Collaborative Imputation of Urban Time Series through Cross-city Meta-learning

Tong Nie and Wei Ma

tong.nie@connect.polyu.hk

ICLR'25 Workshop on Weight Space Learning

Background

Background:

- Urban time series, such as mobility flows, energy consumption, and pollution records, include complex urban dynamics and structures.
- Data collection in each city is impeded by technical challenges such as budget limitations and sensor failures, needing effective imputation techniques.
- Existing learning-based and analytics-based paradigms, face **the trade-off between capacity and generalizability**.

Our Solution

Collaborative imputation scheme:

- Collaborative learning to reconstruct data across multiple cities holds the promise of breaking this trade-off.
- Urban data's inherent **irregularity and heterogeneity** issues exacerbate challenges of knowledge sharing and collaboration across cities.
- We propose a novel collaborative imputation paradigm leveraging **meta-learned implicit neural representations (INRs)**.

Formulation

INRs for Time Series

$$\mathcal{C}(t, \mathbf{x}(t), \nabla_t \mathbf{x}(t), \nabla_t^2 \mathbf{x}(t), \dots) = 0, \quad \min \mathcal{L} = \int_{\mathcal{T}} \| \mathcal{C}(\mathbf{x}(t), \Phi, \nabla_t \Phi, \nabla_t^2 \Phi, \dots) \| dt,$$

Dynamics Reconstruction by Embedding Theory

$$\Psi : \mathcal{R} \mapsto \mathcal{R}^{d_r}, \ \vec{\mathbf{x}}(t) = \Psi(x(t)), \ \ \Psi^{-1} : \vec{\mathcal{M}} \mapsto \mathcal{S}, \Psi^{-1}(\vec{\mathbf{x}}(t)) = x(t),$$

$$\tilde{x}(t) = \Psi^{-1} \circ \mathcal{F}_{\Theta} \circ \Psi(x(t)).$$

Learning the parameterized reconstructor

Coordinate Delay Embedding

$$\vec{\mathbf{X}} = \Psi(\mathbf{x}) = [\Psi(x(1)), \Psi(x(2)), \dots, \Psi(x(T-(m-1)\delta))]^T,$$

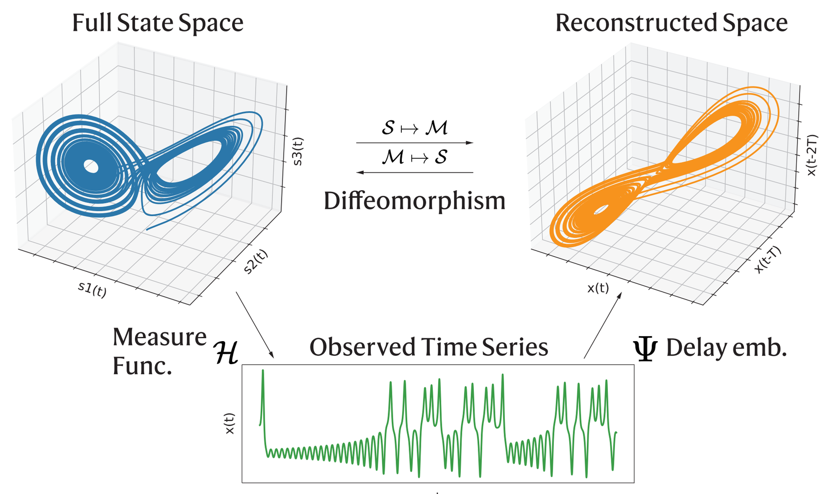
$$\hat{\Psi}(\mathbf{x}) \approx \mathcal{F} \circ \Psi(\tau) \in \mathbb{R}^{(T-(m-1)\delta) \times m\delta}, \ \ \Psi^{-1} \circ \hat{\Psi}(\mathbf{x}) = \mathbf{D}^{+} \operatorname{vec}(\hat{\Psi}(\mathbf{x})) \in \mathbb{R}^{T},$$

Frequency-decomposed Multi-scale INRs

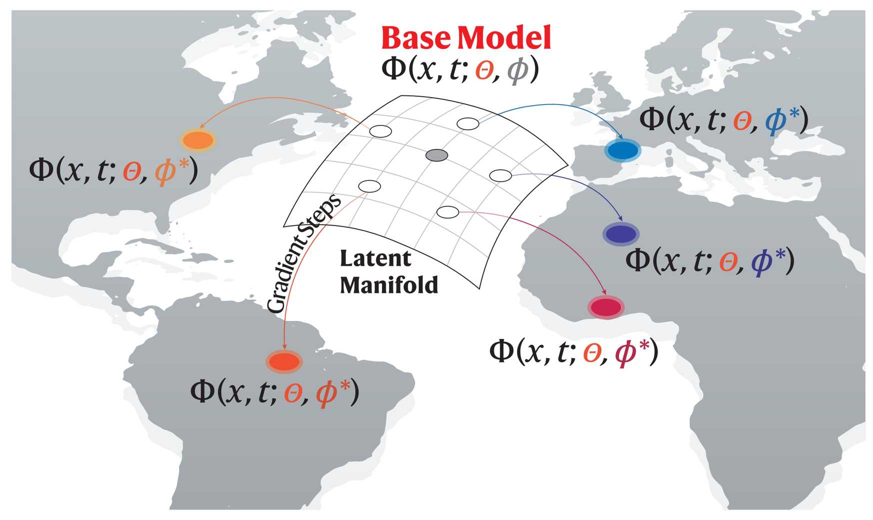
$$\Phi_{\Theta}(\vec{\tau}_t) = \mathcal{C} \times_1 \Phi_{\theta^1}(\vec{\tau}_t^1) \times_2 \Phi_{\theta^2}(\vec{\tau}_t^2), \forall (\vec{\tau}_t^1, \vec{\tau}_t^2) \in \vec{\tau}_t, \ \Phi_{\theta^i} : \vec{\tau}_t^i \mapsto \Phi_{\theta^i}(\vec{\tau}_t^i) \in \mathbb{R}^{n_i}, \ \forall i \in \{1, 2\},$$

Introduction

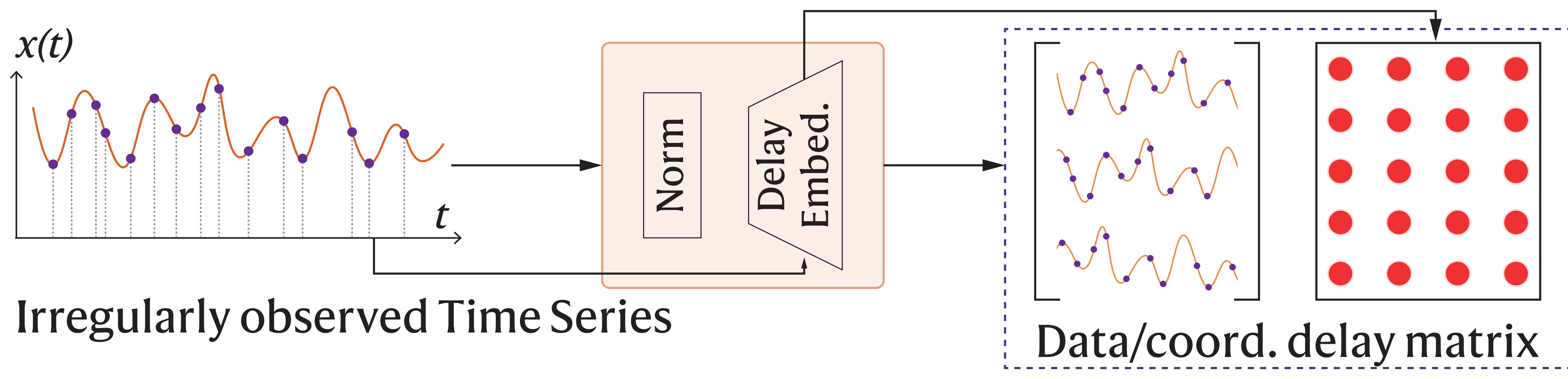
a. Dynamics reconstruction using delay embedding.



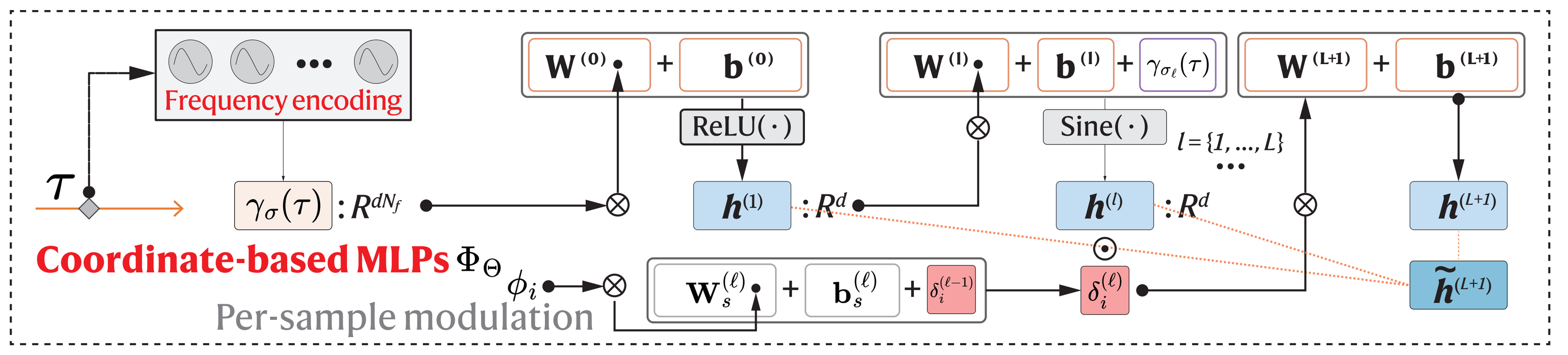
b. Meta-learning enables cross-city collaborative learning.



c. Imputation via implicit neural representations.



d. Architecture of MetaTSI.



$$\mathbf{h}^{(1)} = \texttt{ReLU}(\mathbf{W}^{(0)} \gamma_{\sigma}(\vec{\tau}_t^i) + \mathbf{b}^{(0)}), \ \mathbf{h}^{(\ell+1)} = \delta_i^{(\ell)} \odot \sin(\mathbf{W}^{(\ell)} \mathbf{h}^{(\ell)} + \mathbf{b}^{(\ell)} + \gamma_{\sigma_\ell}(\vec{\tau}_t^i)),$$

$$\tilde{\mathbf{h}}^{(L+1)} = \mathbf{W}^{(L+1)} \mathbf{h}^{(L+1)} + \mathbf{b}^{(L+1)},$$

High-frequency components with Fourier features

Cross-city Generalization by Meta Learning

$$\min_{\Theta, \phi} \ell(\mathcal{X}; \Theta, \{\phi^i\}_{i=1}^N) = \mathbb{E}_{\mathbf{x} \sim \mathcal{X}} [\ell^i(\mathbf{x}^i; \Theta, \phi^i)] = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{T_i} \sum_{t=1}^{T_i} \| x_t^i - \Phi_{\Theta, \phi}(\tau_t; \phi^i) \|_2^2,$$

$$\phi^{(k+1),i} \leftarrow \phi^{(k),i} - \alpha \nabla_{\phi^i} \ell(\Phi_{\Theta, h_\omega(\phi)}, \{\mathbf{x}^i\}_{i \in \mathcal{B}}),$$

Inner loop update and modulation

$$\delta^{(\ell),i} = s^i + \delta^{(\ell-1),i}, \ \ell = \{1, \dots, L\}, s^i = h_\omega^{(\ell)}(\phi^i) = \mathbf{W}_s^{(\ell)} \phi^i + b_s^{(\ell)}, \delta^{(0),i} = \mathbf{W}_s^{(0)} \phi^i + b_s^{(0)},$$

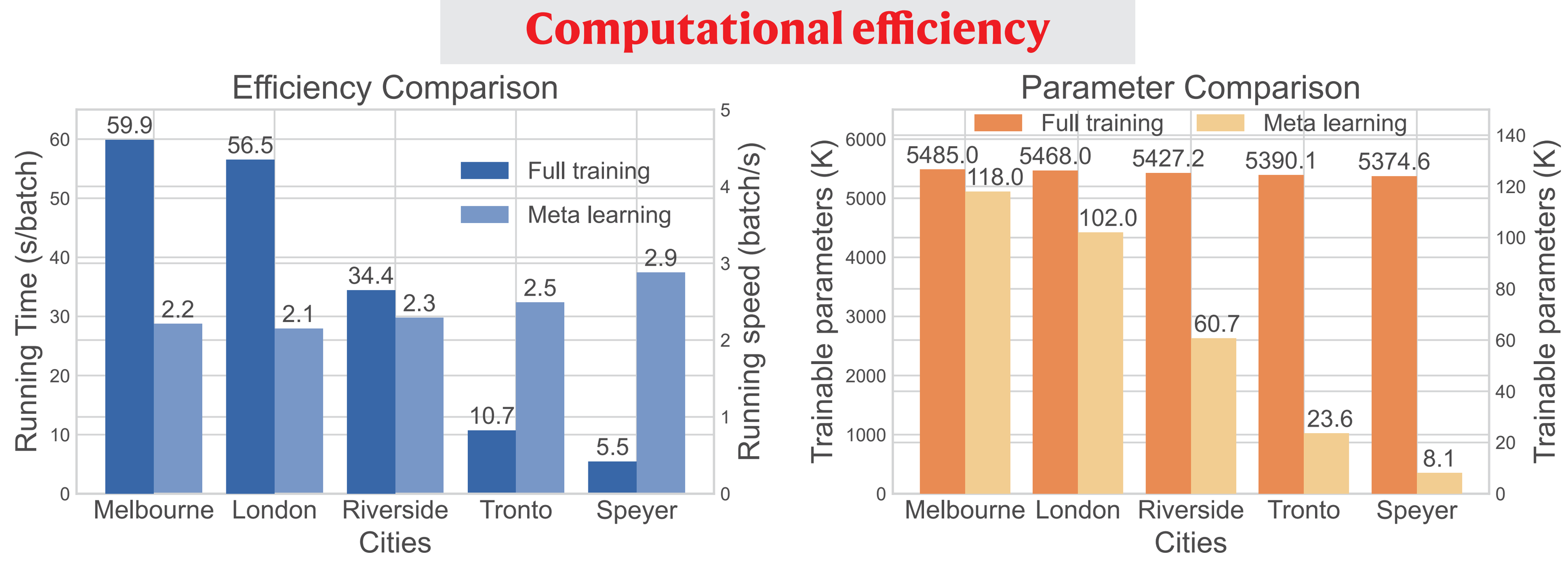
Experiments & Findings

A heterogeneous urban dataset from 20 global cities

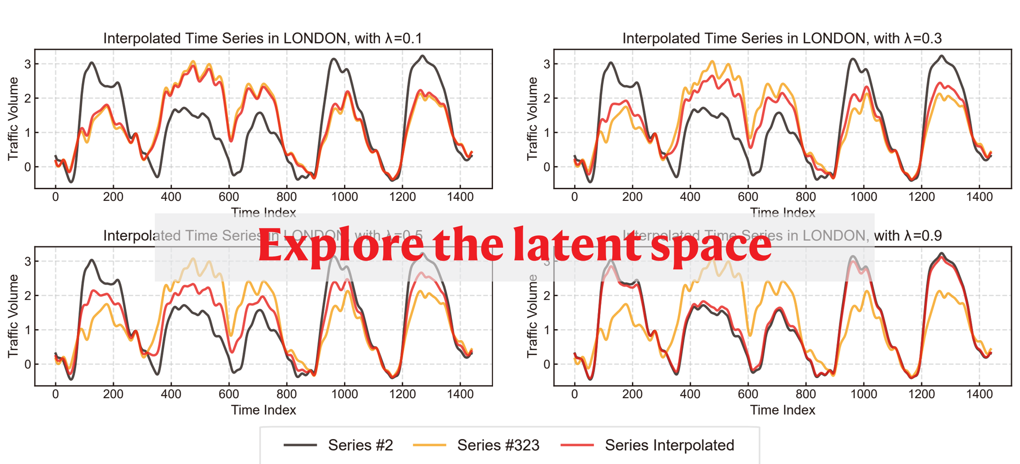
Model generalization on unseen time series with varying missing rates (in terms of MAE).

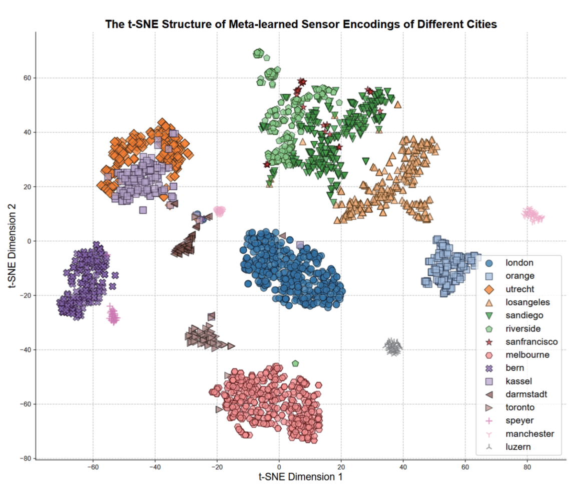
	Models	MetaTSI-1D		MetaTSI-2D		TimeFlow	
	Cities	MAE	MSE	MAE	MSE	MAE	MSE
Unseen Series	London	**0.242**	**0.125**	0.868	1.285	0.493	0.468
	Orange	**0.239**	**0.108**	0.448	0.360	0.421	0.294
	Utrecht	**0.289**	**0.180**	0.686	0.870	0.776	1.154
	Los Angeles	**0.186**	**0.076**	0.555	0.523	0.446	0.335
	San Diego	**0.176**	**0.067**	0.552	0.522	0.468	0.365
	Riverside	**0.181**	**0.073**	0.579	0.589	0.375	0.245
	San Francisco	**0.256**	**0.132**	0.522	0.461	0.498	0.411
	Melbourne	**0.341**	**0.257**	0.754	0.955	0.439	0.374
	Bern	**0.297**	**0.179**	0.581	0.635	0.758	1.056
	Kassel	**0.336**	**0.254**	0.668	0.828	0.775	1.074
	Darmstadt	**0.410**	**0.348**	0.755	0.981	0.746	0.934
	Toronto	**0.330**	**0.238**	0.872	1.245	0.442	0.378
	Speyer	**0.454**	**0.387**	0.549	0.522	0.862	1.255
	Manchester	**0.178**	**0.082**	0.886	1.383	0.657	0.822
	Luzern	**0.338**	**0.200**	0.571	0.579	0.896	1.281
	Average	**0.284**	**0.180**	0.656	0.783	0.603	0.696
New Cities	Taipei	**0.435**	**0.375**	0.910	1.437	0.631	0.746
	Contra Costa	**0.223**	**0.098**	0.592	0.610	0.561	0.513
	Hamburg	**0.433**	**0.363**	0.776	1.035	0.740	0.997
	Munich	**0.266**	**0.189**	0.678	0.815	0.583	0.701
	Zurich	**0.416**	**0.332**	0.752	0.997	0.721	0.947
	Average	**0.354**	**0.271**	0.742	0.979	0.647	0.781

Computational efficiency



Explore the latent space





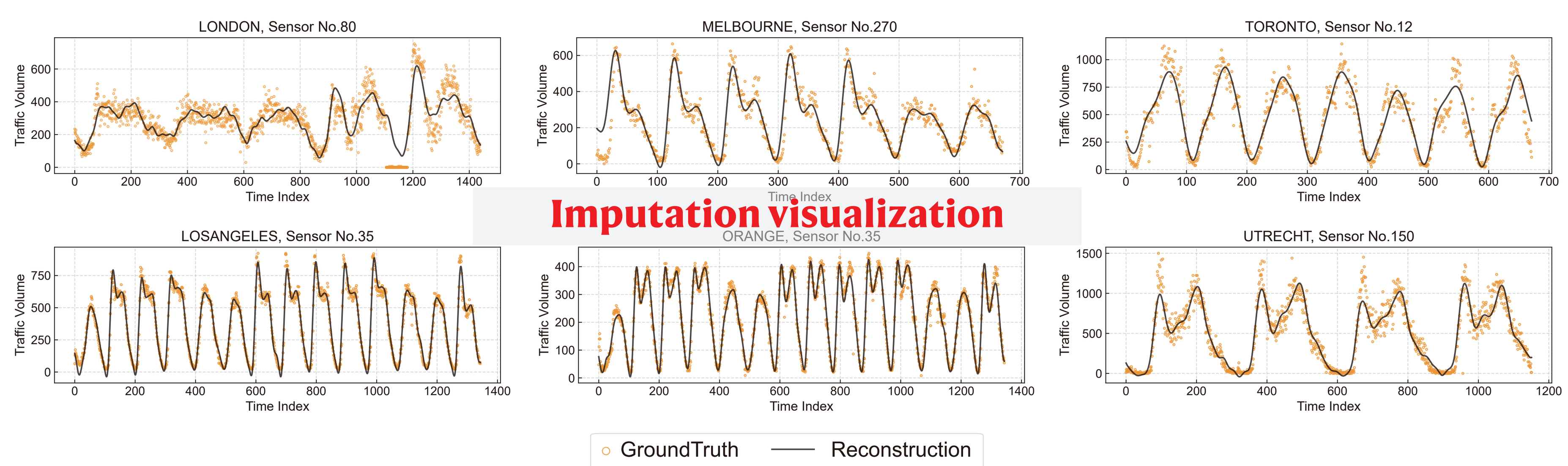
Takeaways

- MetaTSI originally pretrained to distinguish different instances can then generalize to heterogeneous cities to achieve collaborative imputation tasks.
- MetaTSI recovers the global structure of the signal manifold, allowing easy interpolation between nearby signals.
- Limitation: pretraining it with meta learning in the source data is computationally expensive and less stable than standard supervised learning.

Model generalization on unseen time series with varying missing rates (in terms of MAE). MetaTSI only performs meta learning with a small number of steps.

New series	San Diego			Riverside			Orange			Los Angeles		
Models	Observed rate			Observed rate			Observed rate			Observed rate		
	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
FourierNet†	0.254	0.274	0.565	0.264	0.268	0.501	0.212	0.339	0.589	0.254	0.271	0.492
LRTFR†	0.227	0.235	0.597	0.204	0.289	0.471	0.270	0.316	0.581	0.227	0.330	0.513
TimeFlow‡	0.389	0.505	0.759	0.320	0.424	0.695	0.479	0.661	0.876	0.391	0.568	0.764
TimeFlow†	0.403	0.468	0.782	0.390	0.398	0.726	0.340	0.513	0.856	0.404	0.487	0.802
Functa‡	0.301	0.385	0.871	0.279	0.383	0.864	0.313	0.431	0.976	0.297	0.382	0.873
Functa†	0.332	0.391	0.922	0.269	0.401	0.929	0.308	0.426	0.981	0.311	0.420	0.965
MetaTSI†	0.229	0.308	0.630	0.192	0.255	0.589	0.260	0.358	0.748	0.230	0.310	0.591
MetaTSI	**0.162**	**0.226**	**0.437**	**0.164**	**0.248**	**0.449**	**0.178**	**0.250**	**0.493**	**0.162**	**0.223**	**0.374**

Imputation visualization



SCAN TO VIEW MORE OF OUR WORK

THE HONG KONG POLYTECHNIC UNIVERSITY 香港理工大學

同濟大學 TONGJI UNIVERSITY

TOPS

Mobility AI Lab

WE ARE INTERESTED IN:

- AI for Social Good
- Urban Science
- Smart Transportation
- Large Language Models

wei.w.ma@polyu.edu.hk