Incorporating Spatio-Temporal Smoothness for Air Quality Inference

Xiangyu Zhao\textsuperscript{1,2}, Tong Xu\textsuperscript{1}, Yanjie Fu\textsuperscript{3}, Enhong Chen\textsuperscript{1}, Hao Guo\textsuperscript{1}

\textsuperscript{1} University of Science and Technology of China
\textsuperscript{2} Michigan State University
\textsuperscript{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**

\[
\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} F_{ij}^k \| W_{i,k} - W_{j,k} \|_2^2 \\
F_{ij}^k = \cosine(X_{i,k}, X_{j,k}) = \frac{X_{i,k} \cdot X_{j,k}}{\| X_{i,k} \| \| X_{j,k} \|}
\]
Urban Air Quality Inference Framework

\[ \min_{W} L = \sum_{k=1}^{K} \left( \| Y_{:,k} - X_{:,k} \otimes W_{:,k} \|_2^2 + \gamma \| W_{:,k} \|_F^2 + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha D_{ij} + \beta F_{ij}^k) \| W_{i,k} - W_{j,k} \|_2^2 \right) \\
+ \lambda \sum_{n=1}^{N} \left( \sum_{k=2}^{K} \| W_{n,k} - W_{n,k-1} \|_2 \right) \]
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.

<table>
<thead>
<tr>
<th>Temporal</th>
<th>1 hour</th>
<th>3 hour</th>
<th>Spatial</th>
<th>real-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>30.225</td>
<td>45.787</td>
<td>Average</td>
<td>46.563</td>
</tr>
<tr>
<td>VAR</td>
<td>28.756</td>
<td>42.907</td>
<td>IDW+</td>
<td>39.016</td>
</tr>
<tr>
<td>LASSO</td>
<td>25.387</td>
<td>38.653</td>
<td>CoKriging</td>
<td>35.291</td>
</tr>
<tr>
<td>stMTL</td>
<td>18.176</td>
<td>30.009</td>
<td>ANN</td>
<td>29.667</td>
</tr>
<tr>
<td>stMTMV</td>
<td>13.989</td>
<td>24.239</td>
<td>SFST</td>
<td>25.290</td>
</tr>
<tr>
<td>stfMTR</td>
<td><strong>12.595</strong></td>
<td><strong>20.562</strong></td>
<td>stfMTR</td>
<td><strong>22.633</strong></td>
</tr>
</tbody>
</table>

- Our stfMTR 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