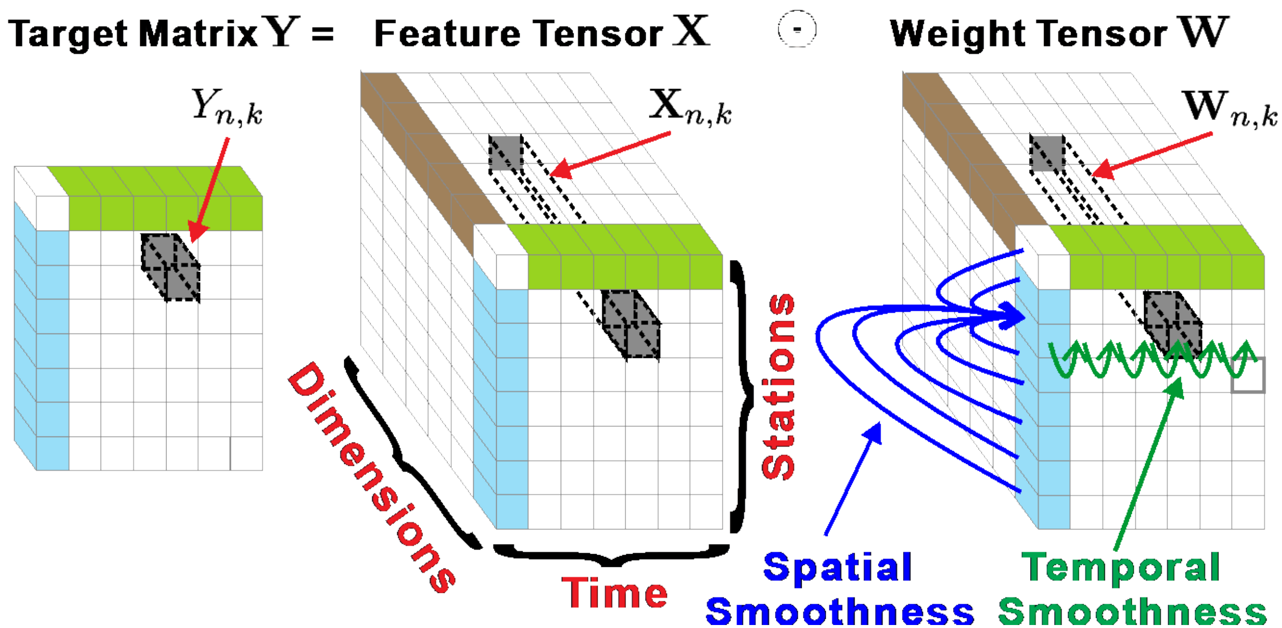


Urban Air Quality Inference Framework

- Spatio-Temporal Smoothness:
  - Real-Time Feature-based Smoothness

$$\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} F_{ij}^{k}\|\mathbf{W}_{i,k}-\mathbf{W}_{j,k}\|_2^2 \qquad F_{ij}^{k}=cosine(\mathbf{X}_{i,k},\mathbf{X}_{j,k})=\frac{\mathbf{X}_{i,k}\cdot\mathbf{X}_{j,k}}{\|\mathbf{X}_{i,k}\|\|\mathbf{X}_{j,k}\|}$$

Urban Air Quality Inference Framework



$$\min_{\mathbf{W}} L = \sum_{k=1}^{K}\left(\|\mathbf{Y}_{:,k} - \mathbf{X}_{:,k} \odot \mathbf{W}_{:,k}\|_2^2 + \gamma\|\mathbf{W}_{:,k}\|_F^2 + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\alpha D_{ij} + \beta F_{ij}^k\right)\|\mathbf{W}_{i,k} - \mathbf{W}_{j,k}\|_2^2\right)$$

$$+\lambda\sum_{n=1}^{N}\left(\sum_{k=2}^{K}\|\mathbf{W}_{n,k} - \mathbf{W}_{n,k-1}\|_2^2\right)$$

Experiment Setting

- Datasets:
  - Shanghai City, China
  - April 1 to April 30, 2015
- 9 stations as training set, 1 station as test set
- Metric
  - Average root-mean-square-error (RMSE)

$$RMSE = \frac{1}{N}\sum_{n=1}^{N}\sqrt{\frac{1}{K}\sum_{k=1}^{K}(y_k - \hat{y_k})^2}$$

Experiment Results

TABLE I
OVERALL PERFORMANCE (RMSE) OF EACH APPROACH.

Temporal	1 hour	3 hour	**Spatial**	real-time
ARIMA	30.225	45.787	Average	46.563
VAR	28.756	42.907	IDW+	39.016
LASSO	25.387	38.653	CoKriging	35.291
stMTL	18.176	30.009	ANN	29.667
stMTMV	13.989	24.239	SFST	25.290
stfMTR	**12.595**	**20.562**	stfMTR	**22.633**

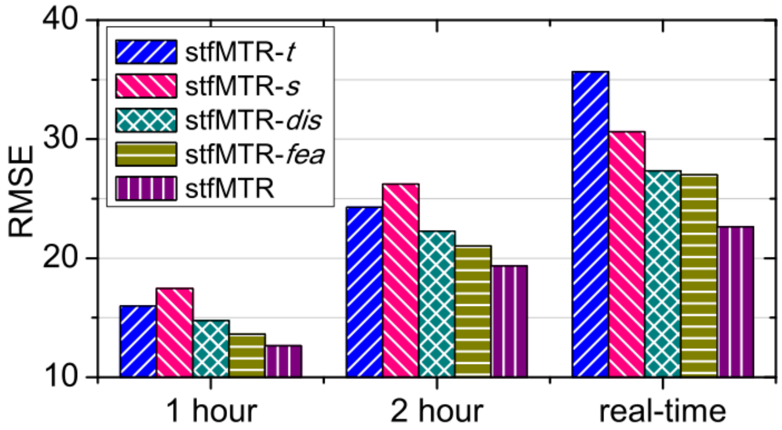


Fig. 3. Performance comparison on model components.

- Our **<u>stfMTR</u>** performs the best with integrating *spatial* and *temporal* smoothness
- *Feature similarity* could be more important compared with *distance proximity*

Conclusion

- *Intra-station* time dependences and the *inter-station* spatial relatedness are both beneficial.
- Feature similarity will enrich the spatial smoothness with removing the bias.
- Theoretically, given the *features* and *historical AQI*, we could predict AQI in any place.

Thanks

tongxu@ustc.edu.cn