



Xinyu Chen

Postdoc, MIT (now)

Advisor: Prof. Jinhua Zhao

PhD, University of Montreal ('23)
Civil Engineering (Transportation)

Interests

- Advanced computing for engineering
- Urban system & human mobility
- Data-driven traffic flow modeling
- Climate system monitoring
- Machine learning & data science
- Optimization & math programming

Interdisciplinary Research

Computational Engineering in CEE

PhD (ML for Transportation)

- **Traffic imputation** w/ tensor decomposition
[Chen et al.'19](#); [Chen et al.'21](#) in TR-C (cited 300+)
[Chen et al.'22](#) in IEEE TITS (cited 100+)
- **Time series imputation** w/ Laplacian convolution
[Chen et al.'24](#) in IEEE TKDE
- **Mobility prediction** w/ Bayesian optimization
[Chen & Sun'22](#) in IEEE TPAMI (cited 250+)
- **Traffic prediction** w/ Hankel factorization
[Chen et al.'24](#) in IJOC
- **Dynamic climate pattern discovery**
[Chen et al.'24](#) in IEEE TKDE

Postdoc (ML + Optimization for Spatiotemporal Data)

- **Tensor decomposition for ML**
[Chen et al.'24](#), major revision in IEEE TPAMI
- **Causal inference from climate systems**
[Chen et al.'24](#), 2nd-round review in IEEE TKDE
- **Mobility periodicity quantification w/ MIP**

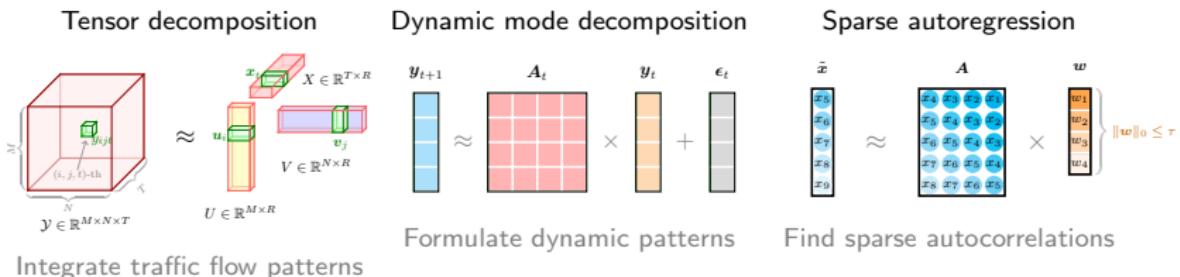


Research Contributions

- Formulating challenging engineering problems



- Advancing ML development with methodological contributions



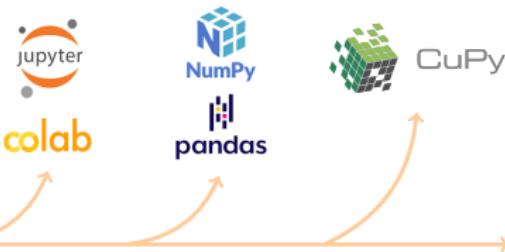
Reproducible Research for Engineering

- The last mile of AI for computational engineering

Human mobility & smart cities
Data-driven transport analytics
Spatiotemporal data modeling
Interpretable ML for causal inference
Tensor decomposition for ML

...

Directions & Topics



Reproducible Research

- Advancing ML development with open-source research



transdim
(1,200+ GitHub stars)

ML for Transport Data Imputation

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

The screenshot shows the 'Tensor4ML' initiative page on the MIT website. The header includes the MIT logo and navigation links for Home, About, Research, People, News, and Publications. The main content area is titled 'Tensor Decomposition for Machine Learning' and features a sub-section titled 'Tensor Decomposition for Machine Learning' by Xinyi Chen, Sheng Dong, Junjie Zhou (2022). Below the title, there is a brief description of the initiative's purpose: 'An easy view of the development of tensor decompositional models and algorithms, along with tutorials on matrix and tensor decompositions, and an tensor decomposition techniques across a wide range of scientific areas and applications.'

Tensor Decomposition for ML
(ML initiative)

Math & ML Tutorials

<https://sites.mit.edu/tensor4ml>

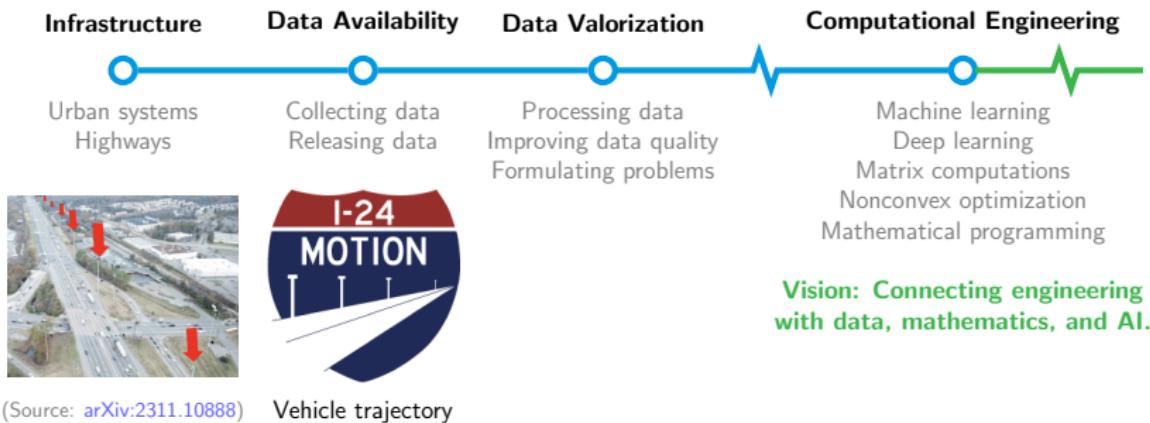


Spatiotemporal data modeling
(Data valorization initiative)

Model Development of ML & Data Science

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

Building Research Impact at Vanderbilt



(Source: [arXiv:2311.10888](https://arxiv.org/abs/2311.10888))

Vehicle trajectory

- Conduct interdisciplinary research in CEE, Transportation Engineering, Computer Science, and Management Science
- Reshape methodological & practical contributions to multidisciplinary research communities
- Initiatives at Vanderbilt: College of Connected Computing

Teaching & Grant

- Teaching interests & plan
 - Formats: Tutorial, data example, LaTeX graphic, Python code, GitHub repository, and course website

Figure 2 shows taxi pickup and dropoff trips (2023) on 77 community areas in the City of Chicago. Note that the average trip duration is 1207.75 seconds and the average trip distance is 6.16 miles.

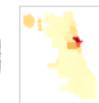


Figure 2. Taxi pickup and dropoff trips (2023) in the City of Chicago, USA. There are 4,783,861 remaining trips after the data processing.

For comparison, Figure 3 shows taxi pickup and dropoff trips (2019) on 77 community areas in the City of Chicago. Note that the average trip duration is 918.62 seconds and the average trip distance is 3.93 miles.



Figure 3. Taxi pickup and dropoff trips (2019) in the City of Chicago, USA. There are 12,484,572 remaining trips after the data processing. See the data processing codes.

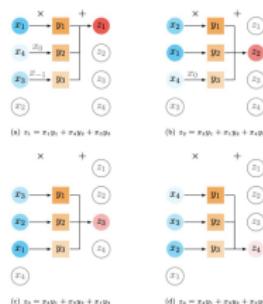
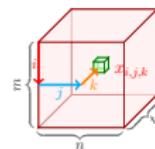
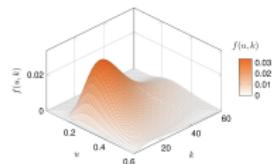


Figure 2. Illustration of the circular convolution between $x = (x_1, x_2, x_3, x_4)^T$ and $y = (y_1, y_2, y_3)^T$.
.. (a) Computing z_1 involves $x_0 = x_4$ and $x_{-1} = x_3$. (b) Computing z_2 involves $x_0 = x_4$. The figure inspired by Prince (2023).



Data-Driven Transportation Analytics

Advanced Computing for CEE

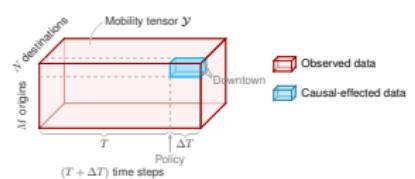
Spatiotemporal Modeling in CEE

- Grant & Funding



Transit-Centric Smart Mobility System with ML (DOE)

(PI: Jinhua Zhao. Role: Senior Researcher)



Causal inference for congestion pricing (NSF, submitted)

(PI: Jinhua Zhao; Co-PI: Ankur Moitra. Role: Senior Researcher)