



**POLYTECHNIQUE
MONTREAL**

UNIVERSITÉ
D'INGÉNIERIE



IVADO

transdim: Machine Learning for Transportation Data Imputation and Prediction

Reproducible Research Workshop

TRB 103rd Annual Meeting · Washington, D.C.

Xinyu Chen

January 11, 2024

Open-source & reproducible research:

- ① GitHub repository: <https://github.com/xinychen/transdim>
- ② Slides: <https://xinychen.github.io/slides/transdim.pdf>

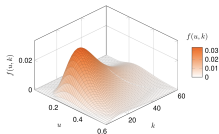
ML algorithms



transdim

(1.1k stars)

Visualization tools



awesome-latex-drawing

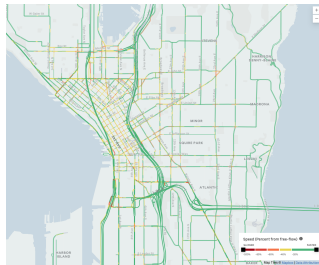
(1.2k stars)

Storytelling with Data

- Uber (hourly) movement speed data



NYC movement



Seattle movement

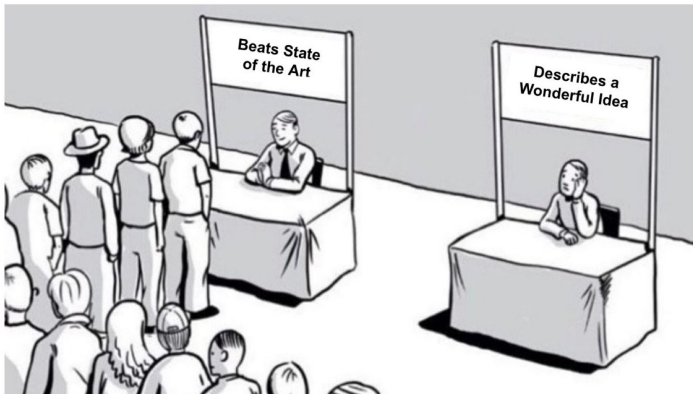
- $\{\text{road segment, time step (hour), average speed}\}$
- $\mathbf{Y} \in \mathbb{R}^{N \times T}$ with N spatial locations $\times T$ time steps
- Computing hourly speed: Road segments have 5+ unique trips.

Issue: Insufficient sampling of ridesharing vehicles on the road network!

Storytelling with Data

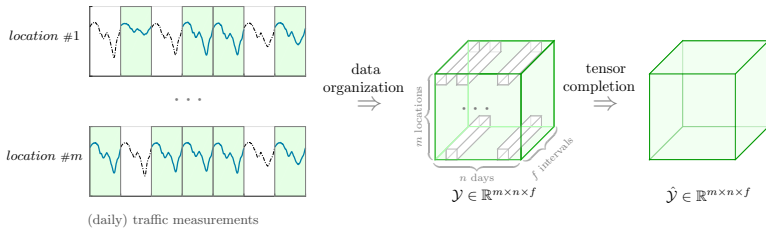
- Data
- Quality
- Sparsity
- Estimation
- Imputation
- Interpolation
- Forecasting

Storytelling with Data



Source: Twitter

Reformulate Traffic Data Imputation



Why?

Academic:

- Provide platform and benchmark for comparison
- Stimulate new algorithmic ideas

Industry:

- Solution to ...

Next-step plan:

- Spatiotemporal data modeling:
<https://spatiotemporal-data.github.io>



**POLYTECHNIQUE
MONTREAL**

UNIVERSITÉ
D'INGÉNIERIE



IVADO

Thanks for your attention!

Any Questions?

About me:

 Homepage: <https://xinychen.github.io>

 GitHub: <https://github.com/xinychen>

 How to reach me: chenxy346@gmail.com