Deep Learning for Human Mobility Analytics

-- L4: Cross-Domain Data Fusion and Multimodal Learning

Yuxuan Liang (梁宇轩)

INTR & DSA Thrust

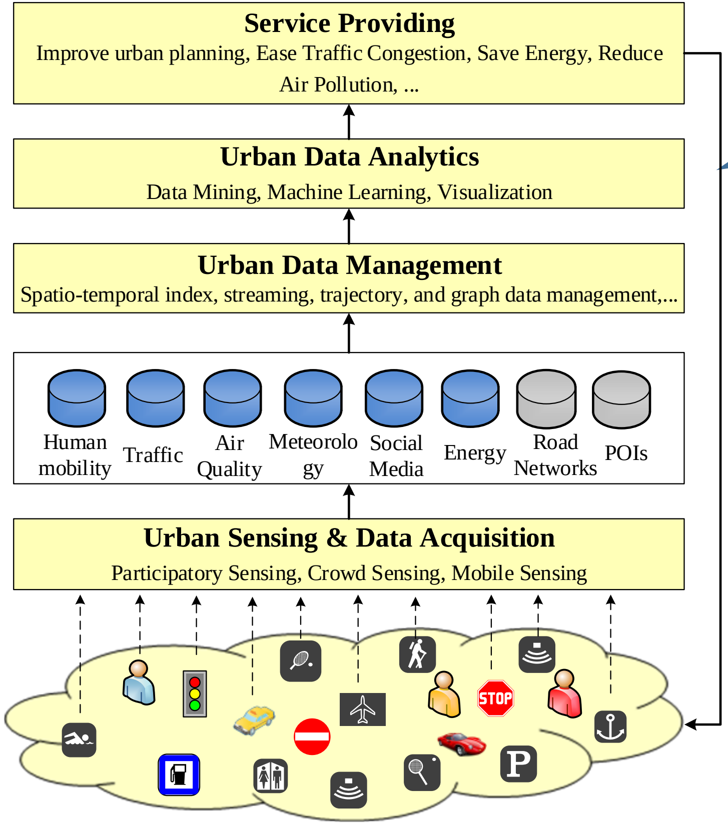
yuxuanliang@hkust-gz.edu.cn

Objectives of this Course

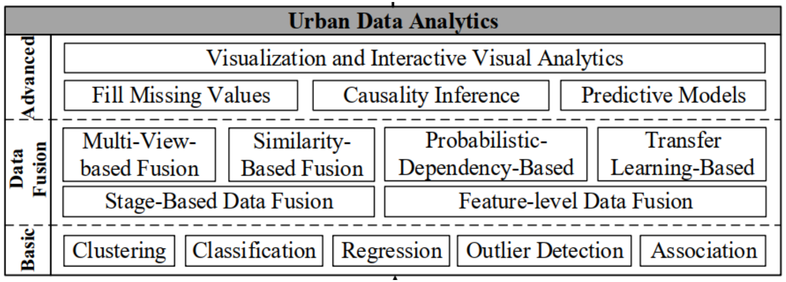
To introduce

- Stage-based data fusion
- Feature-level-based data fusion
  - Feature concatenation + regularization
  - Deep learning-based fusion
- Semantic meaning-based fusion
  - Multiple-view-based: like co-training
  - Similarity-based: Coupled matrix factorization
  - PGM-based
  - Transfer learning-based

3rd Stage: Urban Data Analytics



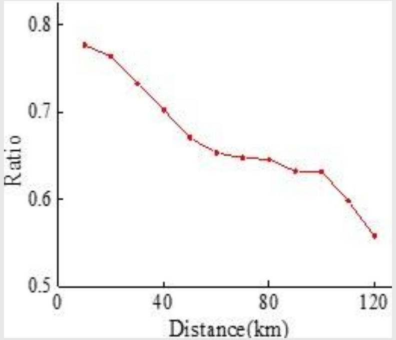
- Texts and images → spatio-temporal data
- A single data source → cross-domain data sources
- Separate data mining algorithms → ML + data management
- Visual and interactive data analytics



Spatio-Temporal Data is Unique

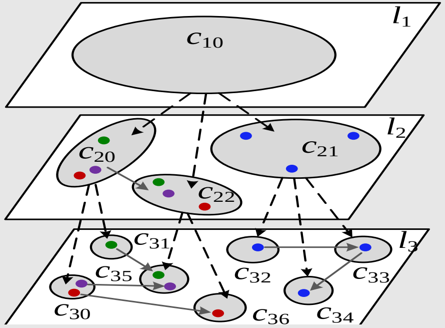
- Spatial property

Spatial closeness



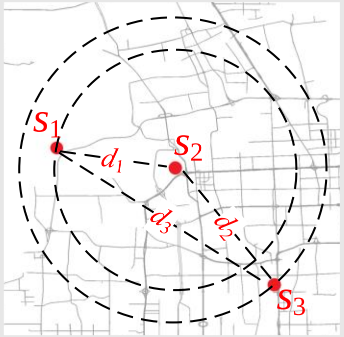
Describing correlations

Spatial hierarchy



Structural constraints between different spatial granularity

Spatial distance

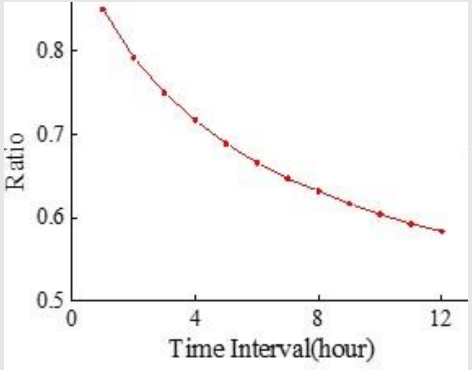


Triangle inequality:
$|d_1 - d_2| \le d_3 \le |d_1 + d_2|$

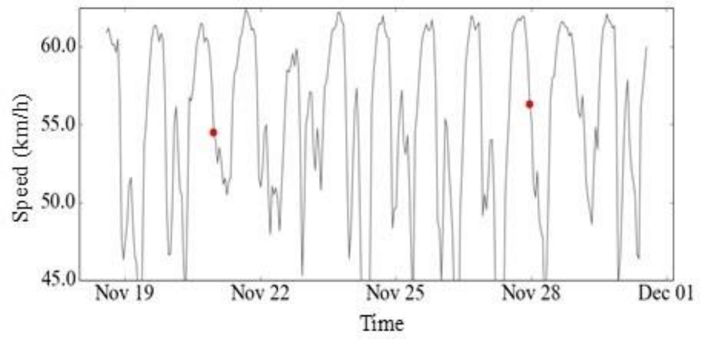
Spatio-Temporal Data is Unique

- Temporal property

Closeness

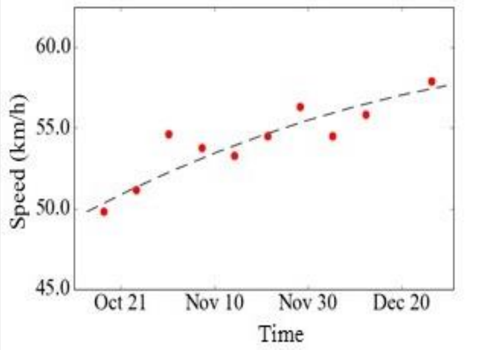


Periodicity



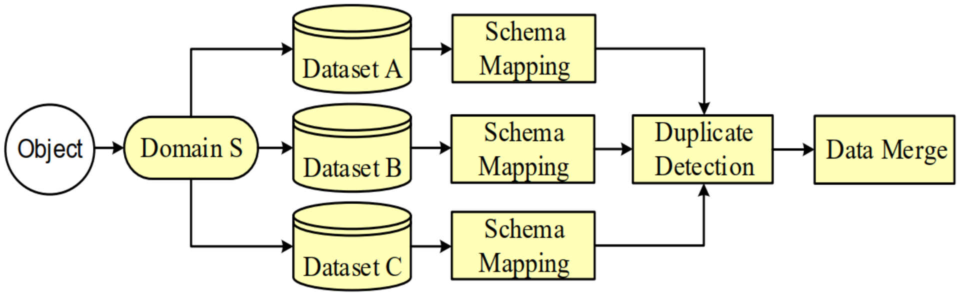
A) Hourly traffic speed on consecutive days

Trend

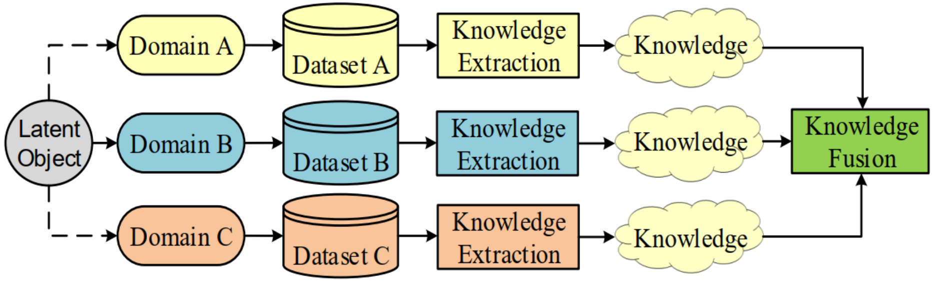


B) Traffic speed at 9-10am on consecutive weekends

Data Integration vs Knowledge Fusion



A) Paradigm of the conventional data fusion



Cross-Domain Data Fusion

Yu Zheng, et al. Methodologies for Cross-Domain Data Fusion: An Overview, IEEE Transaction on Big Data

Cross-Domain Data Fusion

- Case 1: Unlock the power from multiple (**sparse**) data across **different domains**

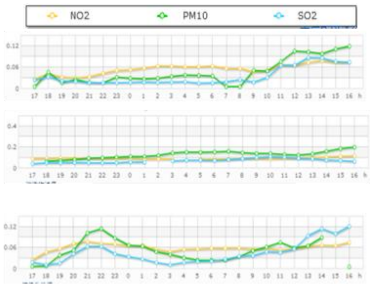
Meteorology

Traffic

Human Mobility

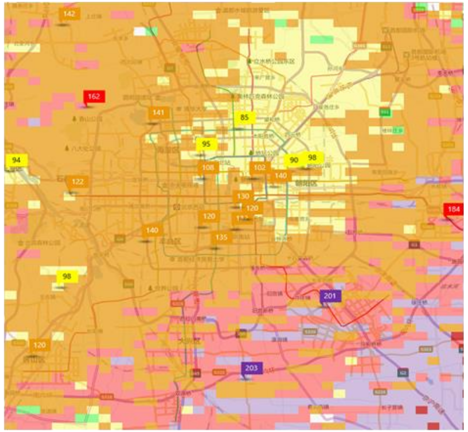
POIs

Road networks



Historical air quality data

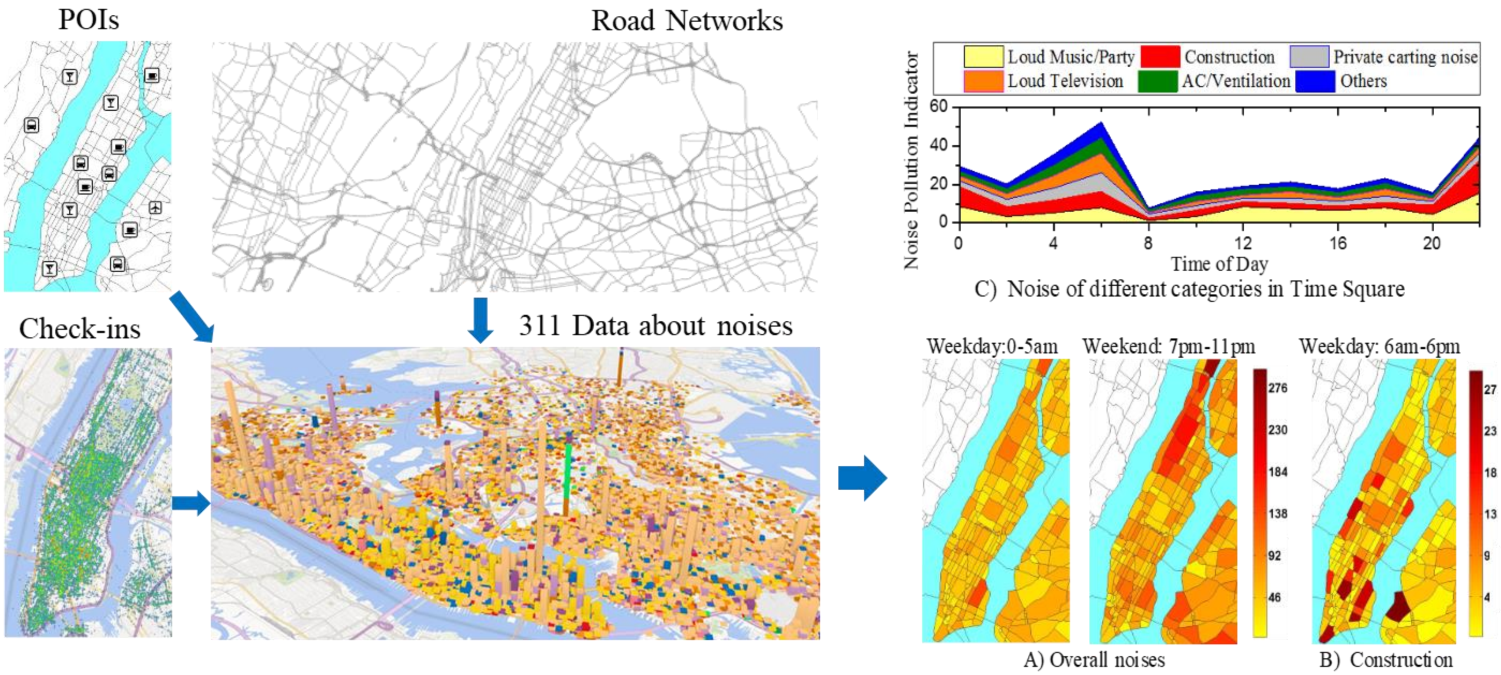
Real-time air quality reports



Yu Zheng, et al. U-Air: When Urban Air Quality Inference Meets Big Data, KDD 2013

Cross-Domain Data Fusion

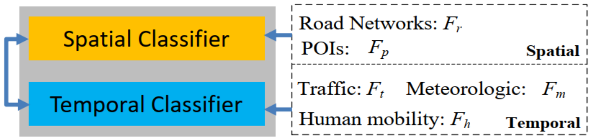
- Case 2: Unlock the power from multiple (**sparse**) data across **different domains**



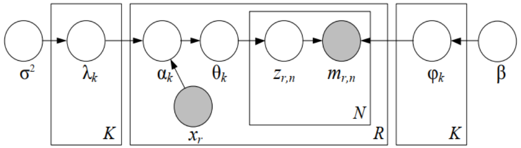
Yu Zheng, et al. Diagnosing New York City's Noises with Ubiquitous Data, Ubicomp 2014

Methodologies for Cross-Domain Data Fusion

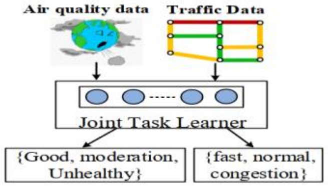
- Stage-based data fusion
- Feature-level-based data fusion
  - Feature concatenation + regularization
  - DNN-based
- Semantic meaning-based fusion
  - Multiple-view-based: like co-training
  - Similarity-based: Coupled matrix factorization
  - PGM-based
  - Transfer learning-based



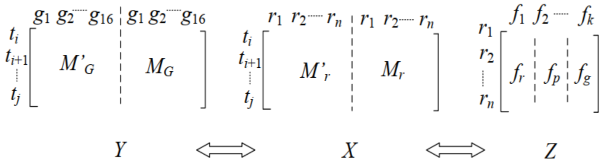
Multi-view learning (Co-training)



Pro. dependency-based (Topic Models)



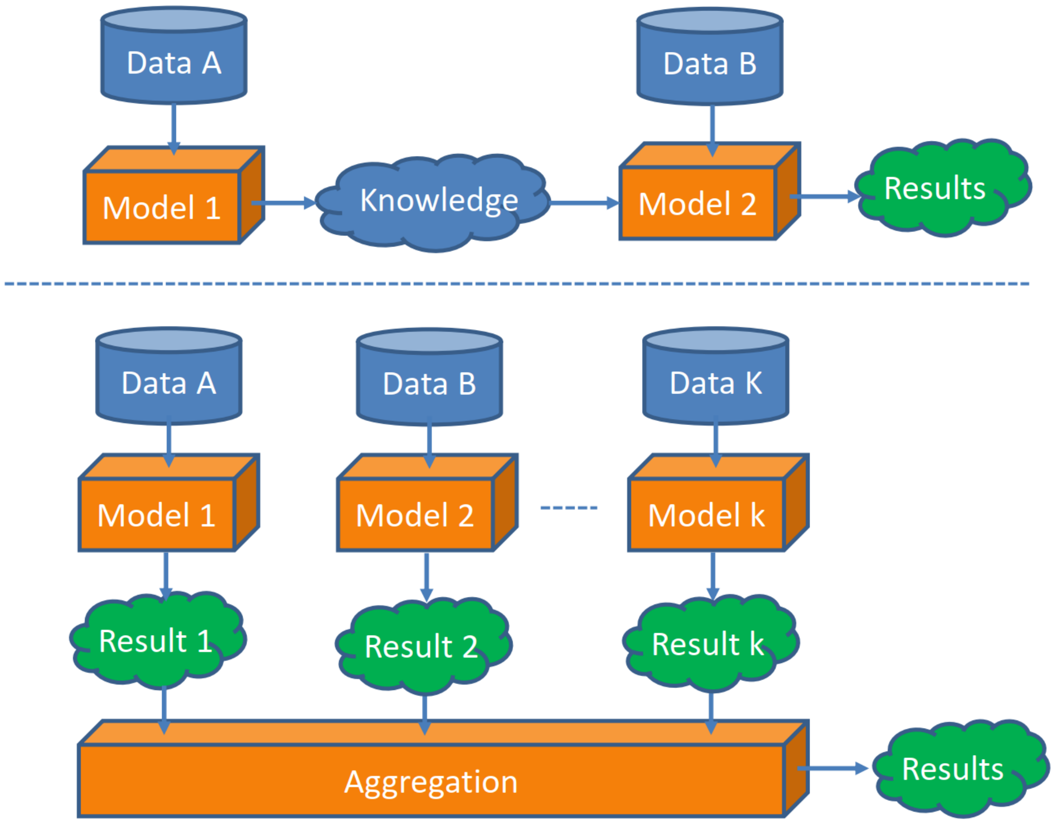
Transfer Learning-based



Similarity-based (matrix factorization)

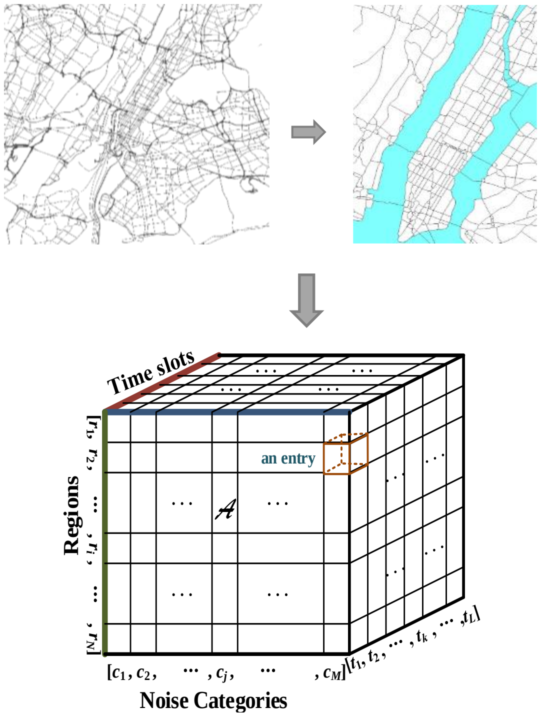
Yu Zheng, et al. Methodologies for Cross-Domain Data Fusion: An Overview, IEEE Transaction on Big Data

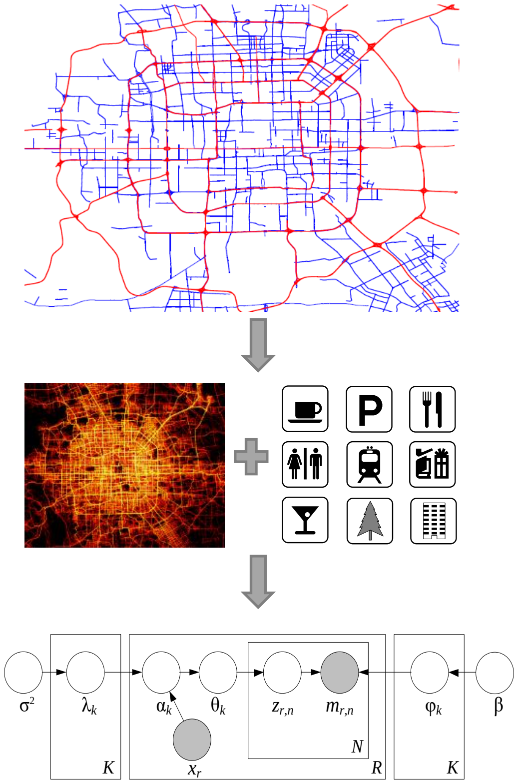
Stage-Based Data Fusion



Methodologies for Cross-Domain Data Fusion

- Stage-based fusion





Yu Zheng, et al. Urban Computing with Taxicabs, In Proc. Of **UbiComp 2011**

Urban Computing for Urban Planning

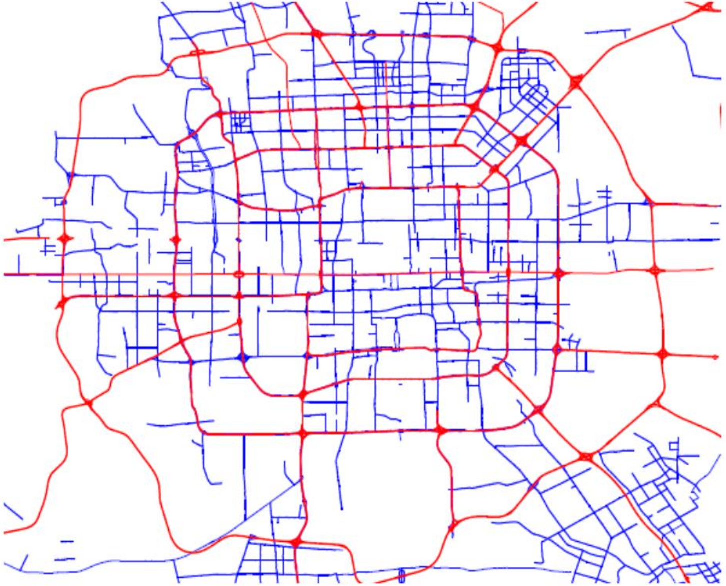
Best Paper Nominee Award at UbiComp 2011

The Most Cited Paper

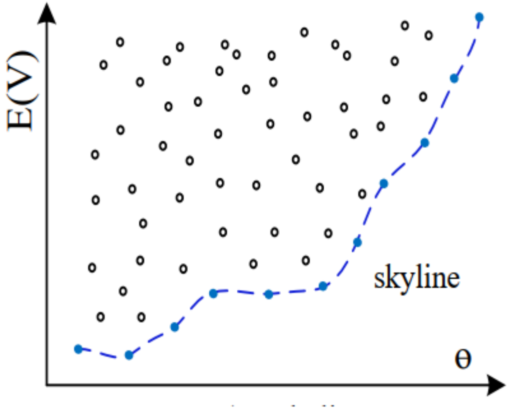
City-Wide Traffic Modeling

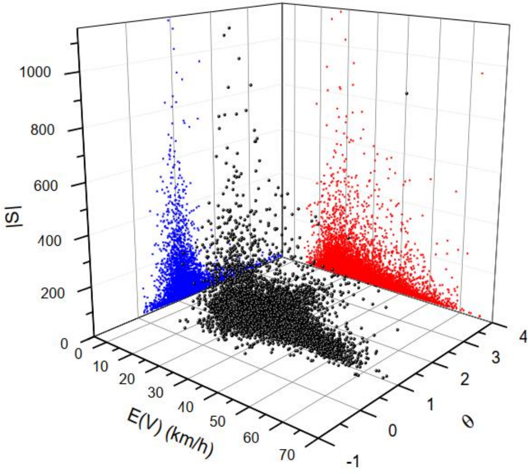
- Partition a city into regions with major roads
- Regions are root causes of the problem



City-Wide Traffic Modeling

- Taxi trajectories onto these regions
- Building a region graph for each time slot
- Find the skyline points (problematic region pairs)



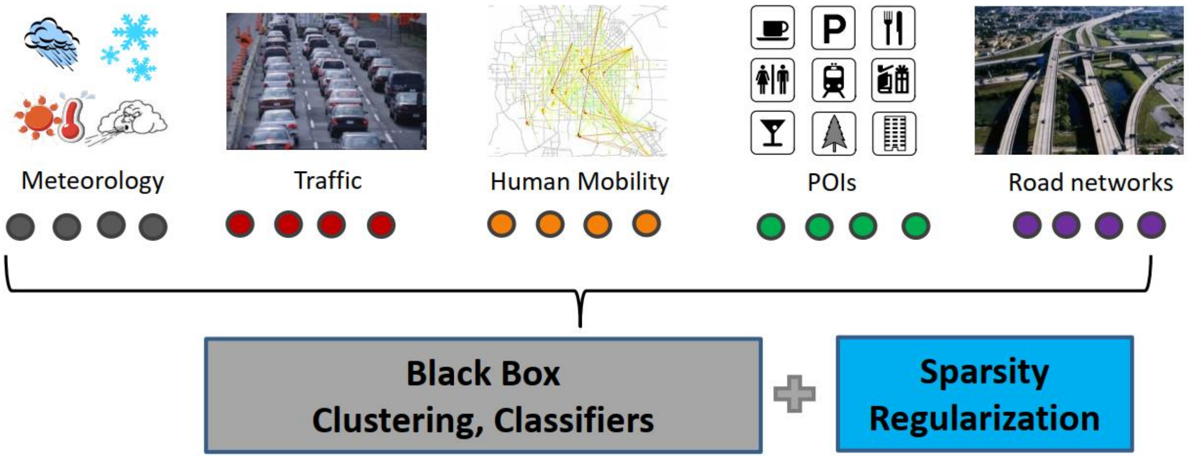


Feature-based Data Fusion

- Feature Concatenation + Regularization
- Deep Learning-Based Fusion

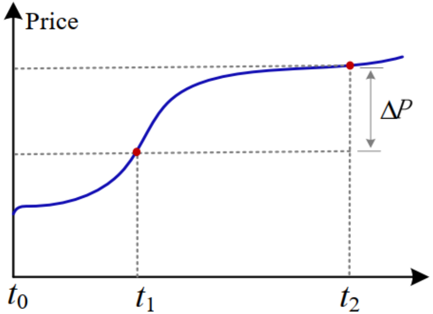
Feature Concatenation + Regularization

- Straightforward Concatenation
  - Overfitting a with small training set
  - Difficult to explore the non-linear relationship between raw features
  - Redundancy and dependency between features



Ranking Real Estates using Big Data

- Values (learned from big data)
  - Increase more in a rising market
  - Decrease less in a falling market

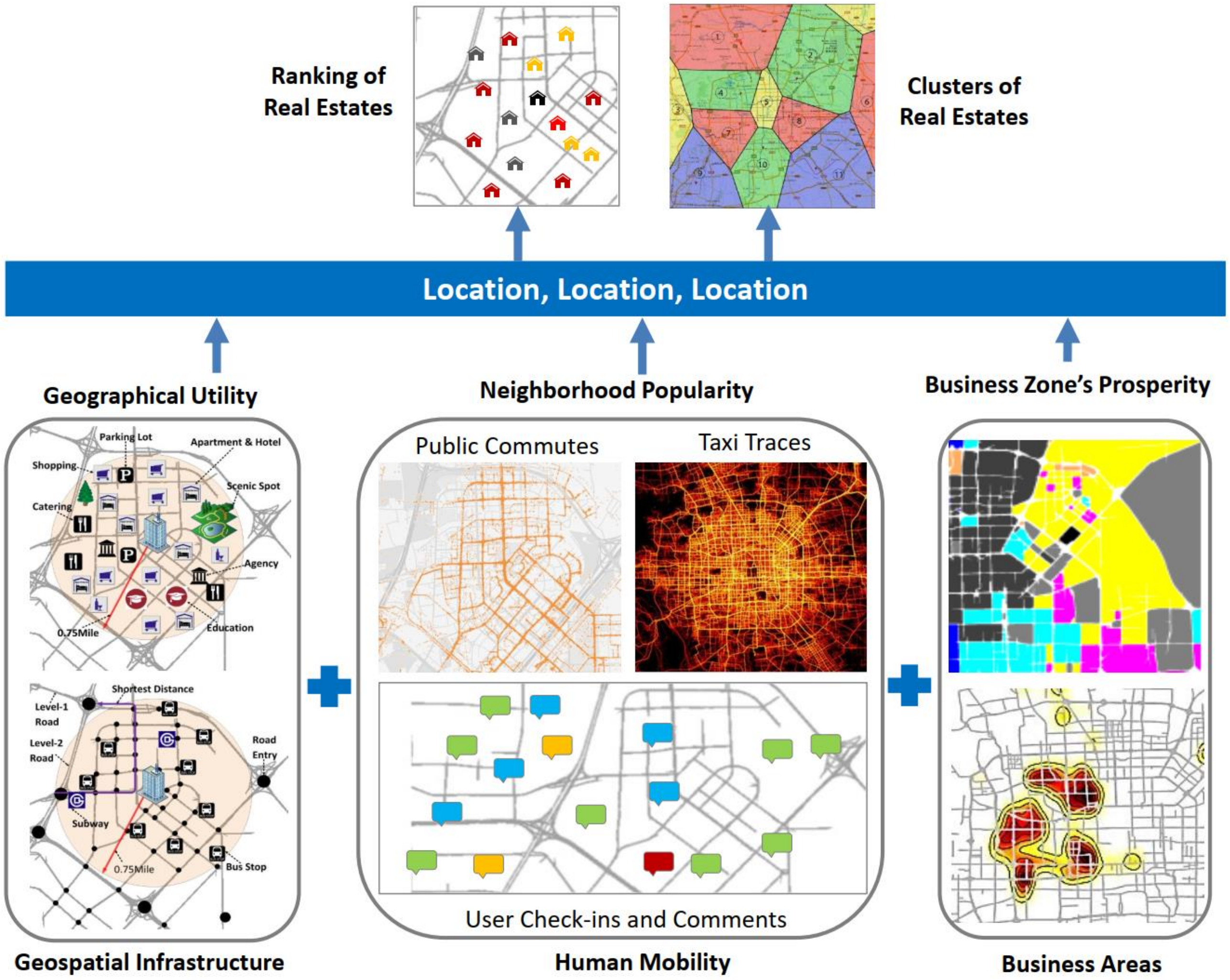


A) The price of a real estate

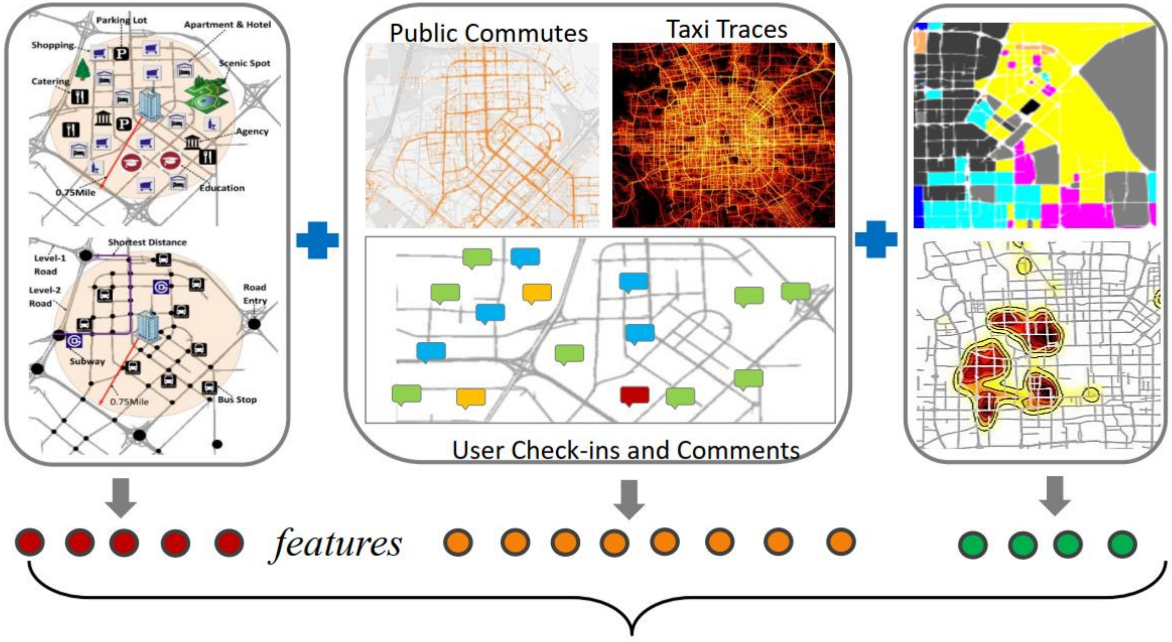
House	Increase	Rank↓
H1	35%	R1
H5	29%	R1
H4	13%	R2
....	....	....
H2	9%	R3
H3	2%	R3
H6	-1.5%	R4
H7	-6.1%	R5

B) Rank of estates by ΔP

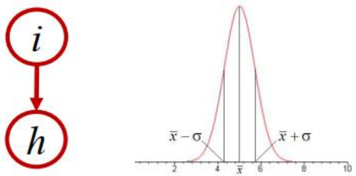
Yanjie Fu et al., [Sparse Real Estate Ranking with Online User Reviews and Offline Moving Behaviors](), ICDM 2014



Ranking of Real Estates
Clusters of Real Estates
Location, Location, Location
Geographical Utility
Neighborhood Popularity
Business Zone's Prosperity
Parking Lot
Apartment & Hotel
Shopping
Scenic Spot
Catering
Agency
Education
0.75Mile
Shortest Distance
Level-1 Road
Level-2 Road
Road Entry
Subway
Bus Stop
0.75Mile
Public Commutes
Taxi Traces
User Check-ins and Comments
Geospatial Infrastructure
Human Mobility
Business Areas



$$f_i(\boldsymbol{x_i};\boldsymbol{w}) = \boldsymbol{w}^\top \boldsymbol{x_i} + \epsilon_i = \sum_{m=1}^{M} w_m \overset{x}{x_{im}} + \epsilon_i$$



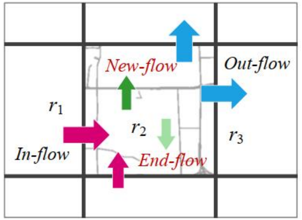
$$P\left(y_i|\boldsymbol{x_i}\right) = \mathcal{N}(y_i|f_i,\sigma^2) = \mathcal{N}(y_i|\boldsymbol{w}^\top\boldsymbol{x_i},\sigma^2)$$

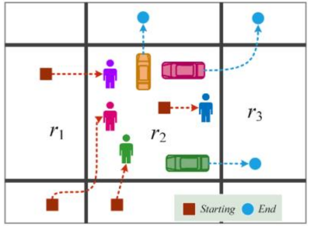
$$P\left(y_i|\boldsymbol{x_i}\right) = \prod_{i=1}^{I}\mathcal{N}(y_i|f_i,\sigma^2)\underbrace{\prod_{i=1}^{I-1}\prod_{h=i+1}^{I}P(i\rightarrow h|\Psi,\Omega)}_{\text{Pair-wised ranking constraint}}\ \underbrace{\prod_{m=1}^{M}\mathcal{N}(w_m|0,\beta_m^2)}_{\text{Sparsity Regularization}}$$

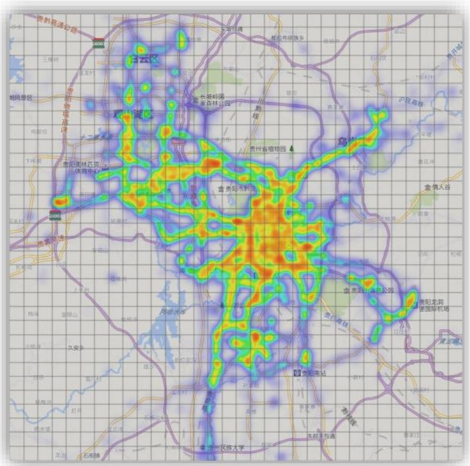
DNN-based Data Fusion

- Will be illustrated in detail in the second half of this course

Predict In-flow and out-flow of crowds in each region at next time interval throughout a city







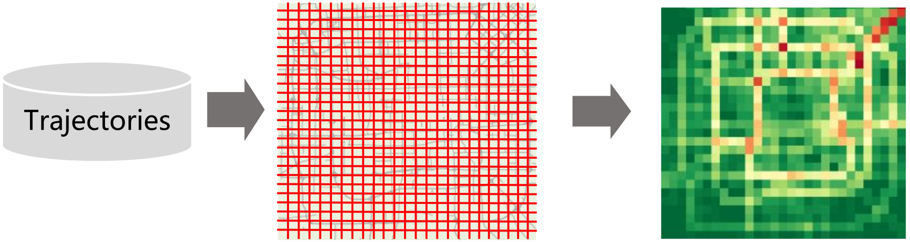
- Important for:
  - Traffic management
  - Risk assessment
  - Public safety

http://urbanflow.sigkdd.com.cn/

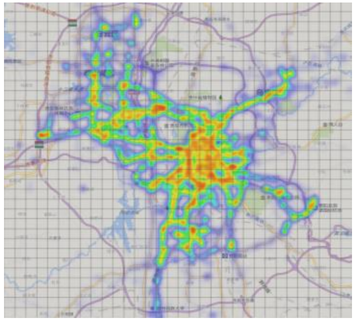
Junbo Zhang et al., Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction, AAAI 2017

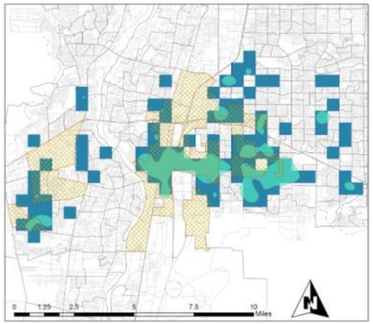
Problem Statement

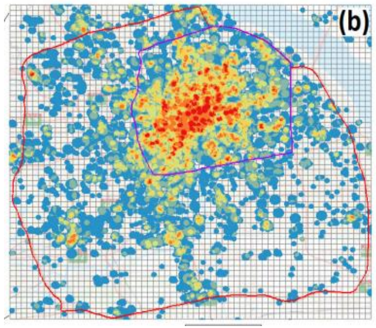
- We partition an area of interest (e.g., a metropolitan) evenly into grid cells, leading to an image-like data format called **ST grid**
  - A pixel → A region
  - RGB → Observations / Attributes



- Real-world examples

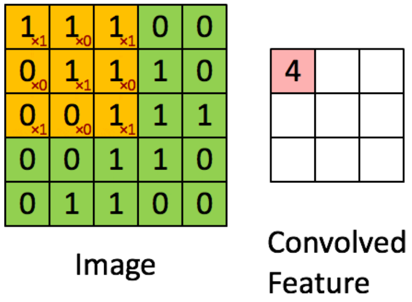

Taxi flows

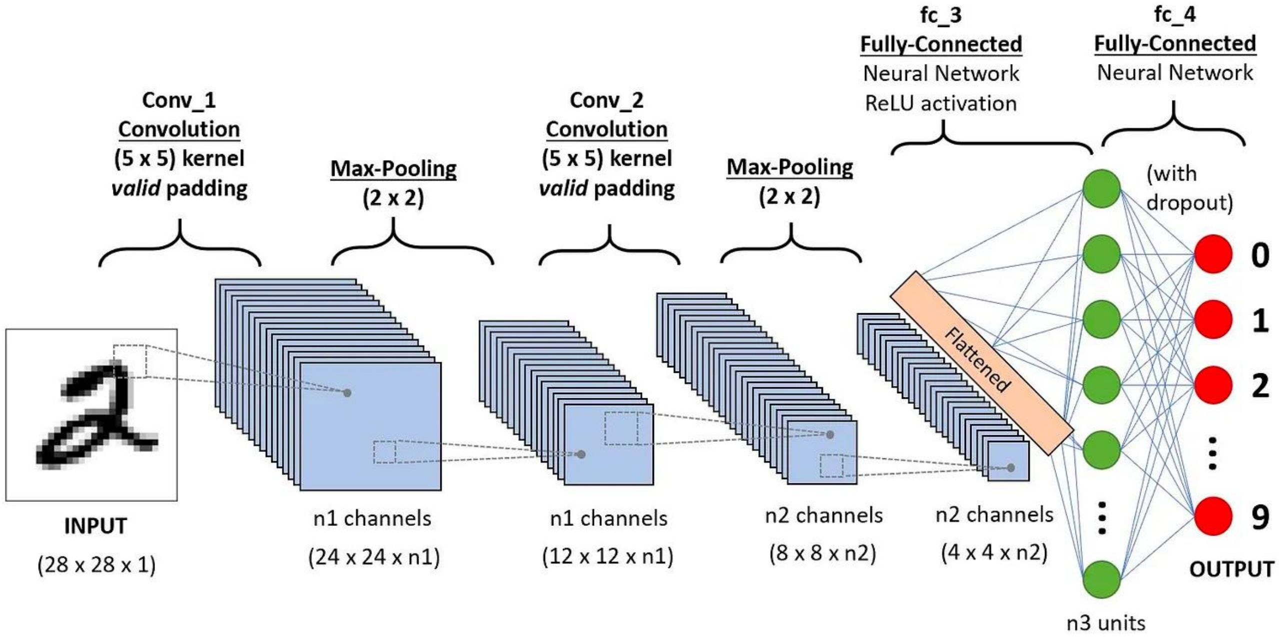

Crime hotspots



Bike-sharing demands

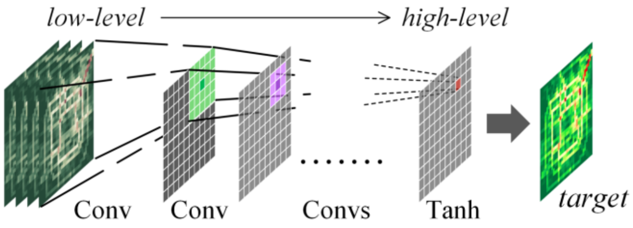
Preliminary: Convolutional Neural Network (CNN)

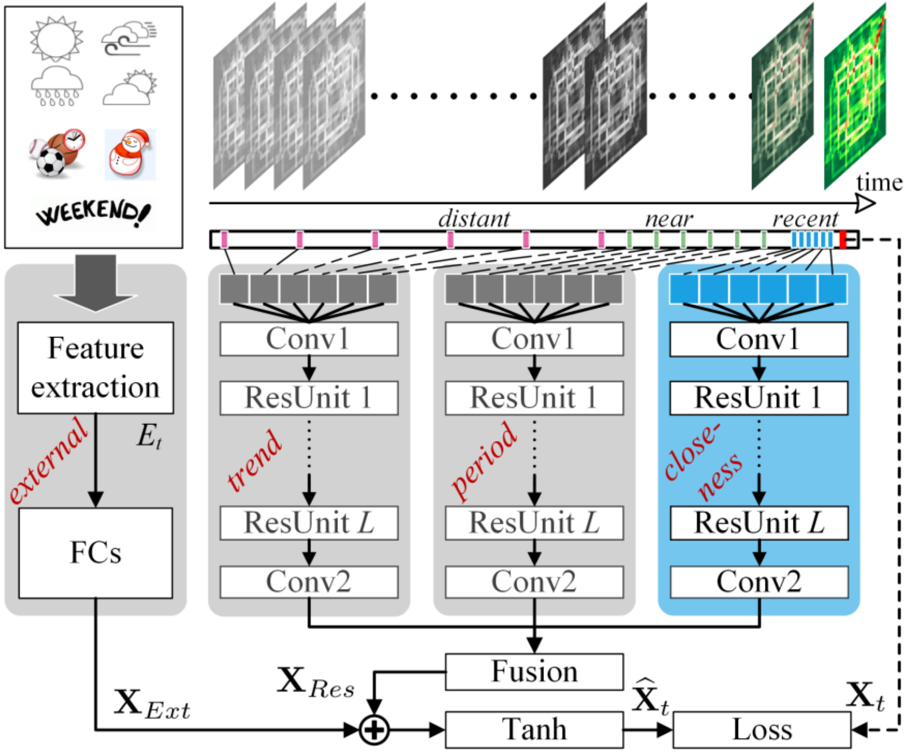




ST-ResNet

- Temporal dependencies
  - Distant
  - Near
  - Recent
- Spatial dependencies





J. Zhang et al., Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. AAAI 2017.

Experiments

- Datasets
  - TaxiBJ
  - BikeNYC

Table 1: Datasets (holidays include adjacent weekends).

Dataset	TaxiBJ	BikeNYC
Data type	Taxi GPS	Bike rent
Location	Beijing	New York
Time Span	7/1/2013 - 10/30/2013 3/1/2014 - 6/30/2014 3/1/2015 - 6/30/2015 11/1/2015 - 4/10/2016	4/1/2014 - 9/30/2014
Time interval	30 minutes	1 hour
Gird map size	(32, 32)	(16, 8)
Trajectory data		
Average sampling rate (s)	~ 60	\
# taxis/bikes	34,000+	6,800+
# available time interval	22,459	4,392
External factors (holidays and meteorology)		
# holidays	41	20
Weather conditions	16 types (*e.g.*, Sunny, Rainy)	\
Temperature / °C	$[-24.6, 41.0]$	\
Wind speed / mph	$[0, 48.6]$	\

Experiments

- Model comparison
- Ablation study

Table 2: Comparison among different methods on TaxiBJ

Model		RMSE
HA		57.69
ARIMA		22.78
SARIMA		26.88
VAR		22.88
ST-ANN		19.57
DeepST		18.18
	ST-ResNet [ours]	
L2-E	2 residual units + E	17.67
L4-E	4 residual units + E	17.51
L12-E	12 residual units + E	16.89
L12-E-BN	L12-E with BN	**16.69**
L12-single-E	12 residual units (1 conv) + E	17.40
L12	12 residual units	17.00
L12-E-noFusion	12 residual units + E without fusion	17.96

Table 3: Comparisons with baselines on BikeNYC. The results of ARIMA, SARIMA, VAR and 4 DeepST variants are taken from (Zhang et al. 2016).

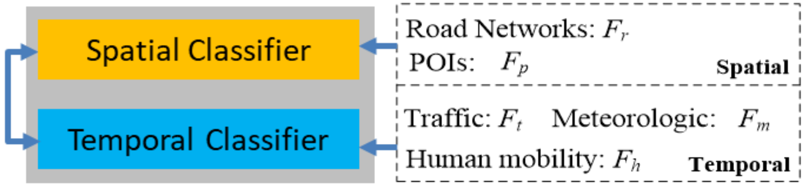
Model	RMSE
ARIMA	10.07
SARIMA	10.56
VAR	9.92
DeepST-C	8.39
DeepST-CP	7.64
DeepST-CPT	7.56
DeepST-CPTM	7.43
ST-ResNet [ours, 4 residual units]	**6.33**

Semantic-based Data Fusion

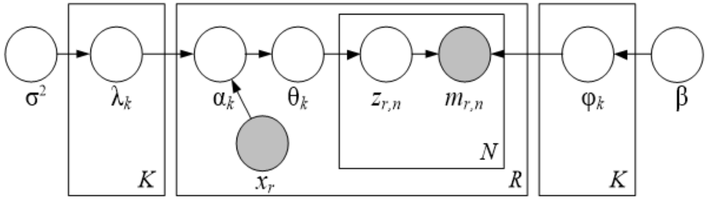
- Multiple-view-based: co-training
- Similarity-based: Coupled matrix factorization
- Probabilistic dependency-based: graphical models
- Transfer learning-based

- The insight of each dataset and relations between features across different datasets
- Derived from the ways that people think of a problem
- Interpretable and meaningful

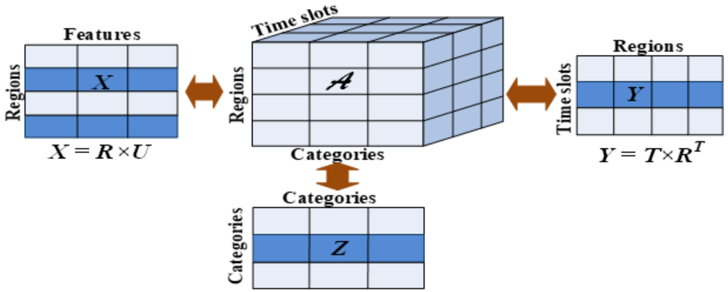
Semantic-based Data Fusion



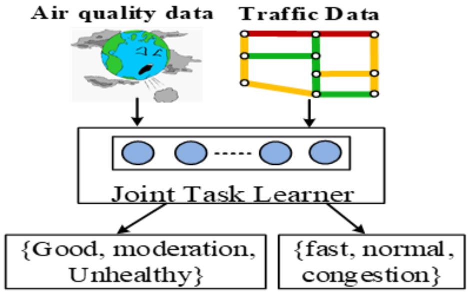
Multi-view learning (Co-training)



Pro. dependency-based (Topic Models)



Similarity-based (matrix factorization)



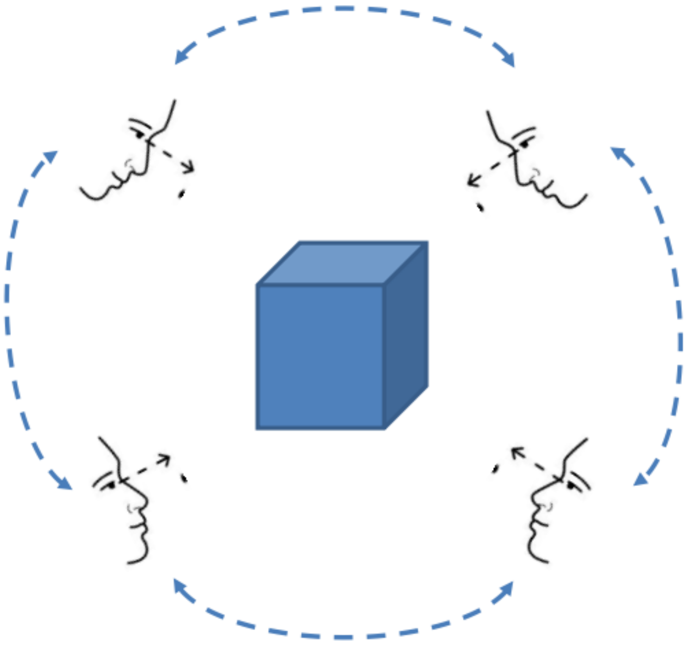
Transfer Learning

Semantic-based Data Fusion

- Multiple-view-based: co-training
- Similarity-based: Coupled matrix factorization
- Probabilistic dependency-based: graphical models
- Transfer learning-based

- The insight of each dataset and relations between features across different datasets
- Derived from the ways that people think of a problem
- Interpretable and meaningful

Multi-View Learning

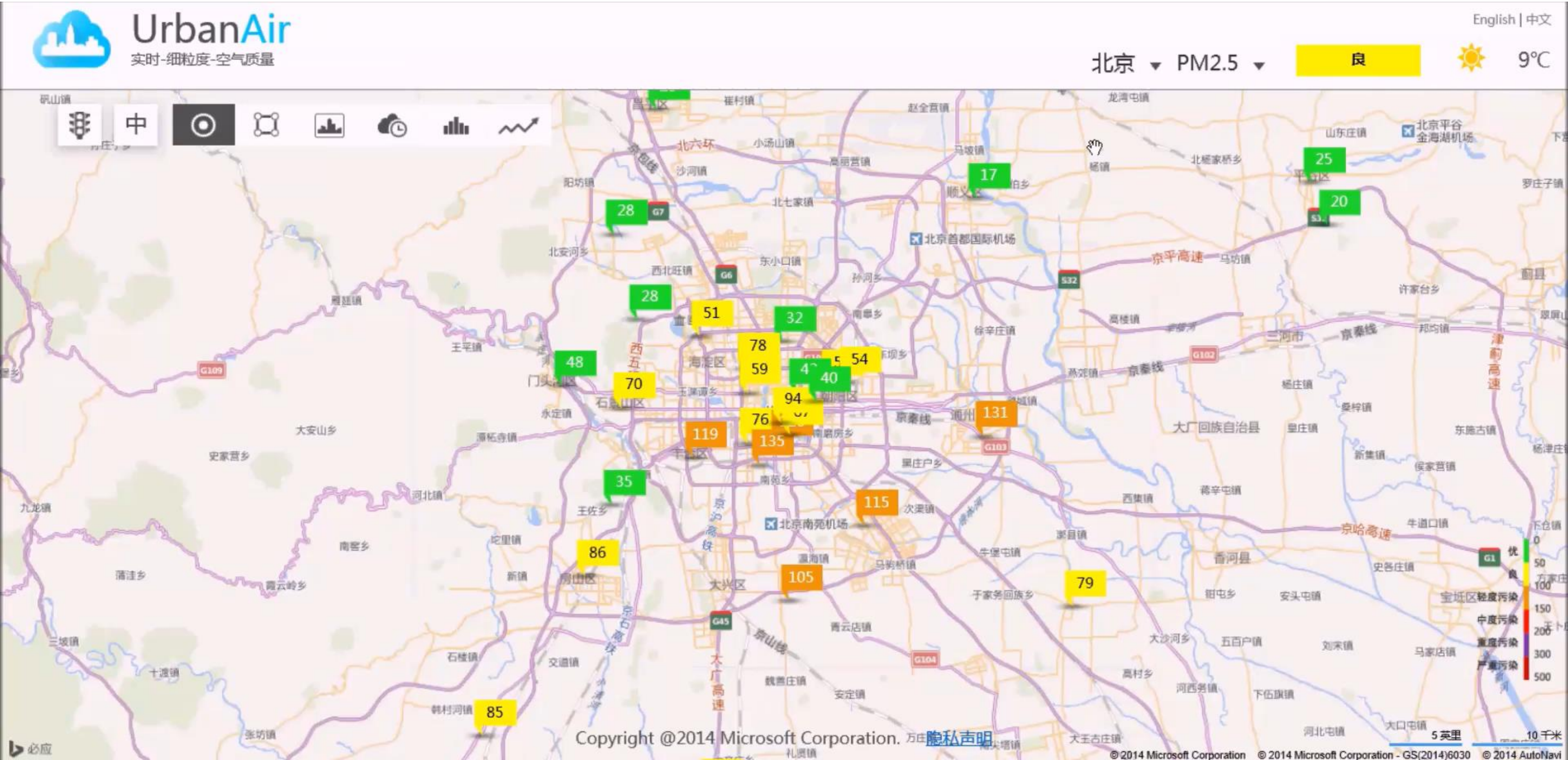


When Urban Air Meets Big Data

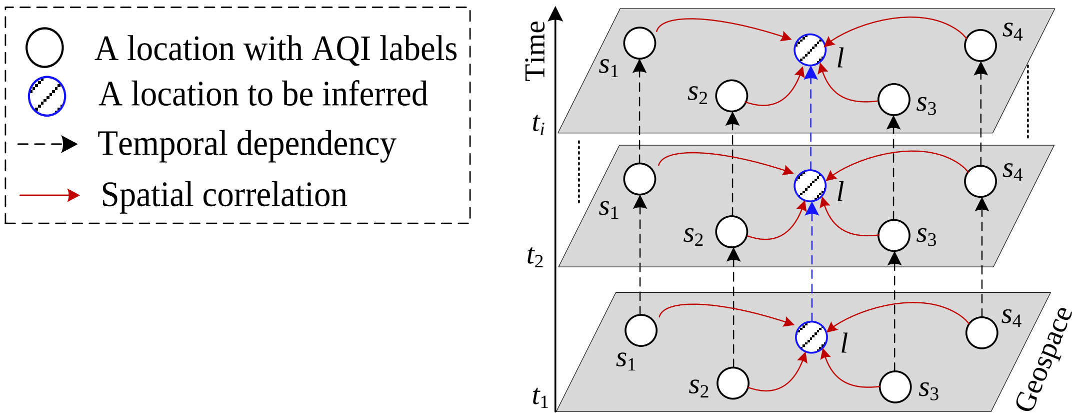
KDD 2013

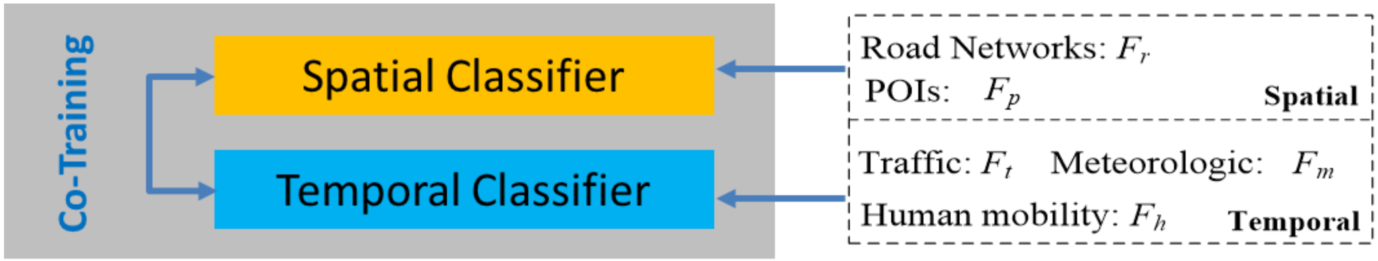
Yu Zheng, et al. [U-Air: When Urban Air Quality Inference Meets Big Data](), KDD 2013

http://urbanair.msra.cn

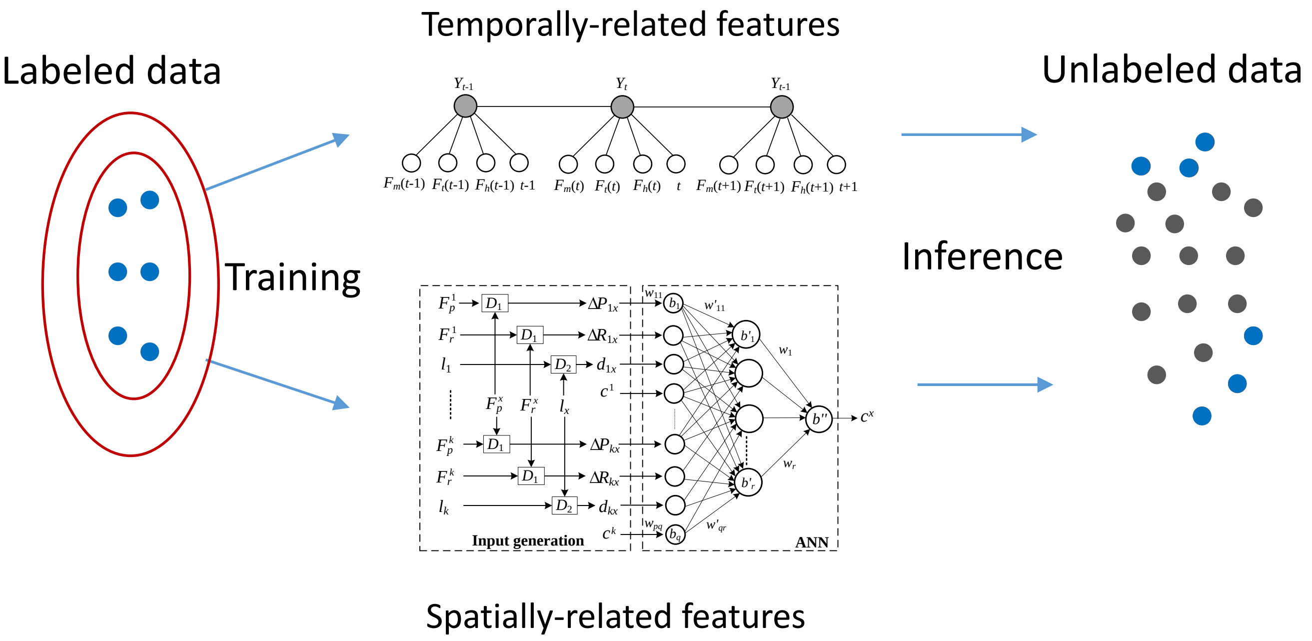


Multi-View Learning

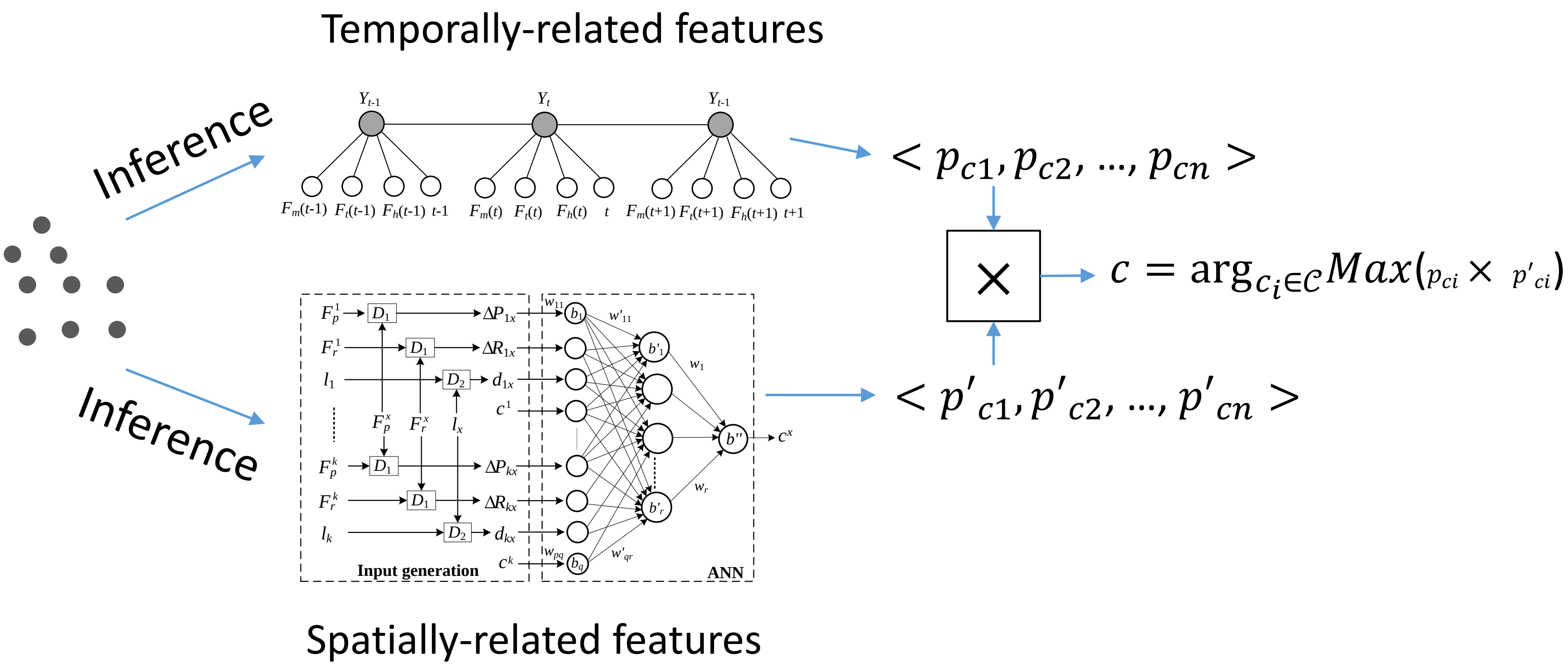




Learning Process



Inference Process

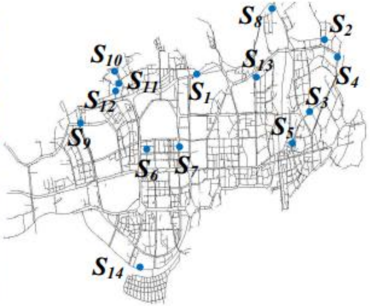


Predicting Urban Water Quality with Ubiquitous Data - A Data-driven Approach

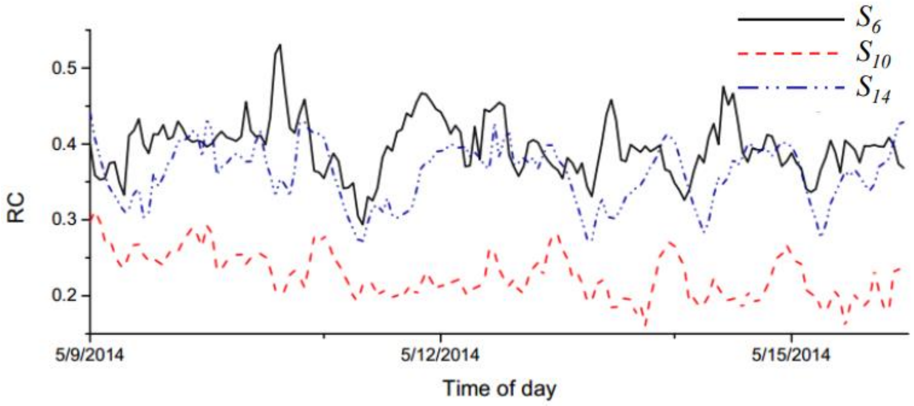
IJCAI 2016 and TBD

Background

(a) Stations on road network



(b) Stations on pipe network

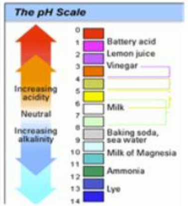


(c) Water quality readings in different stations

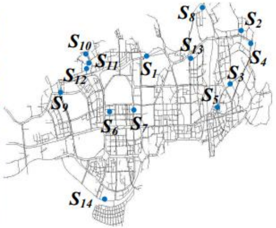
Background

- Urban Water quality is crucial to our life
  - quality index
    - **Residual Chlorine (RC)**
    - Turbidity
    - pH





(a) Stations on road network



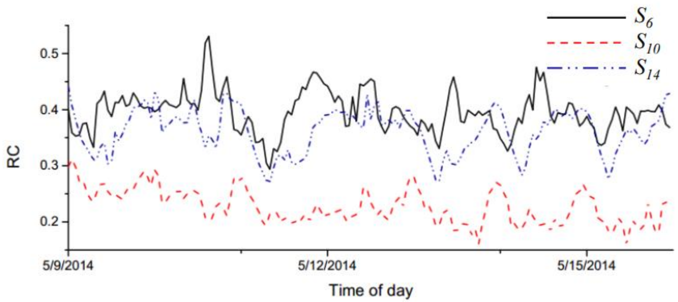
(b) Stations on pipe network

- Predicting urban water quality is of great importance to us
- Applications

Real-time monitoring

water pollution alarming

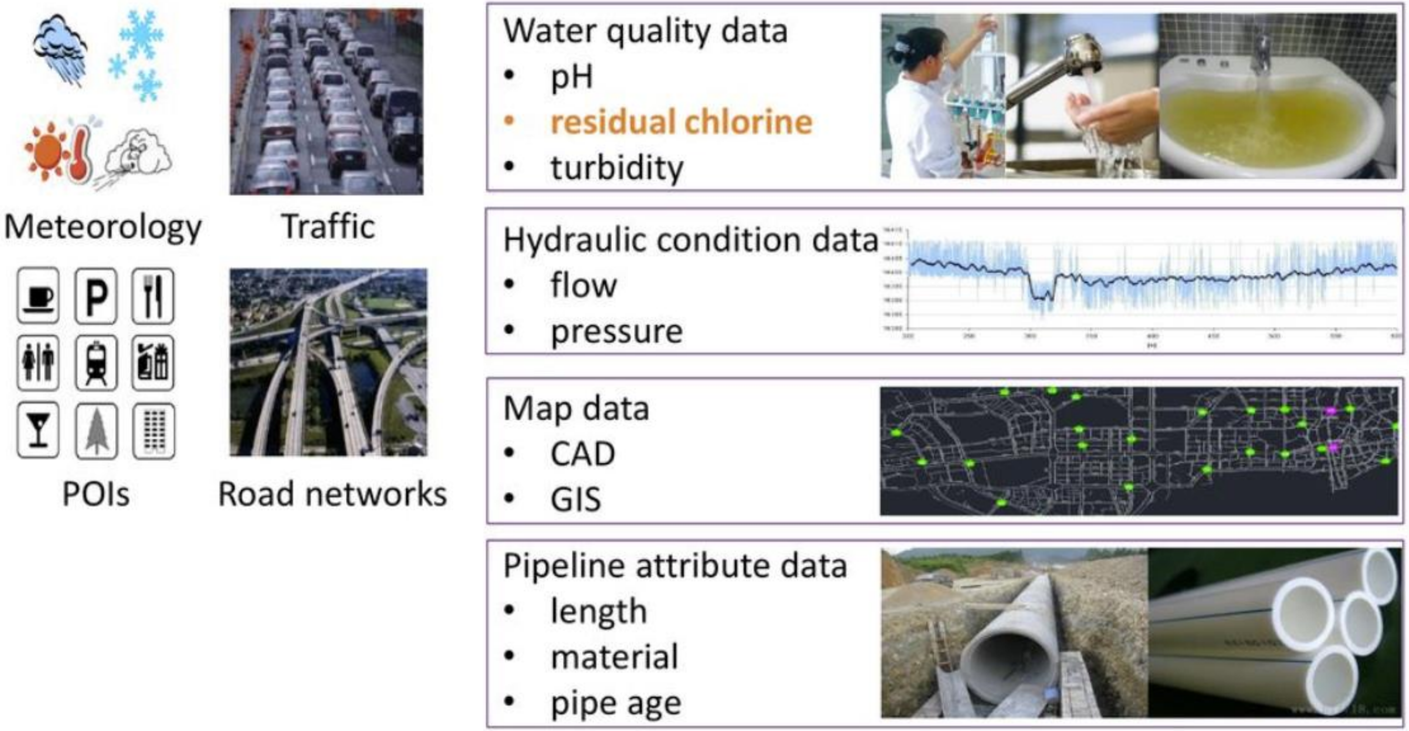
Suggestions for replacements



(c) Water quality readings in different stations

Target

- Predicting the Urban Water Quality from Multi-sources Urban Data



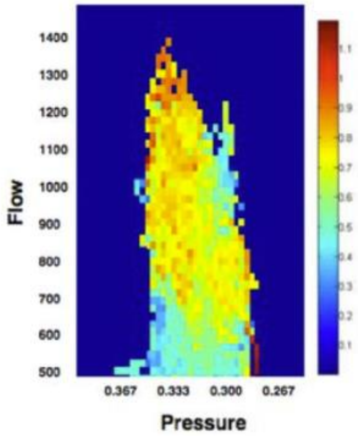
Challenges

- Unknown influential factors that affect water quality
  - Turbidity
  - Flow
  - POIs
  - …
- Water quality varies over time and location non-linearly

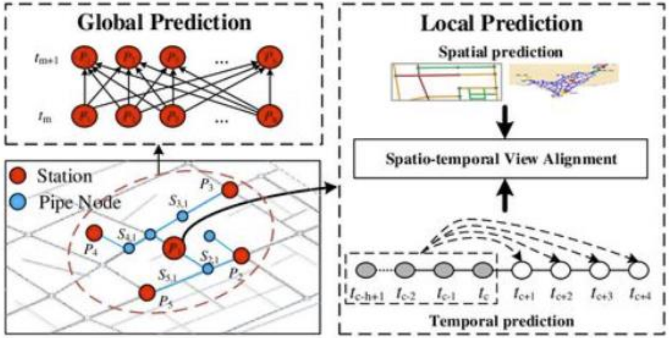
Solutions

- Identifying influencing factors
  - Flow, turbidity, pH, etc.

- Approaching from spatial and temporal perspectives
  - Multi-views: each station has two views:
    - spatial view and temporal view
  - Capturing local information of each station

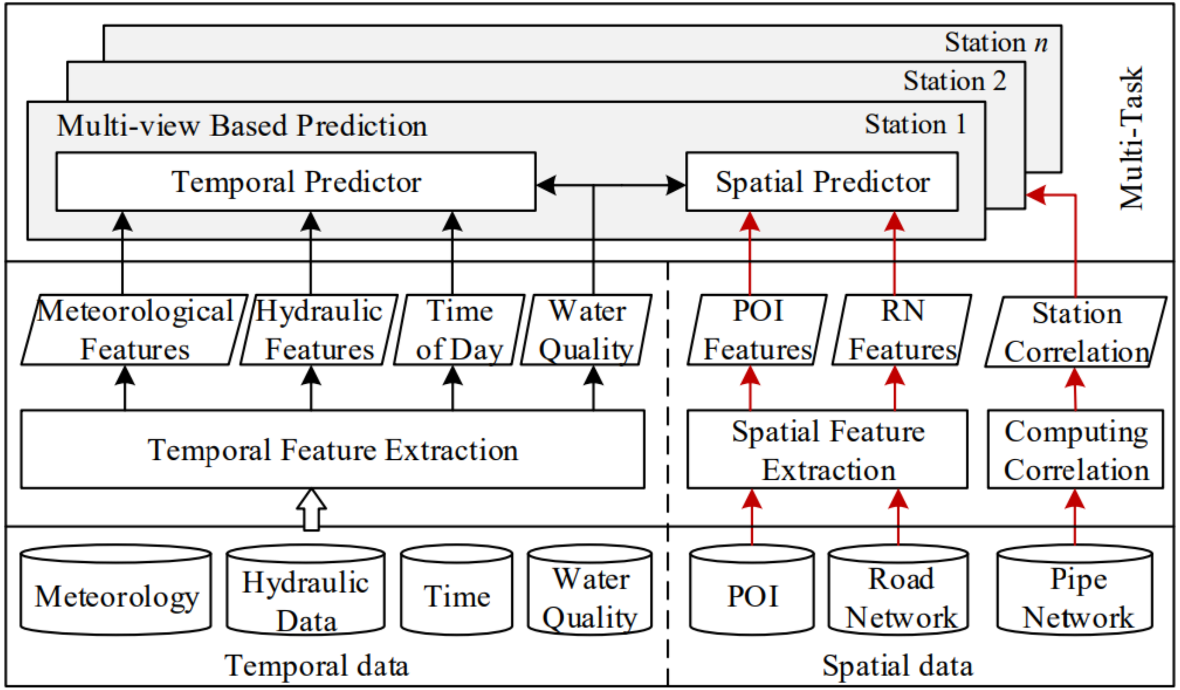


- Approaching from local and global perspectives
  - Multi-tasks: water quality prediction at each station
  - Capturing the global correlations among stations



Multi-Task Multi-View Learning Framework

- Task = Sensor, View = Spatial and Temporal



Methodology

- Multi-task Multi-view Learning
  - Multi-Views: For each station, there are two views
    - Spatial view: predictions based on its neighbors
    - Temporal view: predictions based on its own history
    - Alignment between two views

$$\lambda \sum_{l=1}^{M} ||\mathbf{X}_l^s \mathbf{w}_l^s - \mathbf{X}_l^t \mathbf{w}_l^t||_2^2$$

  - Multi-Tasks:
    - The prediction at each station is a task
    - All stations do the co-prediction
    - Alignments among multiple tasks

$$\gamma \sum_{l,m=1}^{M} C_{l,m} ||\mathbf{w}_l - \mathbf{w}_m||_2^2$$

Multi-Task Multi-View Learning Framework

- Task = Sensor, View = Spatial and Temporal
- LR + MV + MT + Regularization

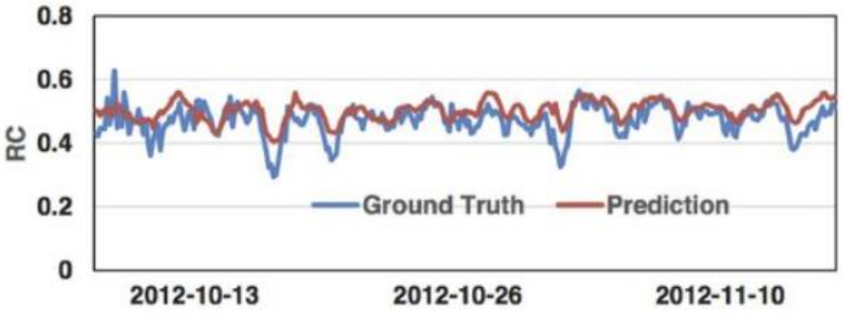
$$\min_{\mathbf{W}} \quad \frac{1}{2}\sum_{l=1}^{M} \|\mathbf{y}_l - \frac{1}{2}\mathbf{X}_l\mathbf{w}_l\|_2^2 + \lambda \sum_{l=1}^{M} \|\mathbf{X}_l^s\mathbf{w}_l^s - \mathbf{X}_l^t\mathbf{w}_l^t\|_2^2$$

$$+\gamma \sum_{l,m=1}^{M} C_{l,m}\|\mathbf{w}_l - \mathbf{w}_m\|_2^2 + \theta\|\mathbf{W}\|_{2,1},$$

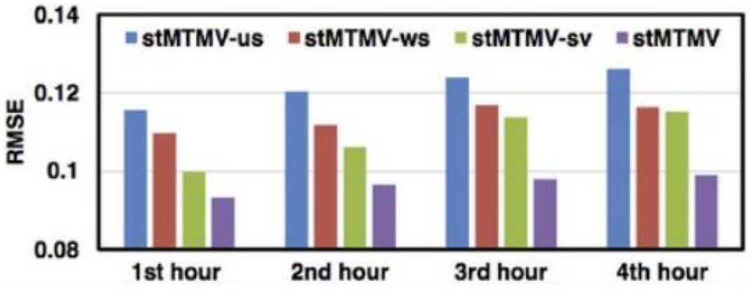
Evaluation

Models	1 hour	2 hour	3 hour	4 hour
RC Decay Model	3.51e-1	3.53e-1	3.59e-1	3.68e-1
ARMA	1.86e-1	2.18e-1	2.46e-1	2.78e-1
LR	1.68e-1	1.99e-1	2.09e-1	2.10e-1
LASSO	1.23e-1	1.42e-1	1.52e-1	1.56e-1
MRMTL	1.32e-1	1.48e-1	1.56e-1	1.58e-1
regMVMT	1.06e-1	1.15e-1	1.18e-1	1.19e-1
stMTMV	**9.33e-2**	**9.66e-2**	**9.80e-2**	**9.90e-2**

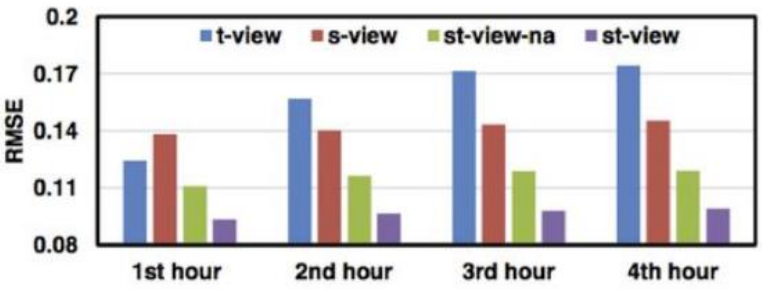
Performance comparison among various approaches



Predictive Performance



Model components comparison



Views comparison

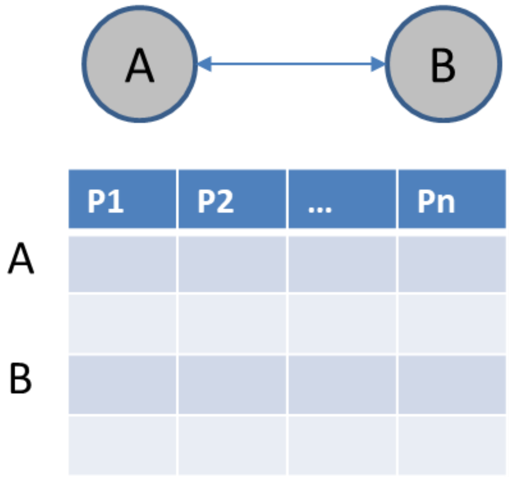
Semantic-based Data Fusion

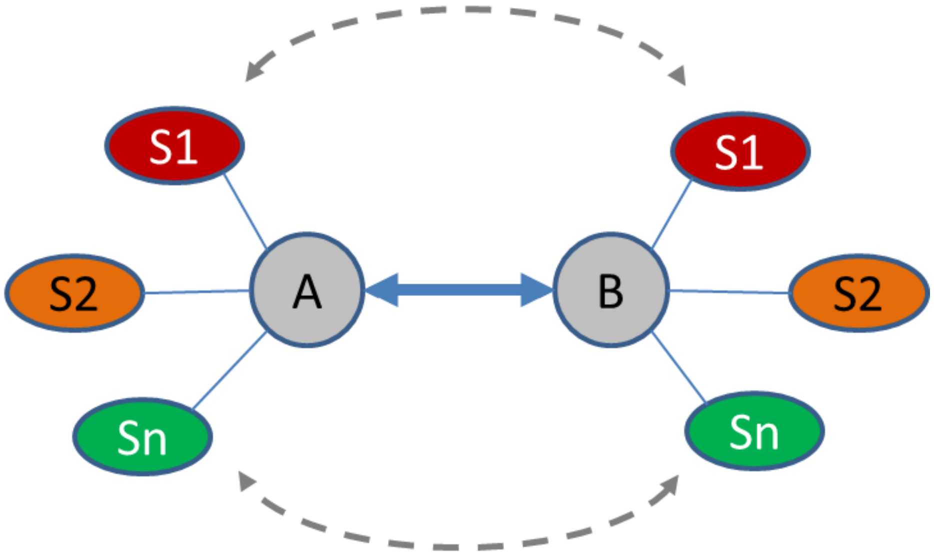
- Multiple-view-based: co-training
- Similarity-based: Coupled matrix factorization
- Probabilistic dependency-based: graphical models
- Transfer learning-based

- The insight of each dataset and relations between features across different datasets
- Derived from the ways that people think of a problem
- Interpretable and meaningful

Semantic-based Data Fusion

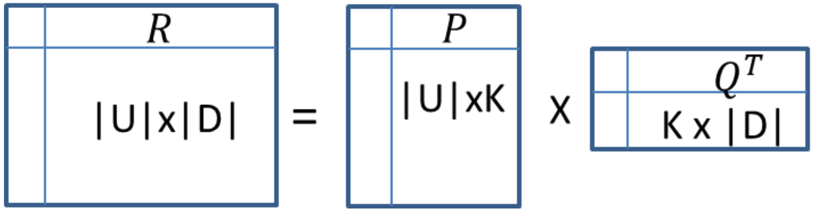
- Similarity-Based Data Fusion
  - Coupled matrix factorization
  - Context-Aware Tensor Decomposition

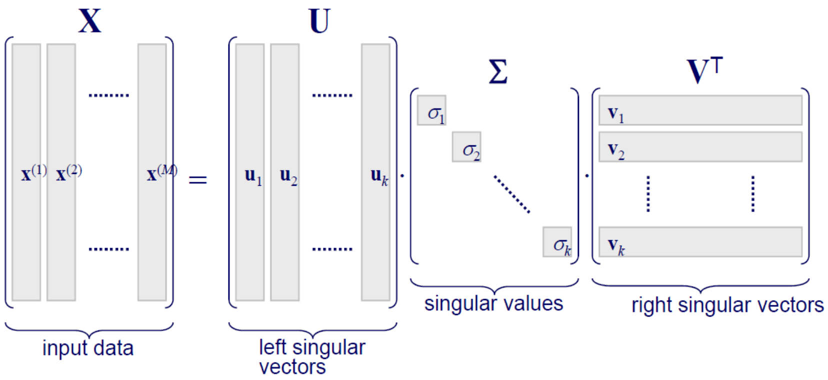




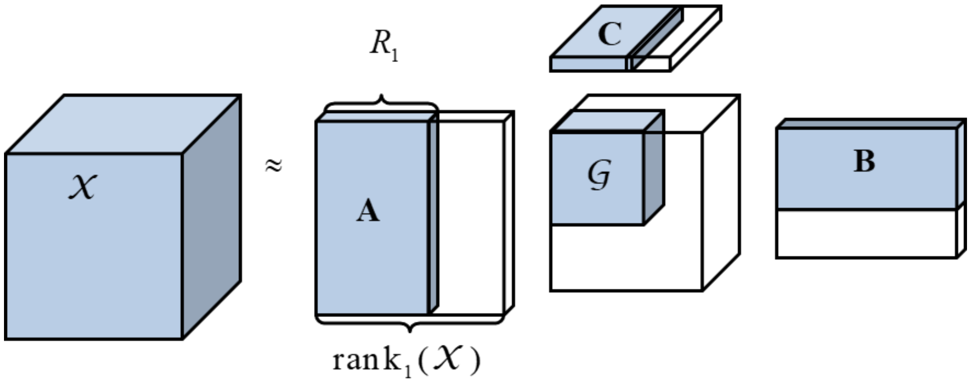
Foundations

- Matrix Factorization

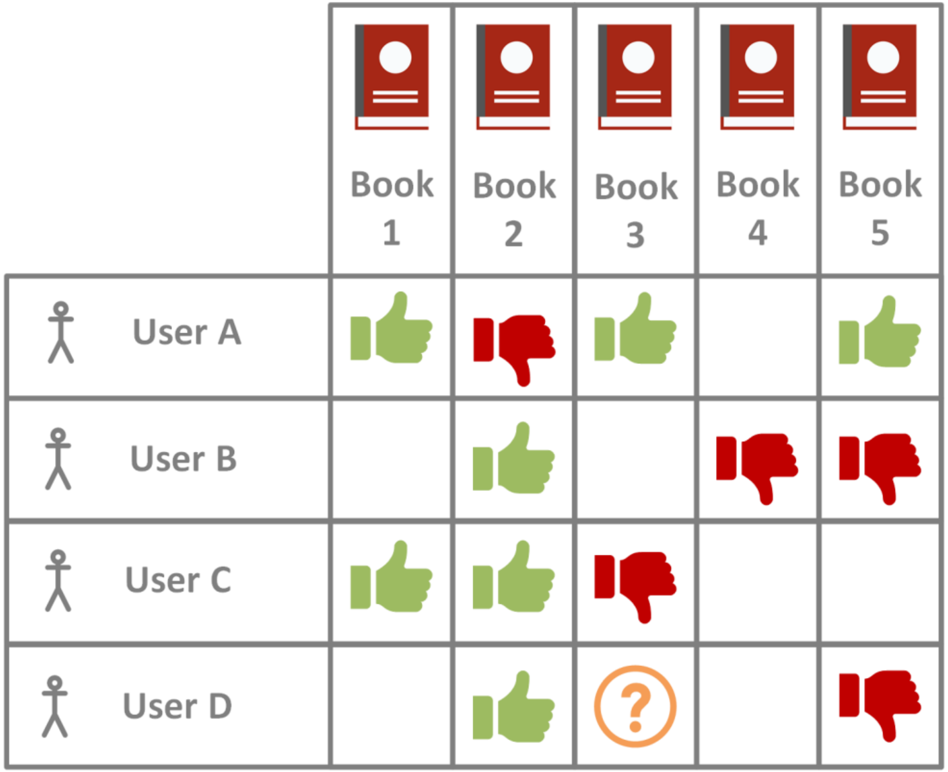




- Tucker Decomposition

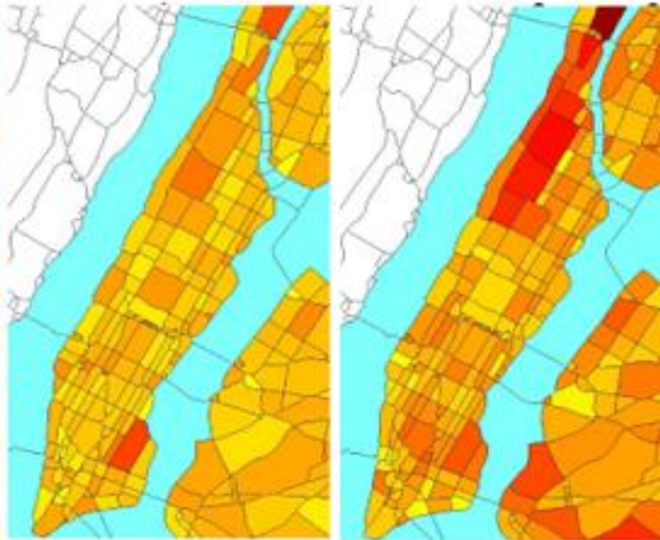


Collaborative Filtering in RecSys



Diagnosing Urban Noises using Big Data

Ubicomp 2014



Background

- Many cities suffer from noise pollutions
  - Traffic, loud music, construction, AC…
  - Compromise working efficiency
  - Reduce sleep quality
  - Impair both physical and mental health

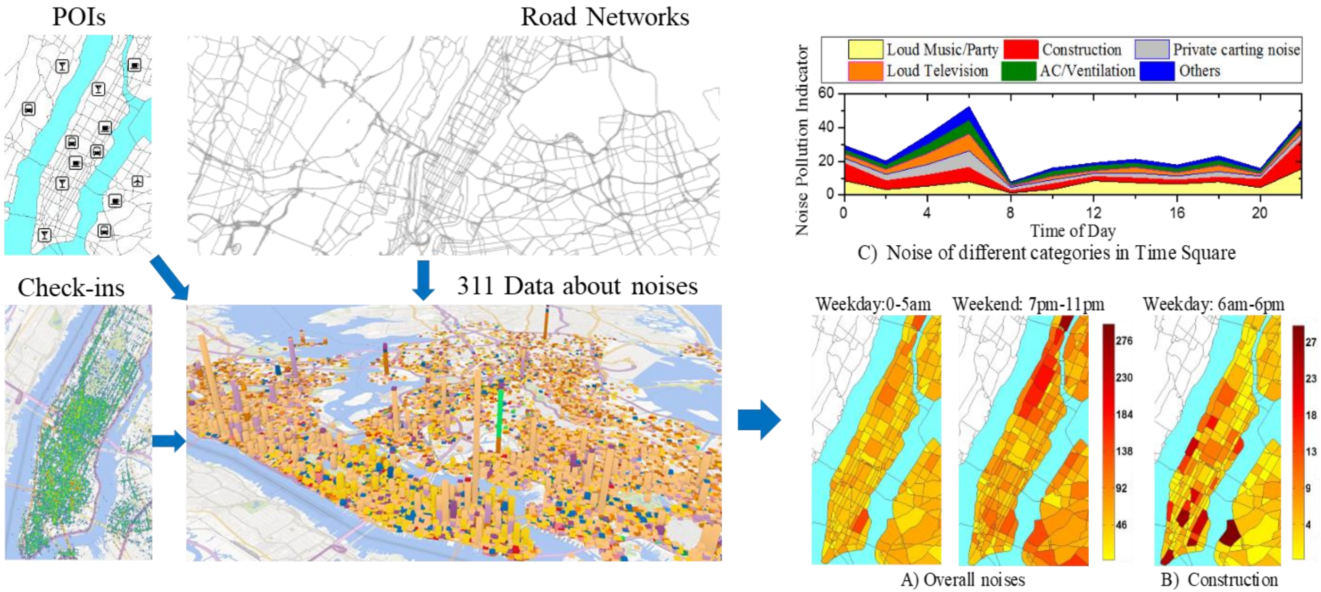
- Urban noise is difficult to model
  - Change over time very quickly
  - Vary by location significantly
  - Depends on sound levels and people’s tolerance
  - The composition of noises is hard to analyze

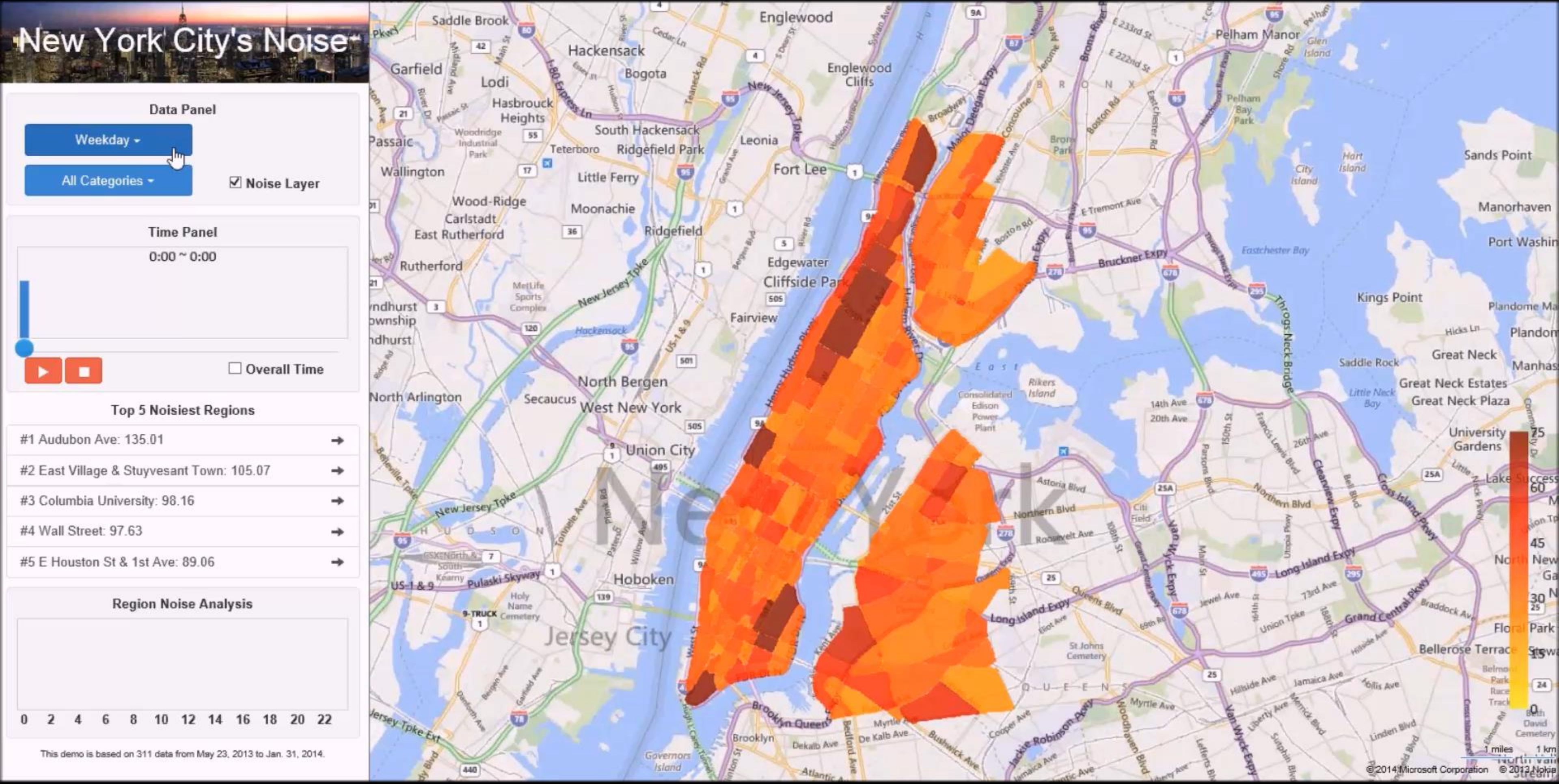
311 in NYC

- 311 Data
  - A platform for citizen's non-emergent complaints
  - Associated with a location, timestamp, and a category
  - Human as a sensor → crowd sensing
  - Implies people's reaction and tolerance to noises

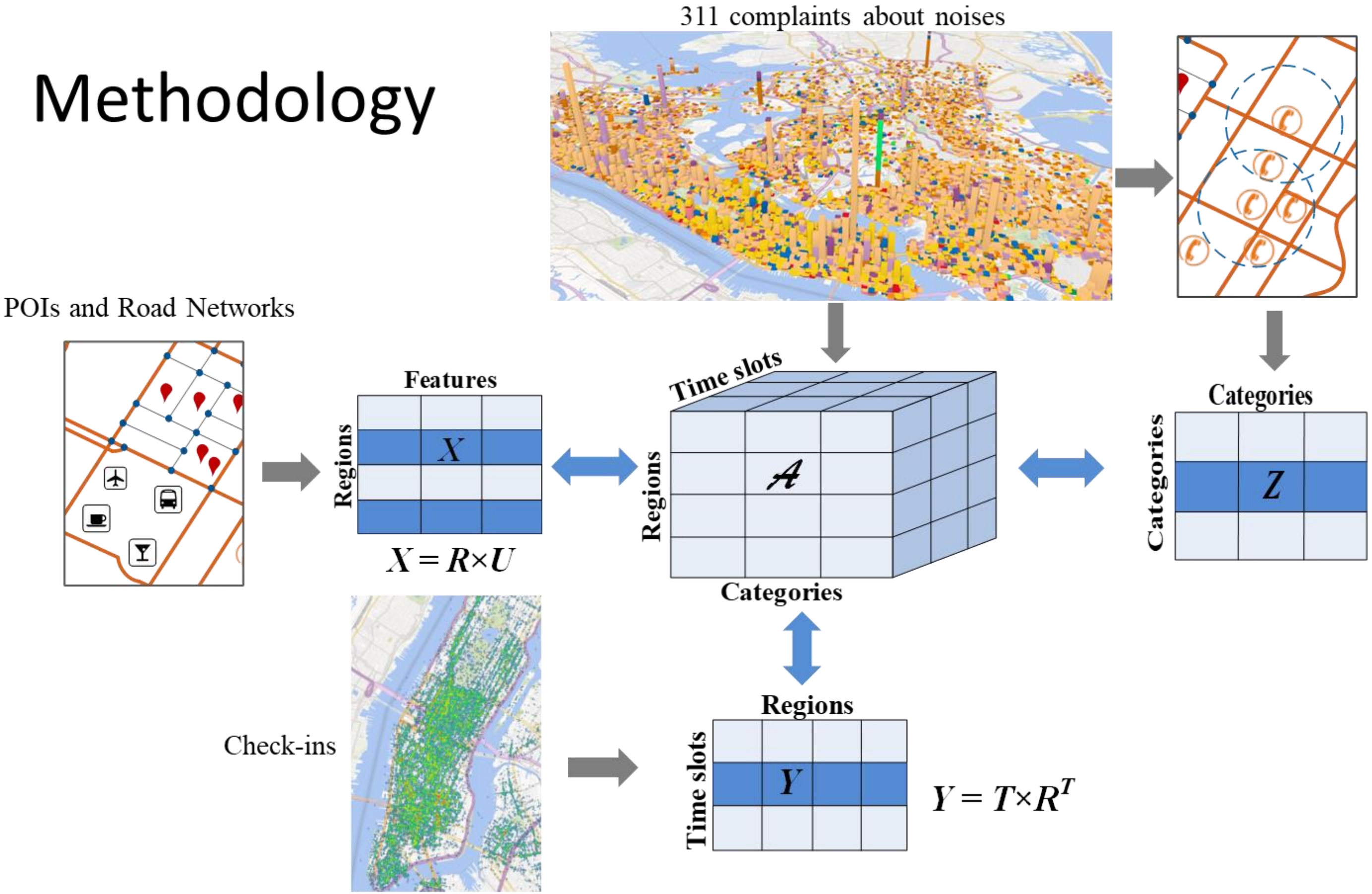
Goal

- Reveal the noise situation of each region in each hour
  - A noise indicator denoting the noisy level
  - Composition of noises in each location



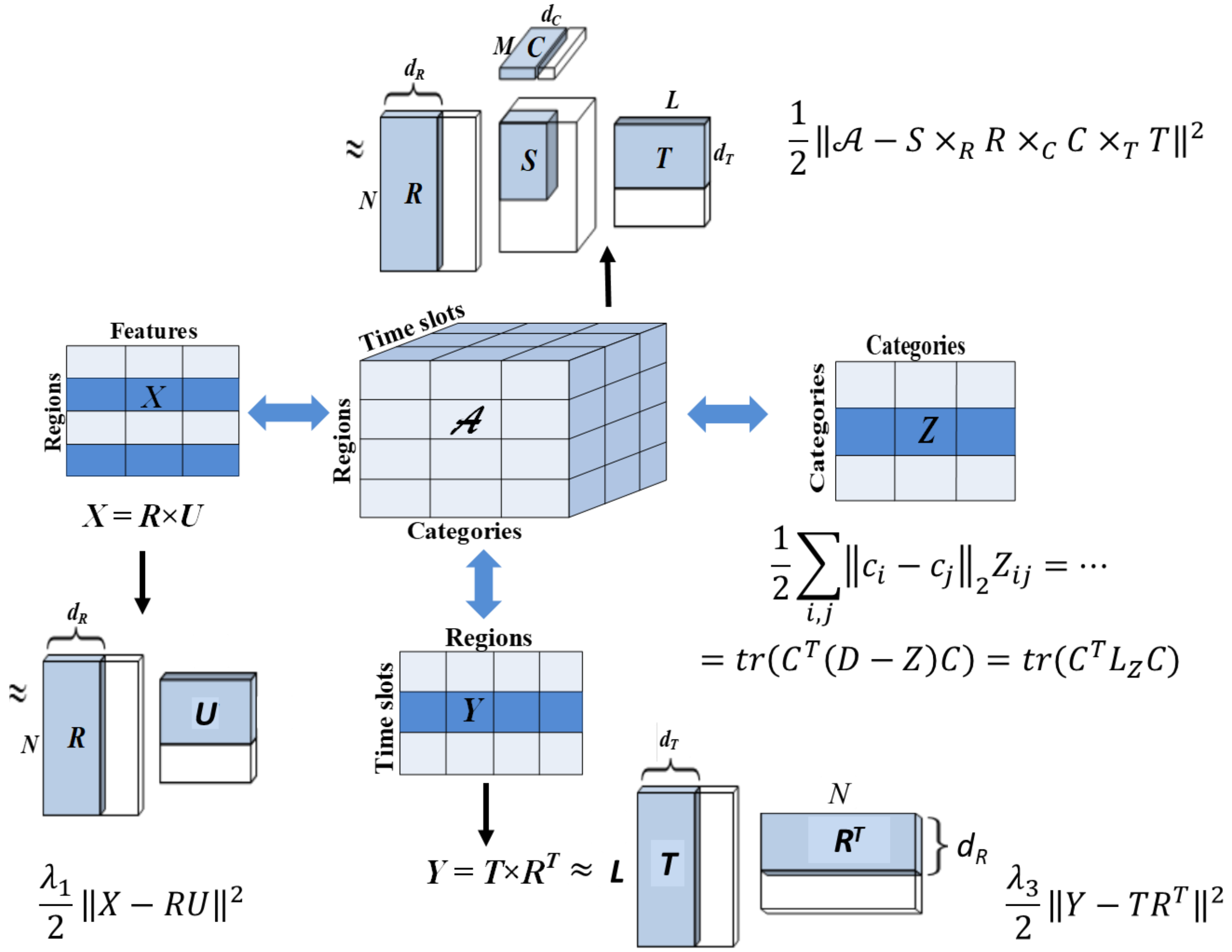

New York City's Noise
Data Panel
Weekday
All Categories
Noise Layer
Time Panel
0:00 ~ 0:00
Overall Time
Top 5 Noisiest Regions
#1 Audubon Ave: 135.01
#2 East Village & Stuyvesant Town: 105.07
#3 Columbia University: 98.16
#4 Wall Street: 97.63
#5 E Houston St & 1st Ave: 89.06
Region Noise Analysis
0 2 4 6 8 10 12 14 16 18 20 22
This demo is based on 311 data from May 23, 2013 to Jan. 31, 2014.
75
60
45
30
15
0
© 2014 Microsoft Corporation © 2013 Nokia

Methodology



$$\mathcal{L}(S,R,C,T,U) = \frac{1}{2}\|\mathcal{A} - S \times_R R \times_C C \times_T T\|^2 + \frac{\lambda_1}{2}\|X - RU\|^2 + \frac{\lambda_2}{2}\mathrm{tr}(C^T L_Z C) + \frac{\lambda_3}{2}\|Y - TR^T\|^2 + \frac{\lambda_4}{2}(\|S\|^2 + \|R\|^2 + \|C\|^2 + \|T\|^2 + \|U\|^2)$$

$$\mathcal{L}(S,R,C,T,U) = \frac{1}{2}\|\mathcal{A} - S \times_R R \times_C C \times_T T\|^2 + \frac{\lambda_1}{2}\|X - RU\|^2 + \frac{\lambda_2}{2}\mathrm{tr}(C^T L_Z C) + \frac{\lambda_3}{2}\|Y - TR^T\|^2 + \frac{\lambda_4}{2}(\|S\|^2 + \|R\|^2 + \|C\|^2 + \|T\|^2 + \|U\|^2)$$



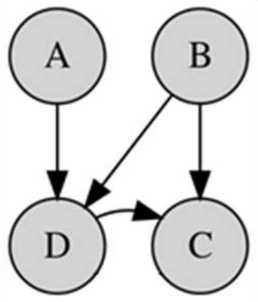
Semantic-based Data Fusion

- Multiple-view-based: co-training
- Similarity-based: Coupled matrix factorization
- Probabilistic dependency-based: graphical models
- Transfer learning-based

- The insight of each dataset and relations between features across different datasets
- Derived from the ways that people think of a problem
- Interpretable and meaningful

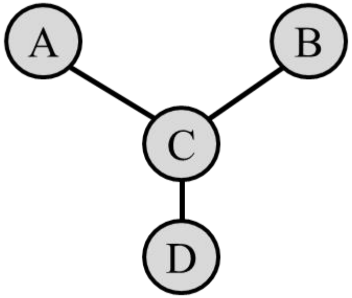
Semantic-based Data Fusion

- Bayesian Networks
  - A.K.A. Belief network
  - Directed acyclic graph
  - Factorize a joint distribution into conditional probabilities
  - Representatives: HMM, LDA, Neural Network



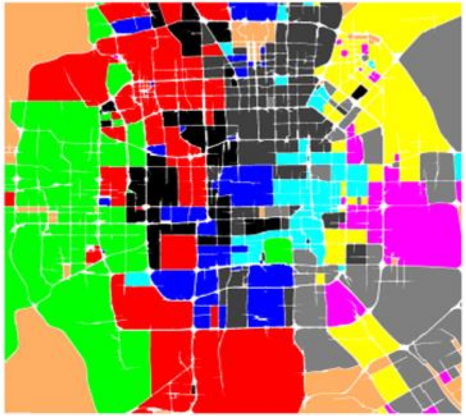
$$P[X_1, \ldots, X_n] = \prod_{i=1}^{n} P[X_i | pa_i]$$

- Markov Networks
  - A.K.A. Markov Random Field
  - Undirected graph and may be cyclic
  - Joint probability
  - Representatives: CRF, GMRF



Discover Regions of Different Functions using Human Mobility and POIs

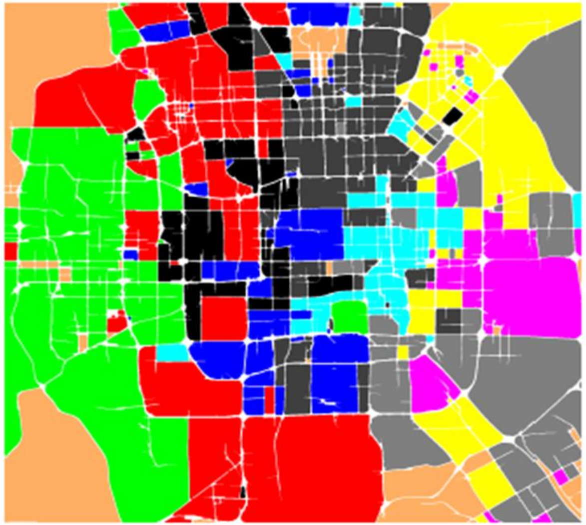
In KDD 2012



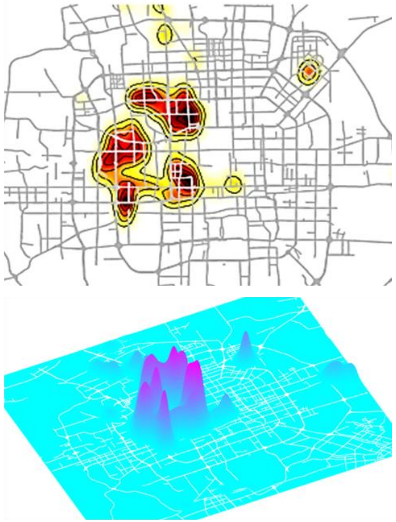
Jing Yuan, Yu Zheng. Discovering regions of different functions in a city using human mobility and POIs. In KDD 2012

Goals

- Discover regions of different functions in urban areas
- Identify the kernel density of a functionality



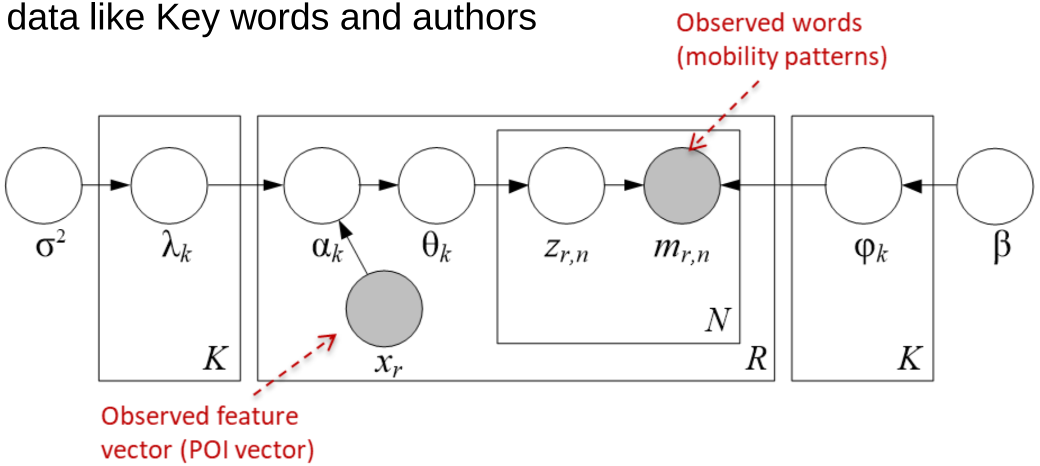
Functional Regions



Functionality Density

Methodology Overview

- **Mapping from regions to documents**
  - Regions → Documents (*R*)
  - Functions → Topics (*K*)
  - Mobility patterns → Words (*N*)
  - POIs → meta data like Key words and authors



Infer the topic distribution using a LDA(Latent Dirichlet allocation)-variant topic model

Semantic-based Data Fusion

- Multiple-view-based: co-training
- Similarity-based: Coupled matrix factorization
- Probabilistic dependency-based: graphical models
- Transfer learning-based

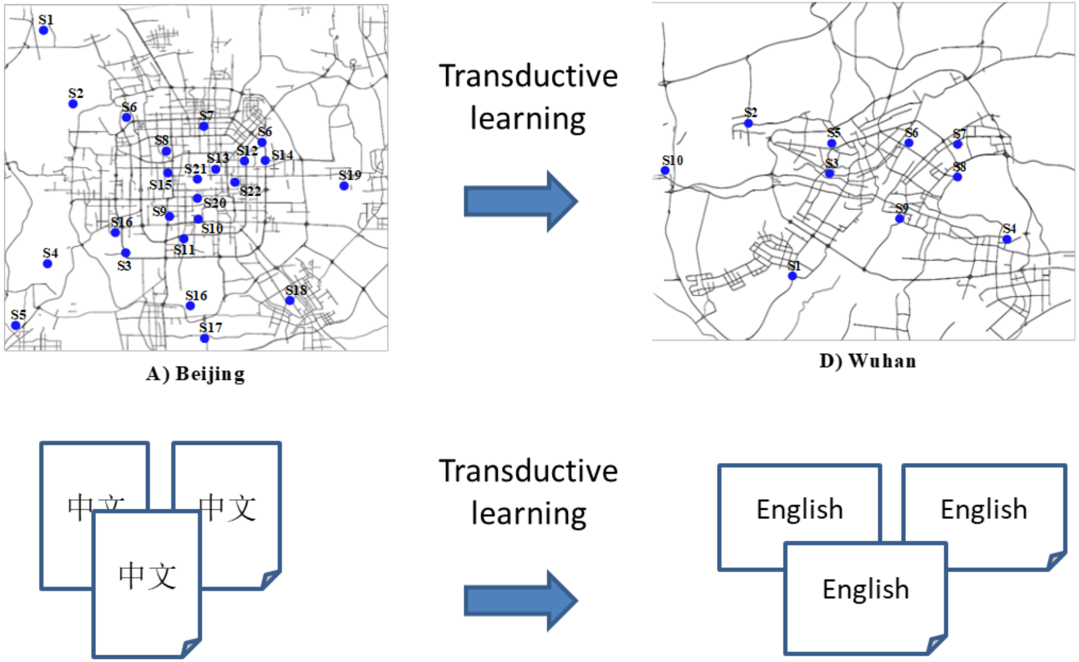
- The insight of each dataset and relations between features across different datasets
- Derived from the ways that people think of a problem
- Interpretable and meaningful

Transfer Learning

- Transfer learning
  - Between Single Type of Datasets
  - Between Multiple-types of datasets

Transfer Learning-based Data Fusion

- Transfer between a Single Type of Datasets

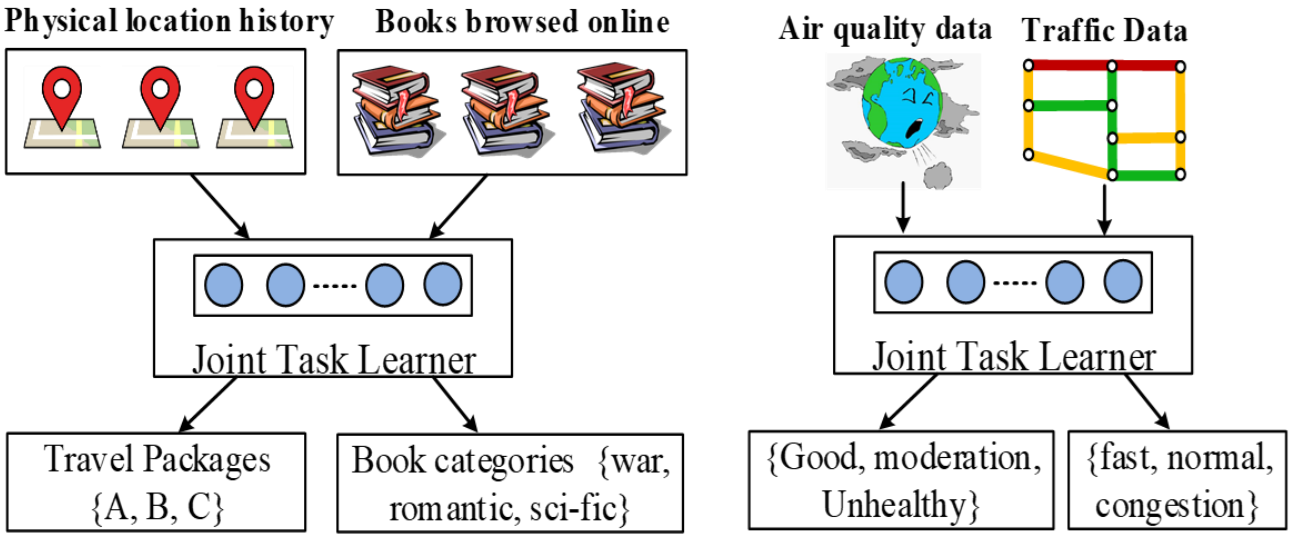


Heterogeneous transfer learning

Transfer Learning-based Data Fusion

- Transfer between a Multiple Type of Datasets

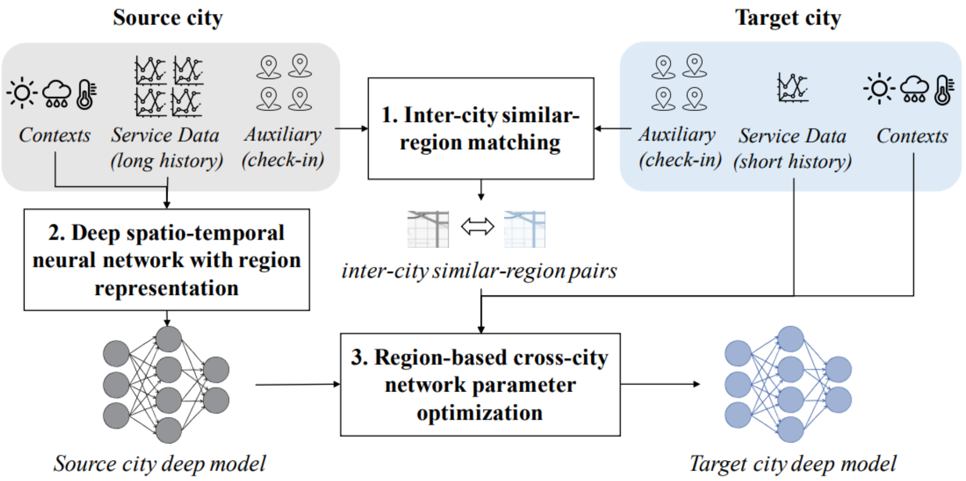
Inductive Learning: Multi-Task Learning

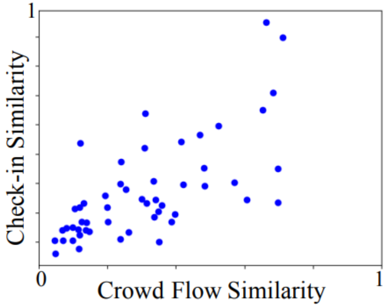


A) Book-travel interests co-estimation B) Air quality-traffic co-prediction

Transfer Learning for Deep Spatio-Temporal Prediction

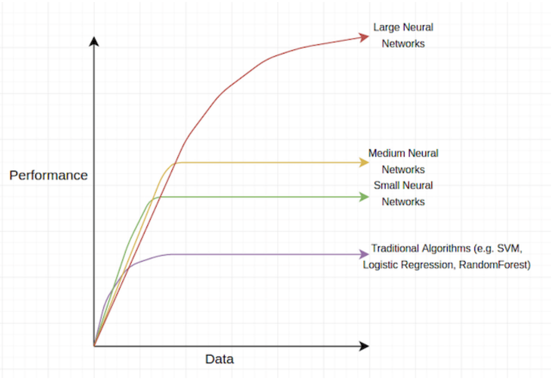
IJCAI 2019





Data Scarcity in Spatio-Temporal Domains

- Machine learning relies on big data
  - More data leads to better performance
- Data scarcity is common in spatio-temporal domains, e.g., in cities
  - Building new cities
  - Planning new urban services
- How to mitigate lack of data for spatio-temporal data mining?



Solution: Transfer Learning

- Transfer learning:
  - Key idea: Borrow knowledge from different but related tasks
  - Effective in visual recognition, text mining, etc.
    - ➢ Pre-training & fine-tuning in computer vision
    - ➢ Cross-domain sentiment classification

Problem Statement

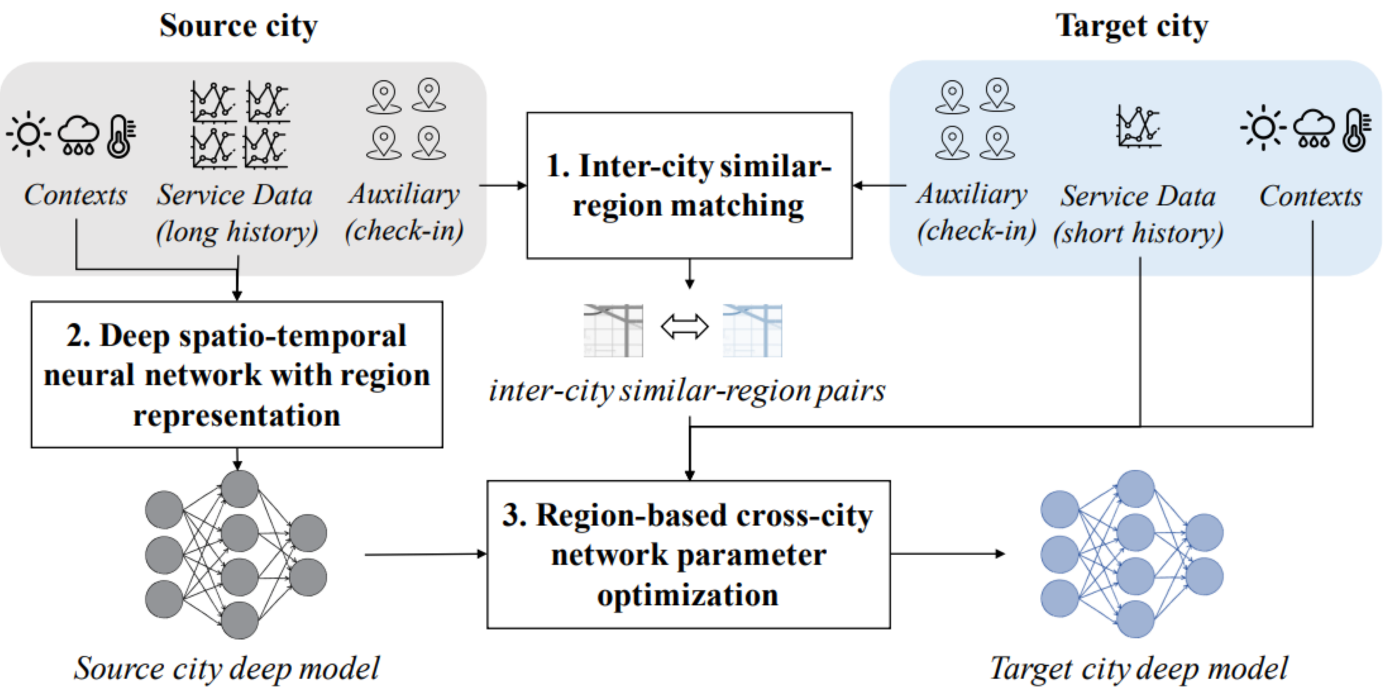
Given the little service data in target city $\mathcal{D}$ and rich service data in source city $\mathcal{D}'$, we aim to learn a function f to predict the citywide service data in the target city $\mathcal{D}$ at the next time-stamp $t_c + 1$:

$$\min_{f} \quad error(\widetilde{\mathcal{S}}_{t_c+1,\mathcal{D}}, \mathcal{S}_{t_c+1,\mathcal{D}}) \tag{5}$$

$$\text{where} \quad \widetilde{\mathcal{S}}_{t_c+1,\mathcal{D}} = f(\mathbb{S}_{\mathcal{D}}, \mathbb{S}_{\mathcal{D}'}), \quad |\mathbb{T}_{\mathcal{D}}| \ll |\mathbb{T}_{\mathcal{D}'}| \tag{6}$$

error metric may be mean absolute error, root mean squared error, etc., according to the real application requirement.

Framework



Source city
Target city
Contexts
Service Data (long history)
Auxiliary (check-in)
Auxiliary (check-in)
Service Data (short history)
Contexts
1. Inter-city similar-region matching
inter-city similar-region pairs
2. Deep spatio-temporal neural network with region representation
3. Region-based cross-city network parameter optimization
Source city deep model
Target city deep model

Inter-City Similar Region Matching

- Matching with a Short Period
  - For each target region, they choose the source region with the largest correlation value

$$
\begin{aligned}
&\mathcal{M}(r) = r^*, && r \in \mathbb{C}_{\mathcal{D}}, r^* \in \mathbb{C}_{\mathcal{D}'} \\
&\rho_{r,r^*} \geq \rho_{r,r'}, && \forall r' \in \mathbb{C}_{\mathcal{D}'} \\
&\rho_{r,r^*} = corr(\{s_{r,t}\}, \{s_{r^*,t}\}), && r \in \mathbb{C}_{\mathcal{D}}, r^* \in \mathbb{C}_{\mathcal{D}'}, t \in \mathbb{T}_{\mathcal{D}}
\end{aligned}
$$

Inter-City Similar Region Matching

- Matching with a Long Period of Auxiliary Data
  - Due to data scarsity in the target city, the above correlation similarity between a source region and a target region may not be very reliable
  - In reality, sometimes we can find another openly-accessible auxiliary data that correlates with the service data, which may help calculate the inter-city region similarity more robustly, such as check-in data

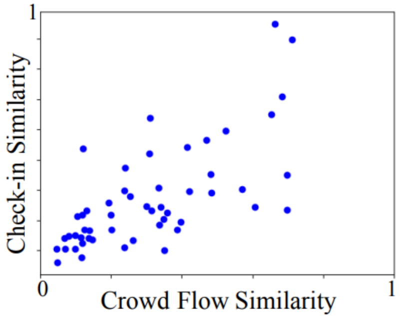
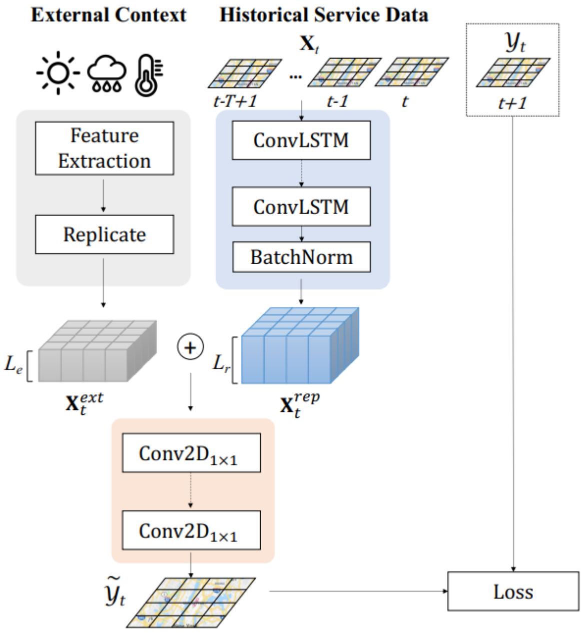


Figure 2: Check-in/crowd flow similarities.

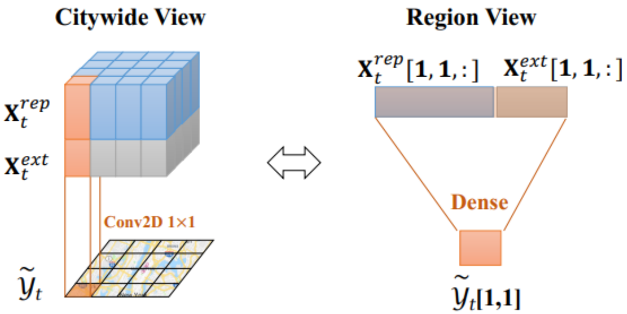
$$\rho_{r,r^*} = corr(\{a_{r,t}\}, \{a_{r^*,t}\}), \quad r \in \mathbb{C}_{\mathcal{D}}, r^* \in \mathbb{C}_{\mathcal{D}'}, t \in \mathbb{T}_{\mathcal{A}}$$

ST Neural Networks with Region Representations

- Existing deep spatio-temporal models often take the whole city data for end-to-end prediction, e.g., ST-ResNet, which cannot be used for region-level transfer



(a)



(b)

$$\textit{ConvLSTM: } f_{\theta_1}: \mathbb{R}^{k\times W\times H} \to \mathbb{R}^{W\times H\times L_r} \tag{8}$$

$$\textit{Region representation: } \mathbf{X}_t^{rep} = f_{\theta_1}(\mathbf{X}_t) \tag{9}$$

$$\textit{Merge: } f_m: (\mathbb{R}^{W\times H\times L_r}, \mathbb{R}^{W\times H\times L_e}) \to \mathbb{R}^{W\times H\times (L_r+L_e)} \tag{10}$$

$$\textit{Conv2D}_{1\times1}\textit{: } f_{\theta_2}: \mathbb{R}^{W\times H\times (L_r+L_e)} \to \mathbb{R}^{W\times H} \tag{11}$$

$$\textit{Prediction}: \widetilde{\mathcal{Y}}_t = f_{\theta_2}(f_m(\mathbf{X}_t^{rep}, \mathbf{X}_t^{ext})) \tag{12}$$

$$= f_{\theta_2}(f_m(f_{\theta_1}(\mathbf{X}_t), \mathbf{X}_t^{ext})) \tag{13}$$

Region-based Cross-city Network Parameter Optimization

- Target 1: more accurate prediction in the target city
- Target 2: we try to minimize the squared error between the network hidden representations of the target region and its matched source region

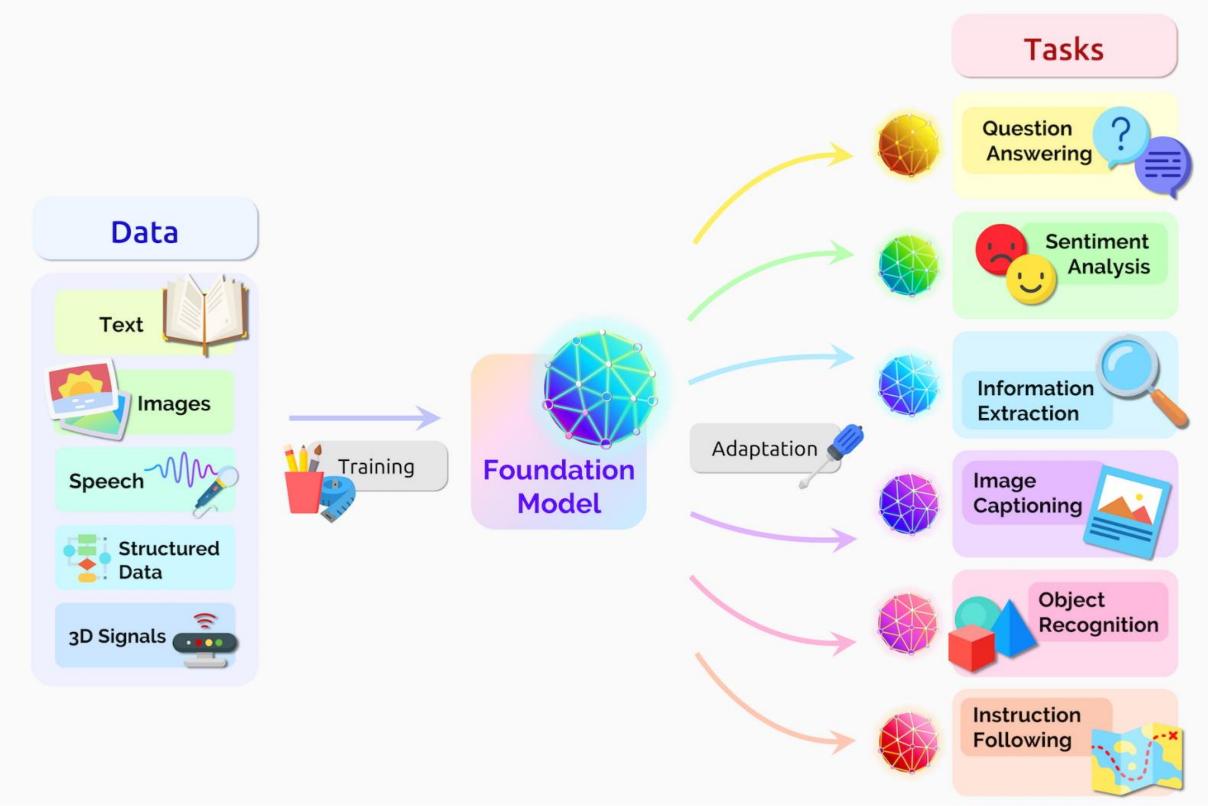
$$\min_{\theta_{\mathcal{D}}} \quad (1-w) \sum_{t \in \mathbb{T}_{\mathcal{D}}} ||\widetilde{\mathcal{Y}}_t - \mathcal{Y}_t||_F^2 + w \sum_{r \in \mathbb{C}_{\mathcal{D}}} \sum_{t \in \mathbb{T}_{\mathcal{D}}} \rho_{r,r^*} \cdot ||\mathbf{x}_{r,t}^{rep} - \mathbf{x}_{r^*,t}^{rep}||^2$$

Evaluation

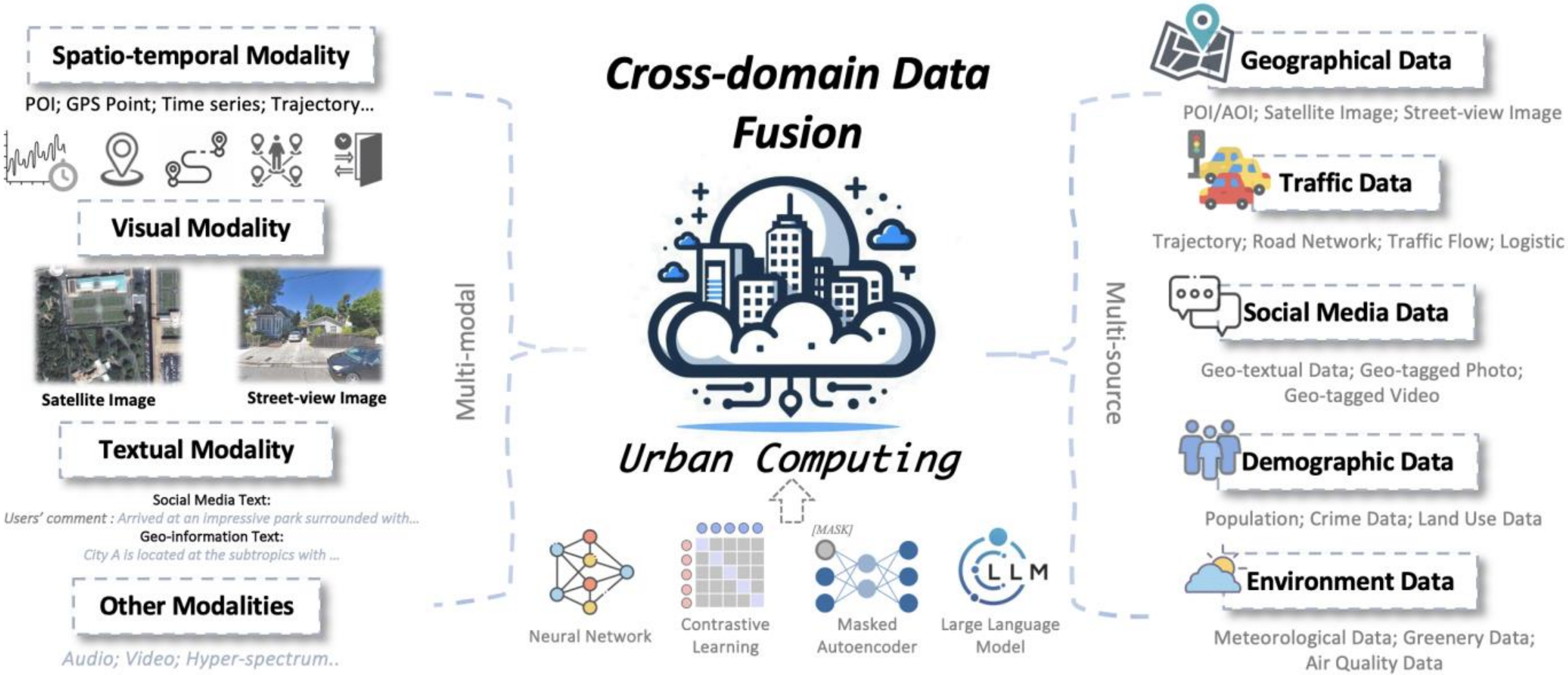
	D.C.→Chicago		Chicago→D.C.		D.C.→NYC		NYC→D.C.	
	1-day	3-day	1-day	3-day	1-day	3-day	1-day	3-day
Target Only								
ARIMA	0.740	0.694	0.707	0.661	0.360	0.341	0.707	0.661
DeepST	0.771	0.711	1.075	0.767	0.350	0.359	1.075	0.767
ST-ResNet	0.914	0.703	0.869	0.738	0.376	0.349	0.869	0.738
Source & Target								
DeepST (FT)	0.652	0.611	0.672	0.619	0.363	0.369	0.713	0.711
ST-ResNet (FT)	0.667	0.615	0.695	0.623	0.385	0.349	0.696	0.691
RegionTrans (S-Match)	0.605	0.594	0.631	0.602	**0.328**	**0.305**	**0.665**	**0.593**
RegionTrans (A-Match)	**0.587**	**0.576**	**0.600**	**0.581**	/	/	/	/

Table 1: Evaluation results. The target city holds 1 or 3-day crowd flow historical data. *RegionTrans (A-Match)* is available for D.C. $\rightleftharpoons$ Chicago as we have collected check-in data for Chicago and D.C.

Foundation Models

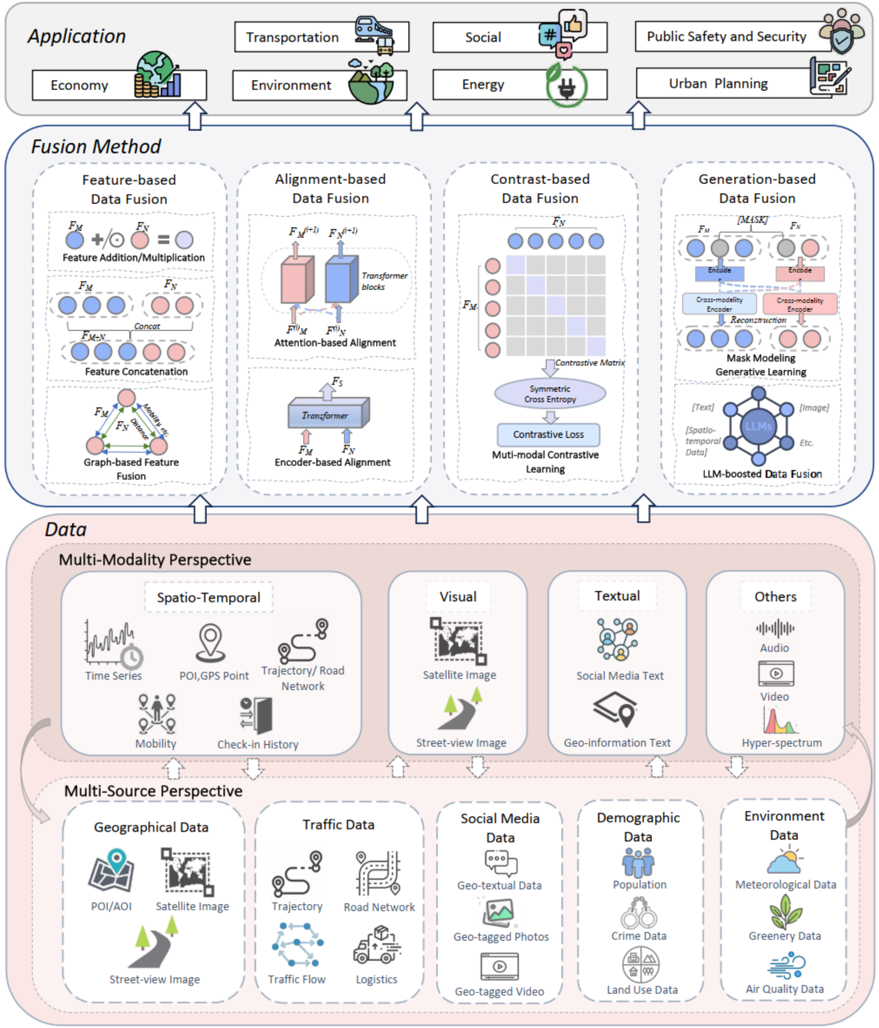


Deep Learning-based Fusion



X. Zou, Y. Yan, **Y. Liang*** et al., Deep Learning for Cross-Domain Data Fusion in Urban Computing:Taxonomy, Advances, and Outlook, ongoing

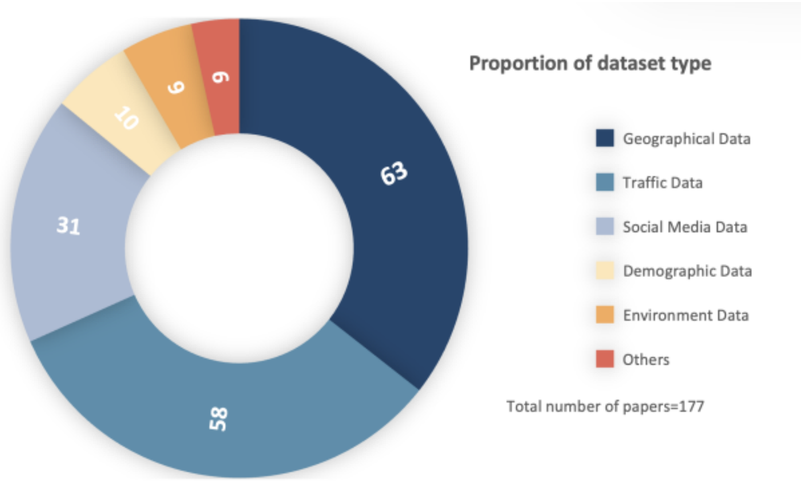
Taxonomy



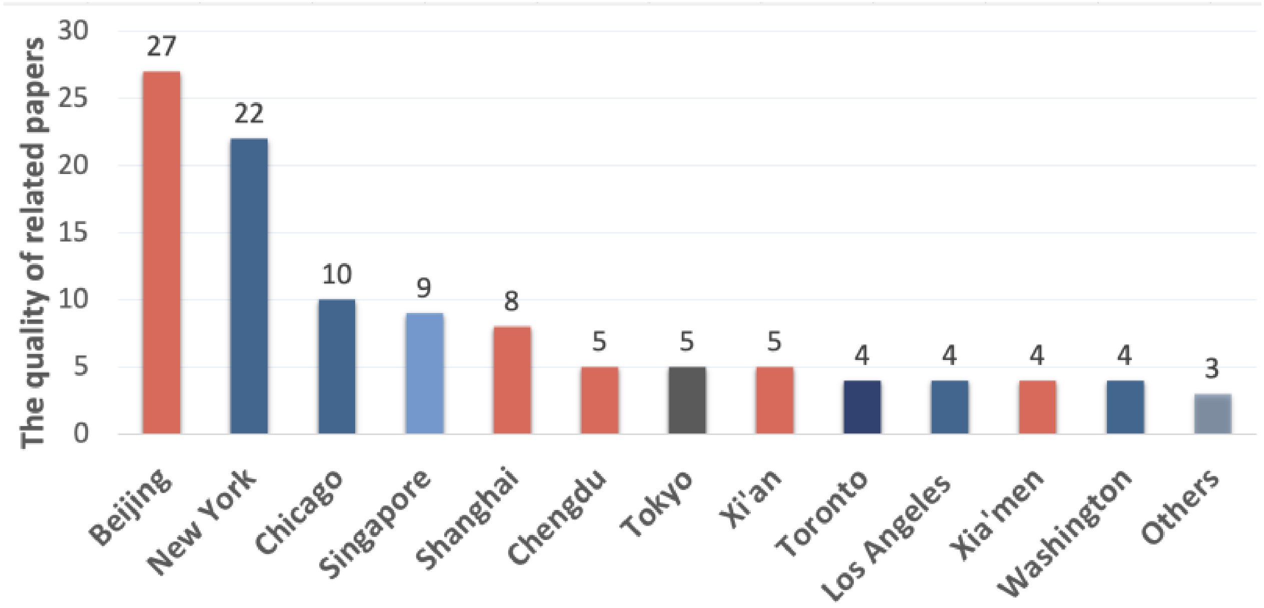
Application
Transportation
Social
Public Safety and Security
Economy
Environment
Energy
Urban Planning
Fusion Method
Feature-based Data Fusion
Feature Addition/Multiplication
Concat
Feature Concatenation
Graph-based Feature Fusion
Alignment-based Data Fusion
Transformer blocks
Attention-based Alignment
Transformer
Encoder-based Alignment
Contrast-based Data Fusion
Contrastive Matrix
Symmetric Cross Entropy
Contrastive Loss
Muti-modal Contrastive Learning
Generation-based Data Fusion
[MASK]
Encode
Cross-modality Encoder
Reconstruction
Mask Modeling Generative Learning
[Text]
[Image]
[Spatio-temporal Data]
LLMs
Etc.
LLM-boosted Data Fusion
Data
Multi-Modality Perspective
Spatio-Temporal
Time Series
POI,GPS Point
Trajectory/ Road Network
Mobility
Check-in History
Visual
Satellite Image
Street-view Image
Textual
Social Media Text
Geo-information Text
Others
Audio
Video
Hyper-spectrum
Multi-Source Perspective
Geographical Data
POI/AOI
Satellite Image
Street-view Image
Traffic Data
Trajectory
Road Network
Traffic Flow
Logistics
Social Media Data
Geo-textual Data
Geo-tagged Photos
Geo-tagged Video
Demographic Data
Population
Crime Data
Land Use Data
Environment Data
Meteorological Data
Greenery Data
Air Quality Data

Data Perspective

- Geographical data
- Traffic data
- Social media data
- Demographic data
- Environment data
- Others

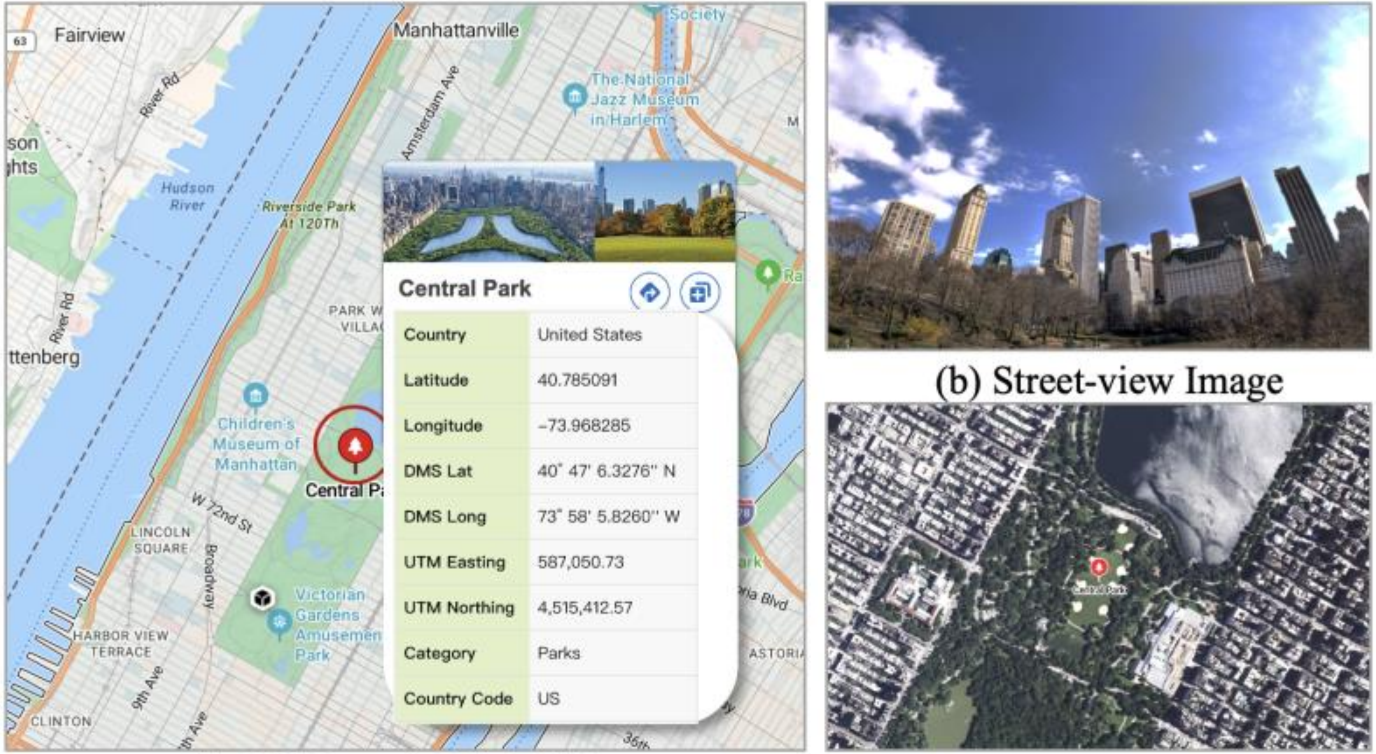


Data Sources



Category	Content	Format	Dataset	Link	Reference
Geographical Data	Satellite Image	Image	ArcGIS	https://developers.arcgis.com	[175]
			PlanetScope	https://developers.planet.com/docs/data/planetscope/	[147]
			Google Earth	https://developers.google.com/maps/documentation/	[112]
			OpenStreetMap	https://www.openstreetmap.org/	[313]
			Baidu Maps	https://lbsyun.baidu.com	[300, 289]
	Street-View Image	Image	Baidu Map	https://lbsyun.baidu.com	[175, 120]
			Google Street	https://developers.google.com/maps/	[175, 4]
			Tencent Map	https://lbs.qq.com/tool/streetview/index.html	[108]
	POIs	Point Vector	Tencent Map Service	https://lbs.qq.com/getPoint/	[286, 220]
			WeChat POIs	https://open.weixin.qq.com	[257]
			Baidu Map POIs	https://lbsyun.baidu.com	[147, 161, 164, 106, 289]
			NYC Open POIs	https://opendata.cityofnewyork.us/	[159, 253, 19, 341, 268]
			Foursquare	https://developer.foursquare.com/docs/checkins/checkins	[19, 355, 13, 41, 103, 112]
			Wikipedia POIs	https://www.wikipedia.org	[360]
			AMap Service	https://lbs.amap.com	[10]
			Yelp POIs	https://www.yelp.com/developers	[13, 354, 357]
			Dianping POIs	https://api.dianping.com/	[32, 62]
			Weibo POIs	https://open.weibo.com/wiki/API	[32, 129, 75]
			Flickr POIs	https://www.flickr.com/services/developer/api/	[95]
			Bing Map POIs	https://www.bingmapsportal.com	[36]
Traffic Data	Traffic Trajectory	Spatio-temporal Trajectory	Shenzhou UCar	https://bit.ly/2MG47xz	[90]
			Chicago Transportation	https://data.cityofchicago.org/	[253, 268, 112]
			VED	https://github.com/gsoh/VED	[197, 346]
			Taxi Shenzhen	https://github.com/cbdog94/STL	[109, 279]
			NYC Open Taxi Data	https://opendata.cityofnewyork.us/how-to/	[343, 341]
			GeoLife	http://urban-computing.com/index-893.htm	[93, 371, 373, 367, 323]
			T-Drive Taxi	http://urban-computing.com/index-58.htm	[326, 327, 204, 180]
			DiDi Traffic	https://outreach.didichuxing.com/research/opendata/	[325, 177, 213, 304, 244]
			Xiamen Taxi	https://data.mendeley.com/datasets/6xg39x9vgd/1	[318, 39, 120, 38]
			Grab-Posisi	https://goo.su/W3yD5m	[313, 315]
	Taffic Flow	Spatio-temporal Graph	California-PEMS	http://pems.dot.ca.gov	[9, 238]
			METR-LA	https://www.metro.net	[138, 160]
			Large-ST	https://github.com/liuxu77/LargeST	[171]
			MobileBJ	https://github.com/FIBLAB/DeepSTN/issues/4	[159, 129, 32]
			TaxiBJ	https://goo.su/aQyjTAz	[156, 11, 211, 116, 343, 72]
			BikeNYC	https://citibikenyc.com/	[159, 11, 211, 116]
	Road Network	Spatial Graph	OpenStreetMap	https://www.openstreetmap.org	[315, 13, 177, 325, 81]
			US Census Bureau	https://www.census.gov/data.html	[341]
	Logistics	Spatio-temporal Trajectory	LaDe	https://cainiaotechai.github.io/LaDe-website/	[282]
			JD Logistics	https://corporate.jd.com/ourBusiness#jdLogistics	[220]

Geographical Data

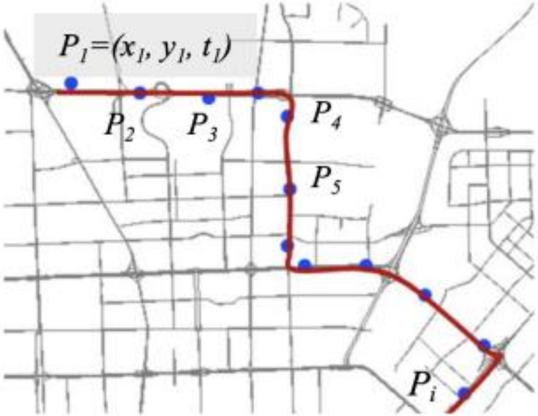


(a) POI Data & Digital Map

(b) Street-view Image

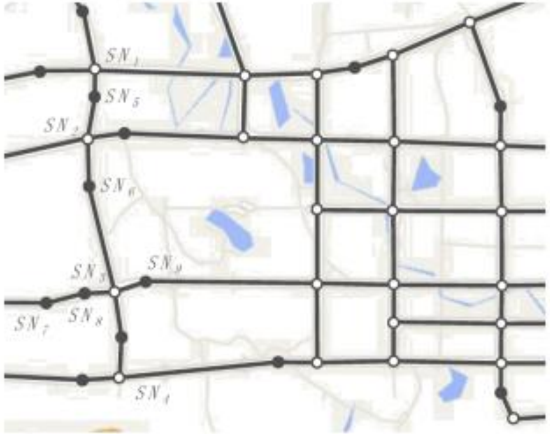
(c) Satellite Image

Traffic Data



(a) Traffic trajectory

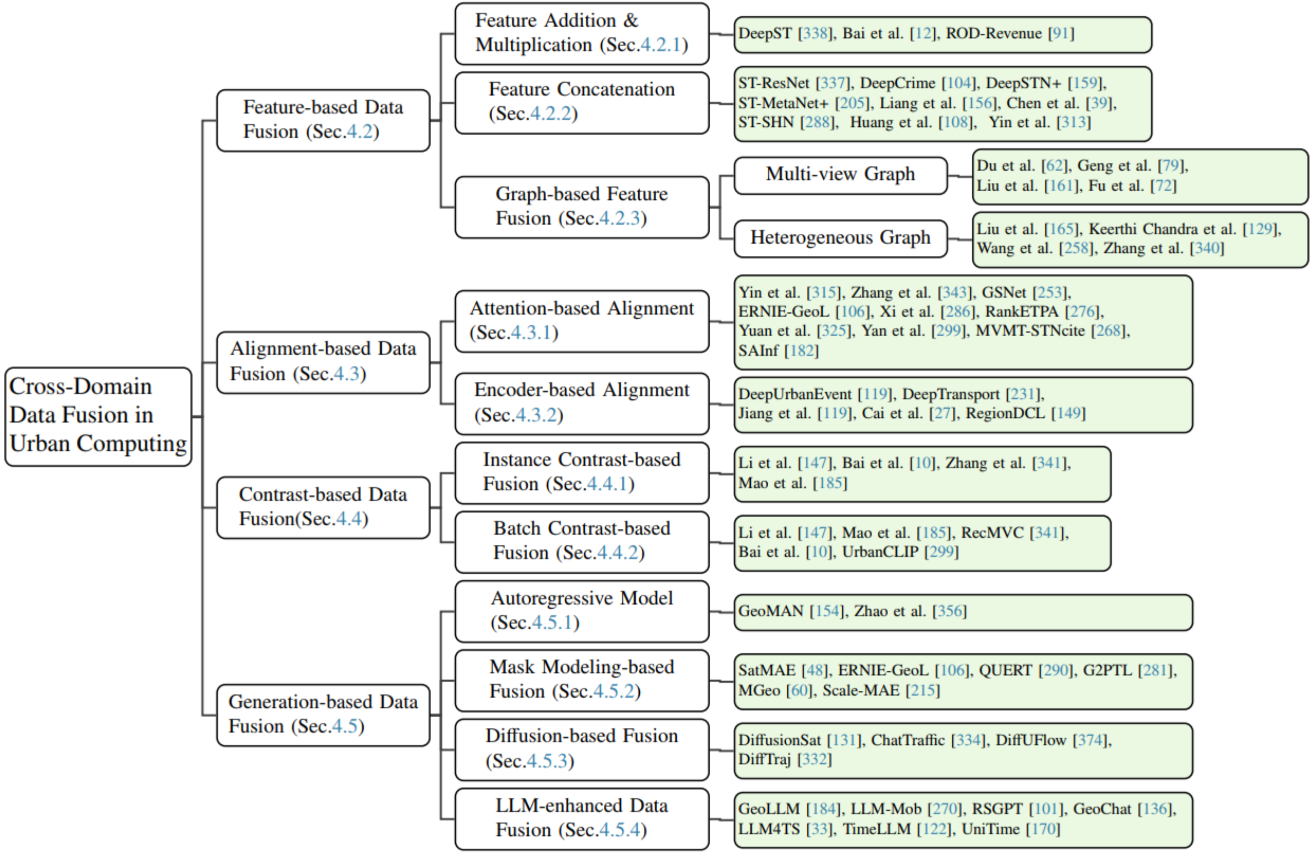
(b) Traffic flow



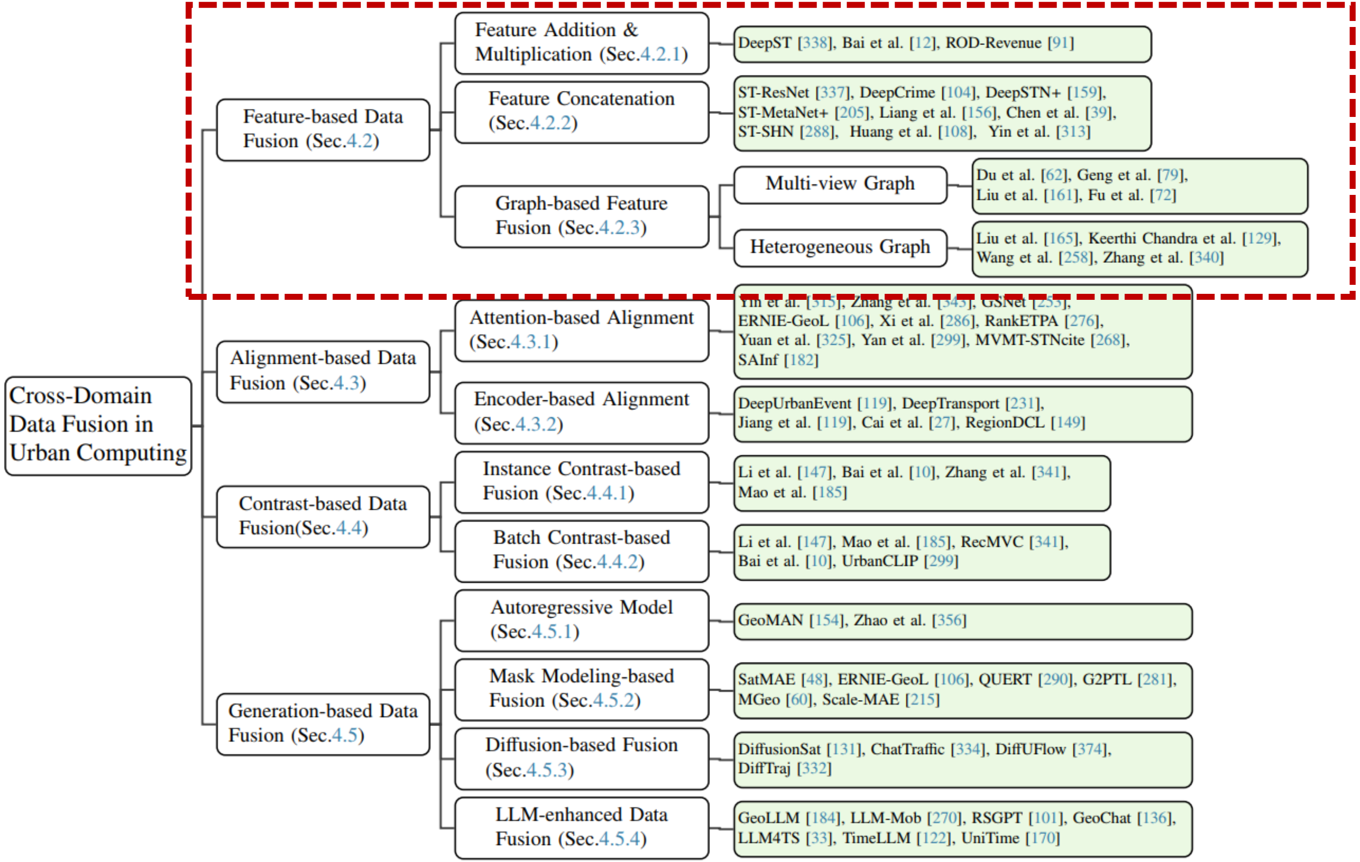
(c) Road network

Social Media Data	Text	Text	Twitter	https://developer.twitter.com/en/docs	[19, 355, 357, 328, 251, 278, 224] [269, 263, 265, 264, 189, 173]
			Common Crawl	https://registry.opendata.aws/commoncrawl/	[354]
			Yelp Reviews	https://www.yelp.com/dataset	[354, 357]
			Weibo Traffic Police	http://open.weibo.com/developers/	[318]
	Geo-tagged Image & Video	Image&Video	YFCC100M	https://goo.su/jzaDU	[360, 316, 95]
			NUS-WIDE	https://goo.su/dWPQZcD	[316, 314]
			GeoUGV	https://qualinet.github.io/databases/video/	[176]
	Users' Info	Time Series	Jiepang User Check-in	https://jiepang.app/	[72]
			Gowalla User Location	http://konect.cc/networks/loc-gowalla_edges/	[41, 328]
			WeChat Mobility	https://open.weixin.qq.com/	[257]
Demographic Data	Crime	Time Series	NYC Crime	https://opendata.cityofnewyork.us/	[343]
	Land Use	Time Series	Land Use SG	https://www.ura.gov.sg/Corporate/Planning/Master-Plan	[149]
			Land Use NYC	https://goo.su/puTuG	[149]
	Population	Time Series	WorldPop	https://www.worldpop.org/	[286, 147, 10]
Environment Data	Meteorology	Time Series	TipDM China Weather	https://www.tipdm.org/	[167]
			DarkSky Weather	https://support.apple.com/en-us/102594	[325]
			WeatherNY	https://opendata.cityofnewyork.us/	[253]
			WeatherChicago	https://data.cityofchicago.org/	[253]
			Weather Underground	https://www.wunderground.com/	[318]
			DidiSY	https://www.didiglobal.com/	[12]
			WD_BJ weather	https://goo.su/DmHFHd	[181]
			WD_USA weather	https://goo.su/RVhBA	[181]
	Greenery	Time Series	Google Earth	https://earth.google.com/	[318]
	Air Quality	Time Series	UrbanAir	https://goo.su/hfzNB53	[372, 369, 366]
			KnowAir	https://github.com/shuowang-ai/PM2.5-GNN	[266, 322, 345, 294]

DL-based Data Fusion

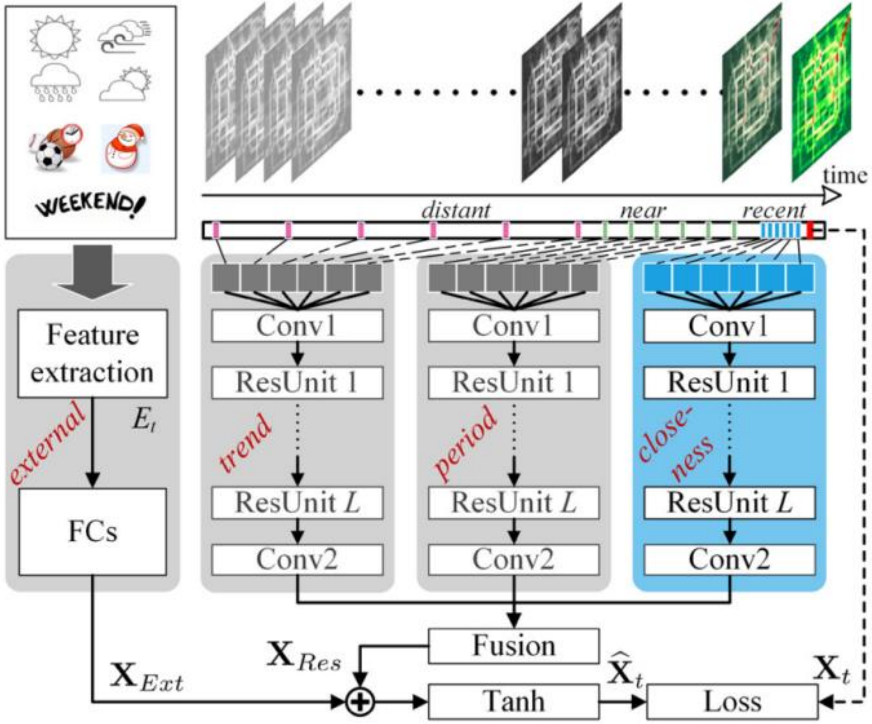


DL-based Data Fusion



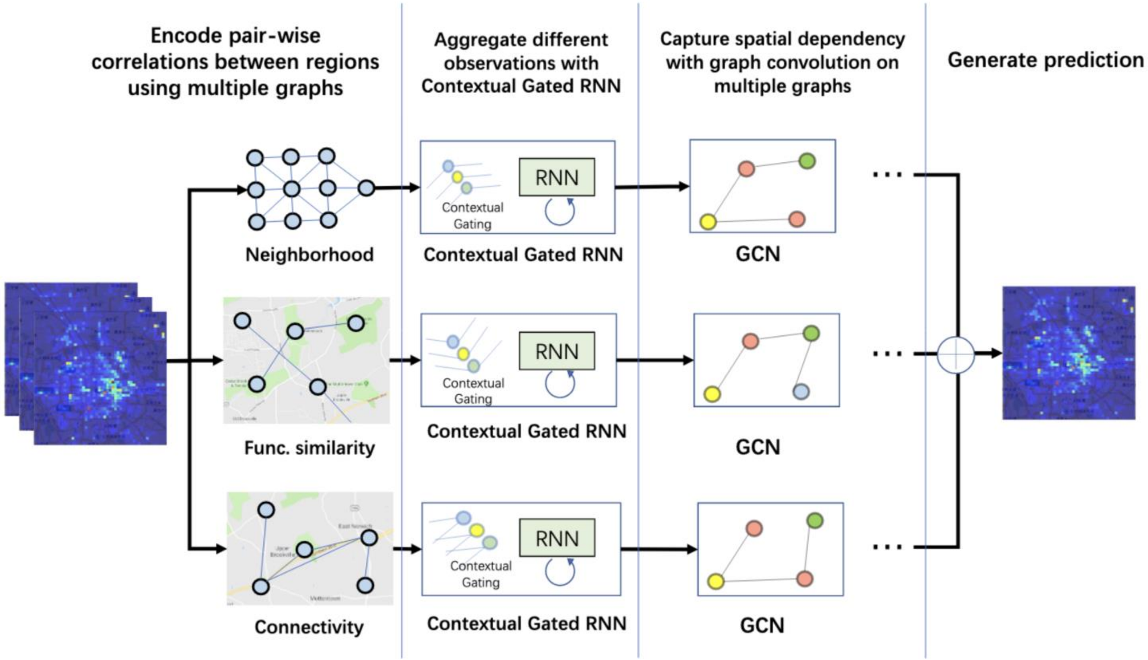
Feature-based Fusion

- Feature Addition and Multiplication
- Feature Concatenation



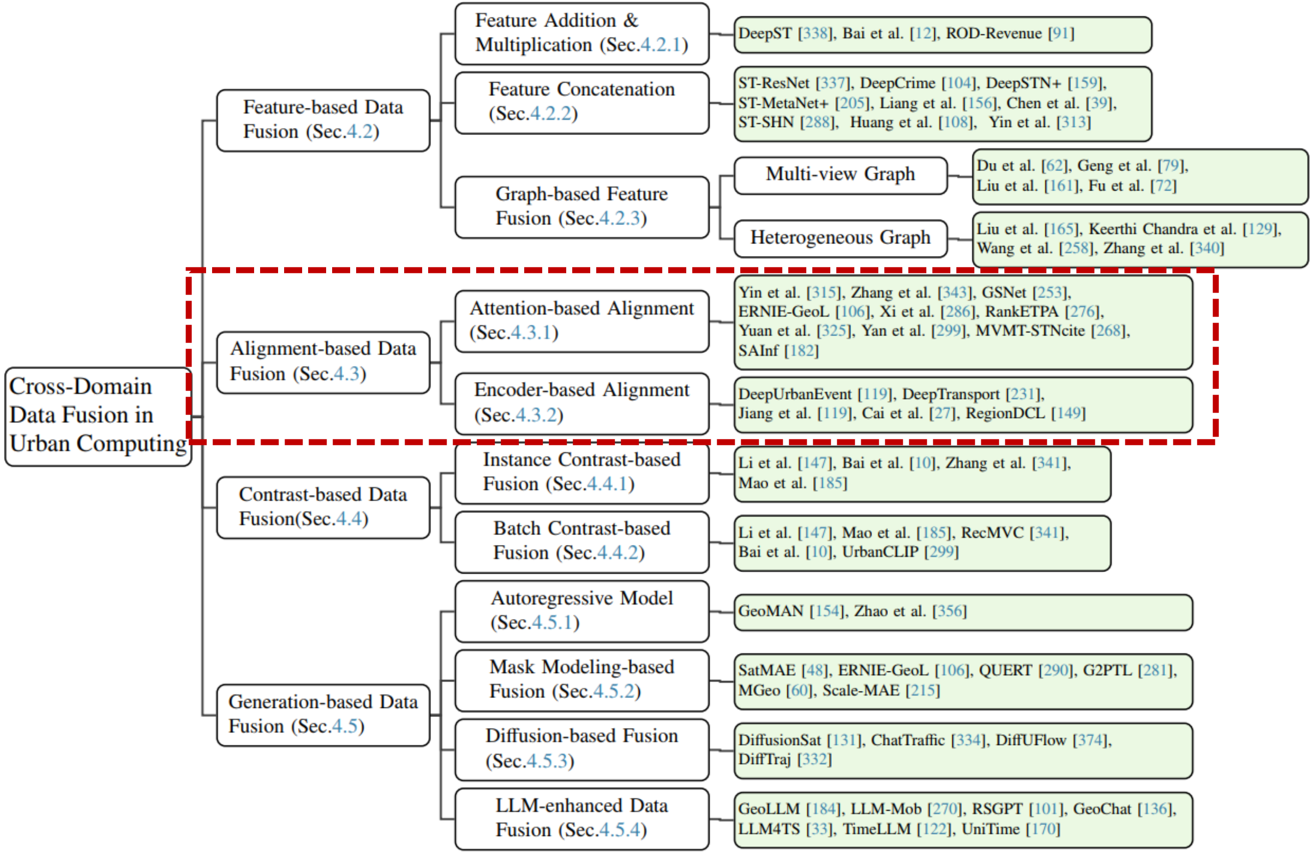
Feature-based Fusion

- Graph-based Data Fusion



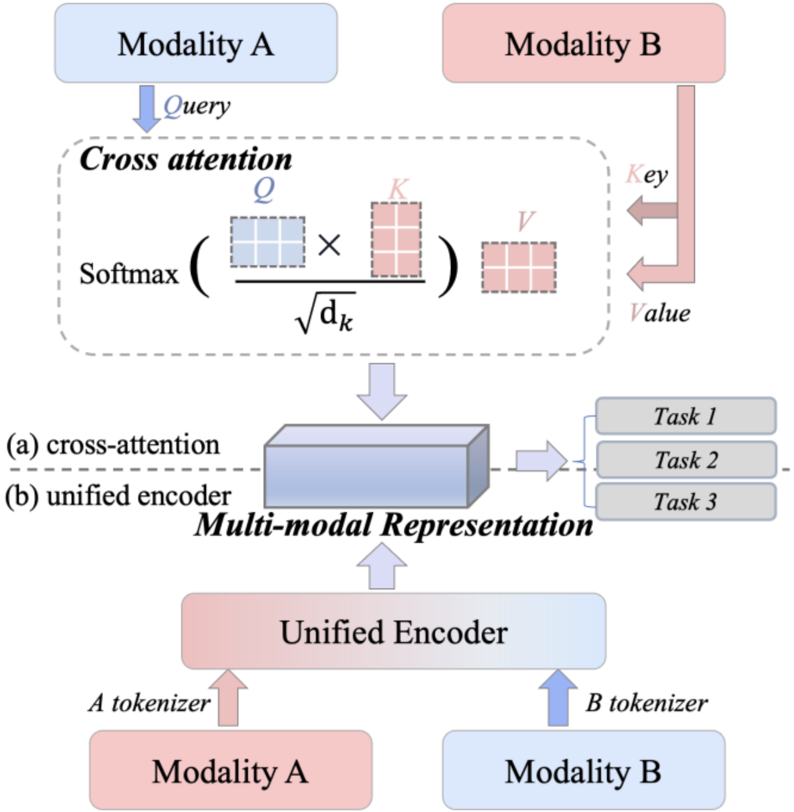
X. Geng et al., [Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting](), AAAI 2019.

DL-based Data Fusion

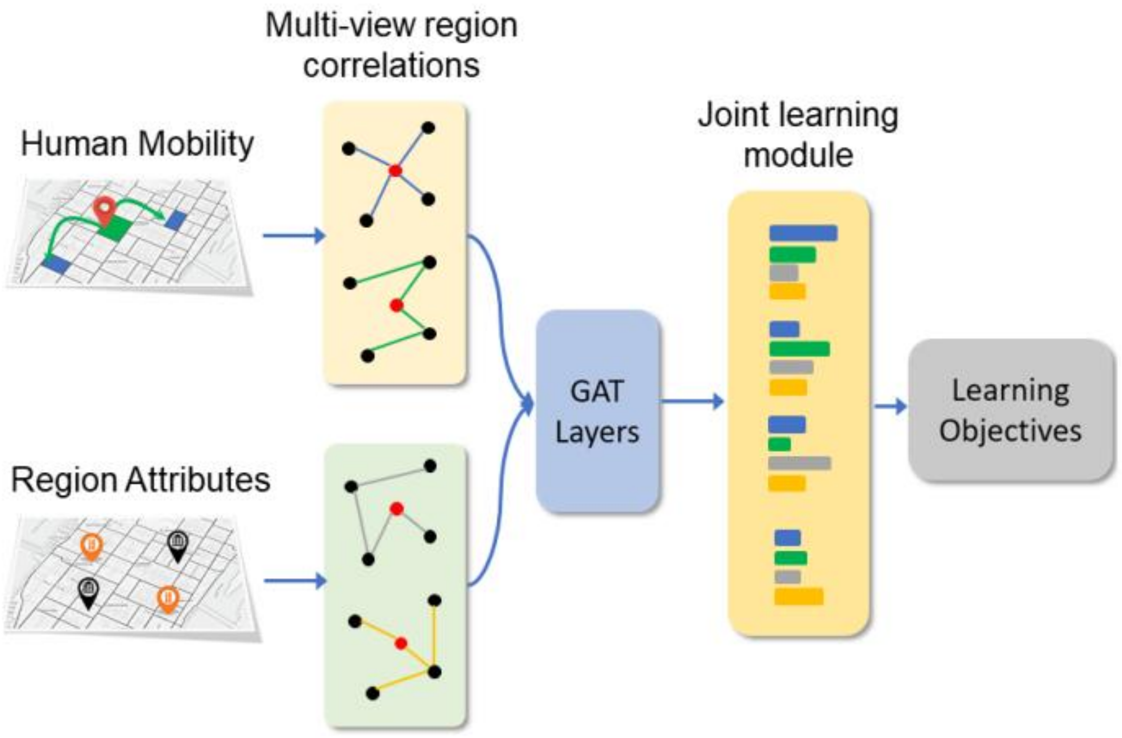


Alignment-based Alignment

- Attention-based fusion
- Encoder-based fusion

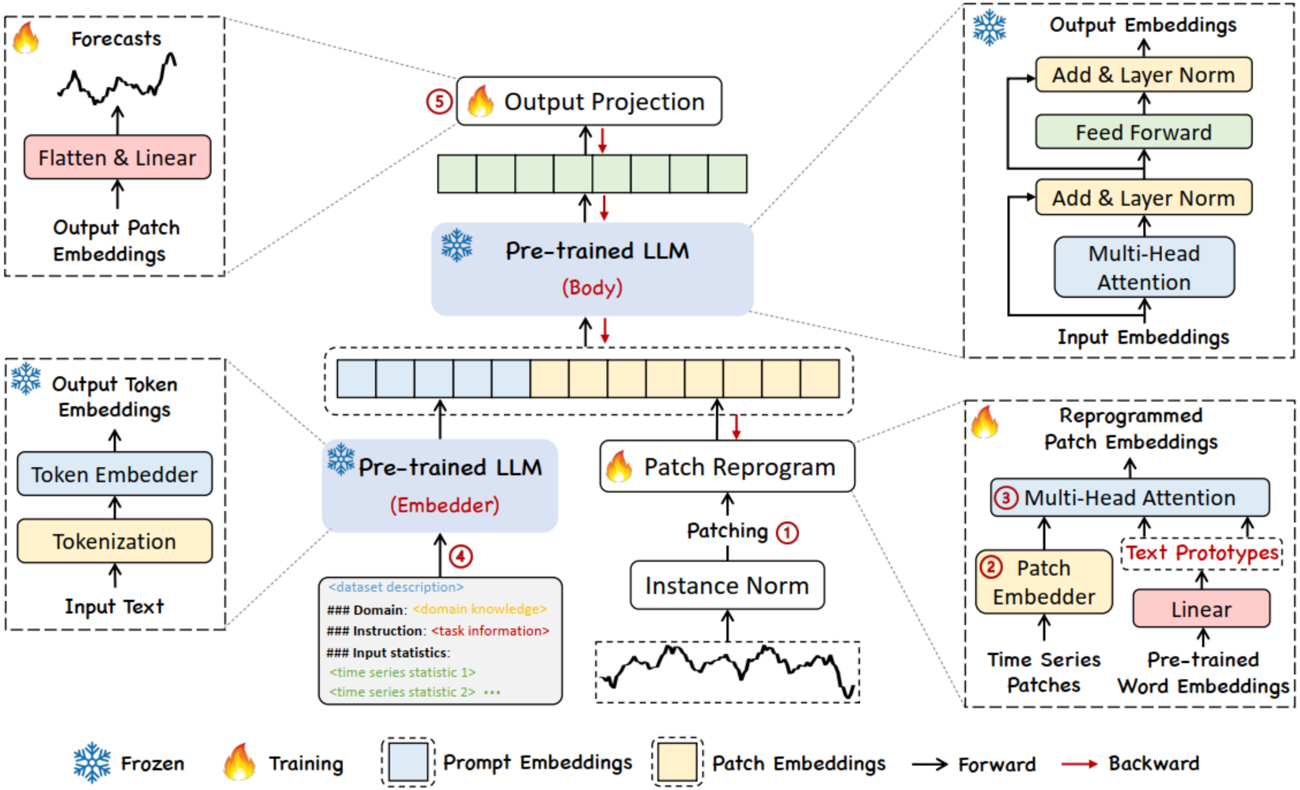


Attention-based Fusion



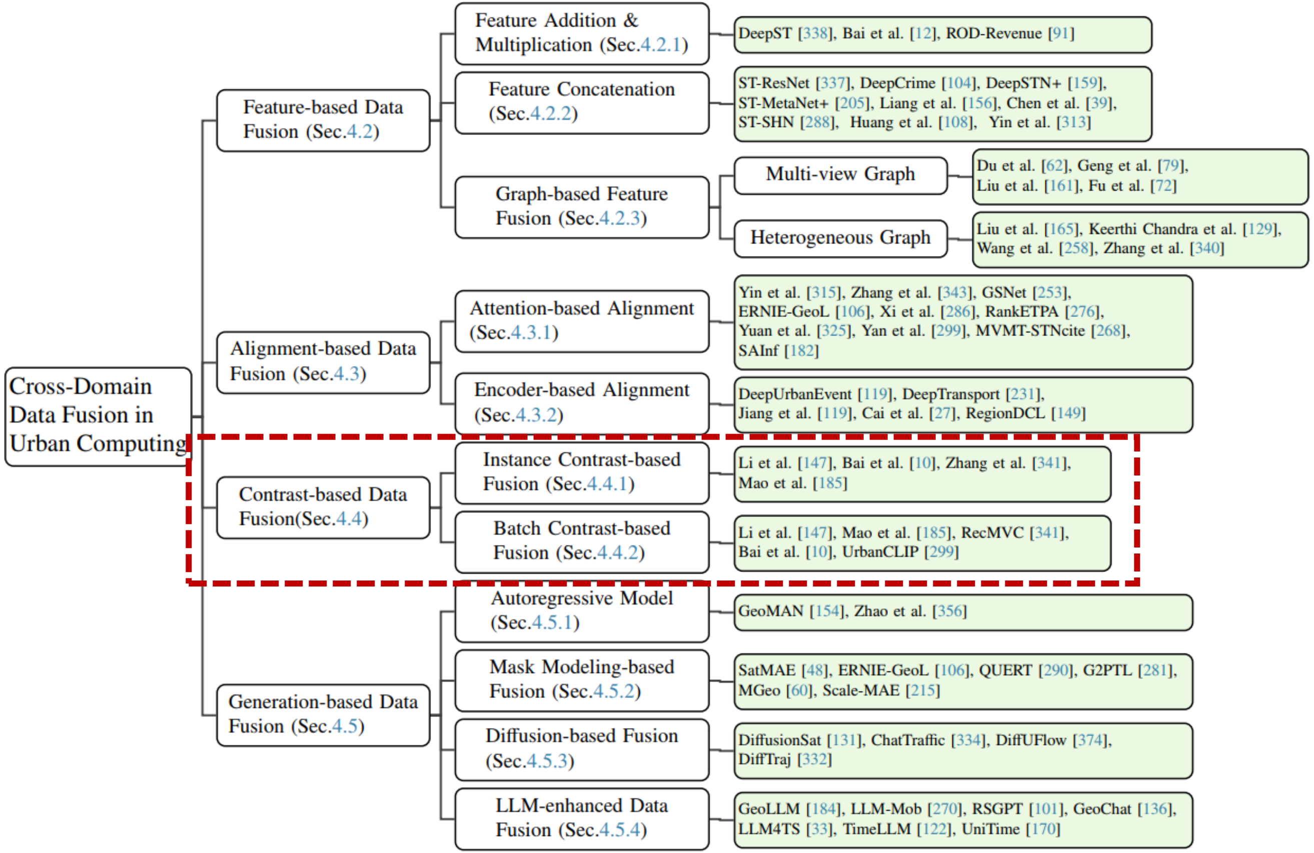
M. Zhang et al., [Multi-view joint graph representation learning for urban region embedding](), IJCAI 2021.

Encoder-based Fusion

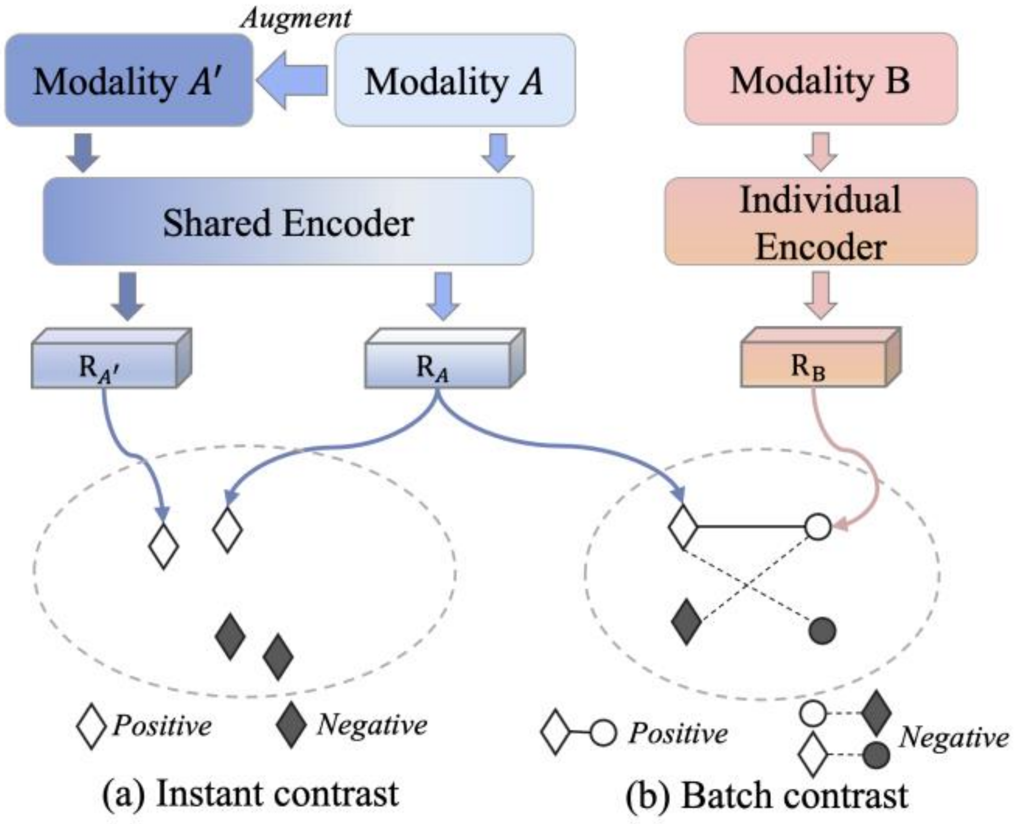


Jin et al. Time-LLM: Time Series Forecasting by Reprogramming Large Language Models. **ICLR** 2024.

DL-based Data Fusion



Contrast-based Data Fusion

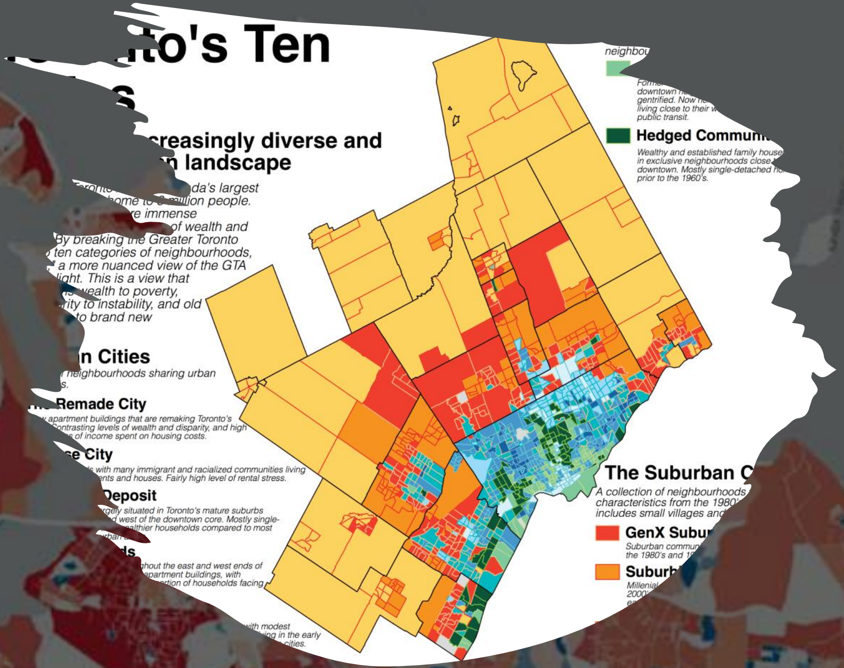


UrbanCLIP: Enhancing Urban Region Profiling with LLMs

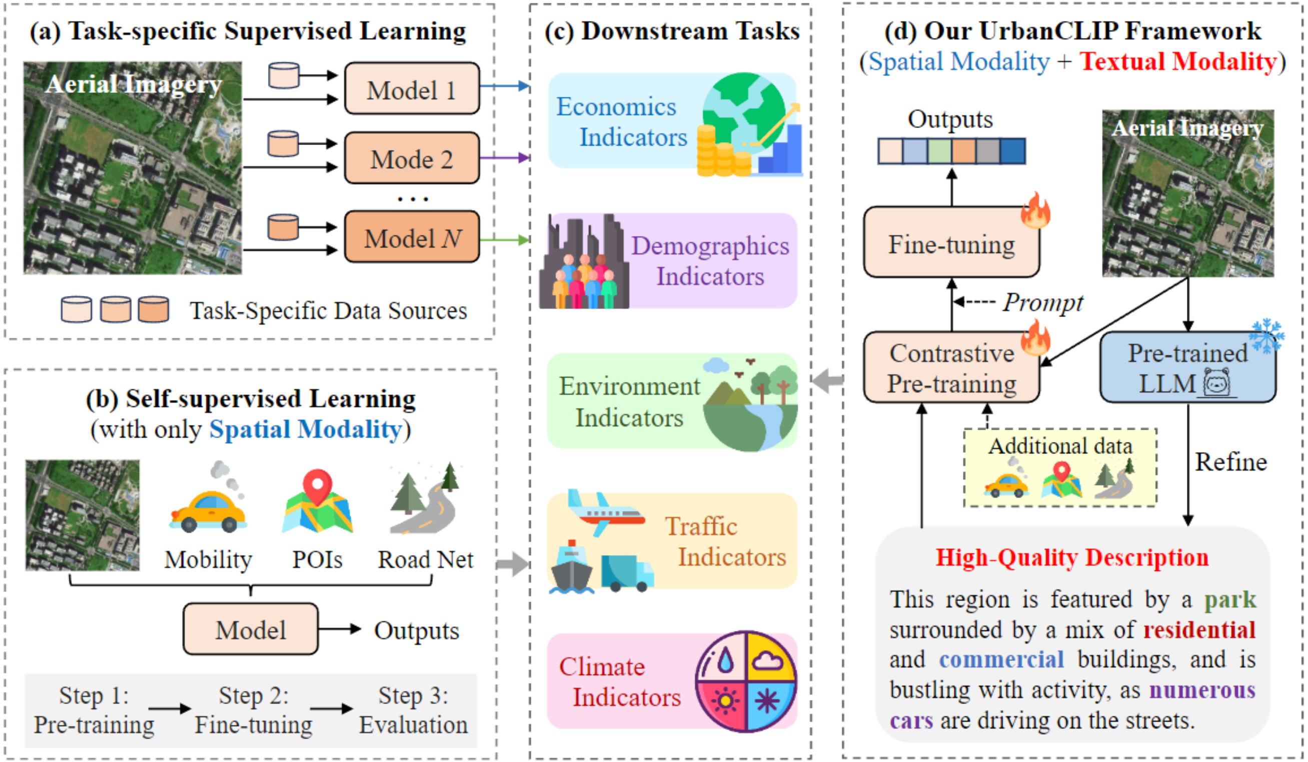
WWW 2024

Background

- The rapid pace of urbanization has led to more than half of the global population, totaling 4.4 billion inhabitants
- **Urban region profiling** is the process of representing and summarizing key features and attributes of urban areas in a lower-dimensional space

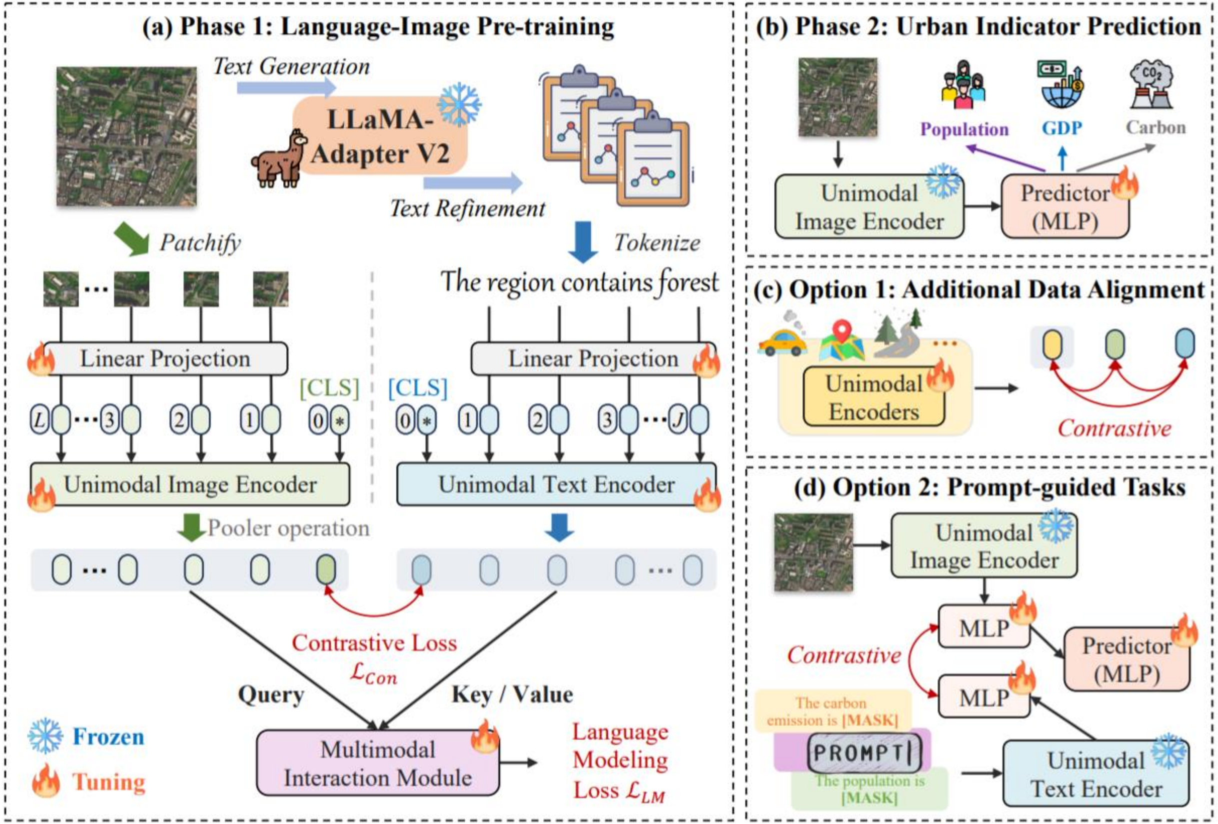


Comparison to Existing Solutions



Y. Yan, **Y. Liang*** et al., UrbanCLIP: Learning Text-enhanced Urban Region Profiling with Contrastive Language-Image Pretraining from the Web. **WWW** 2024.

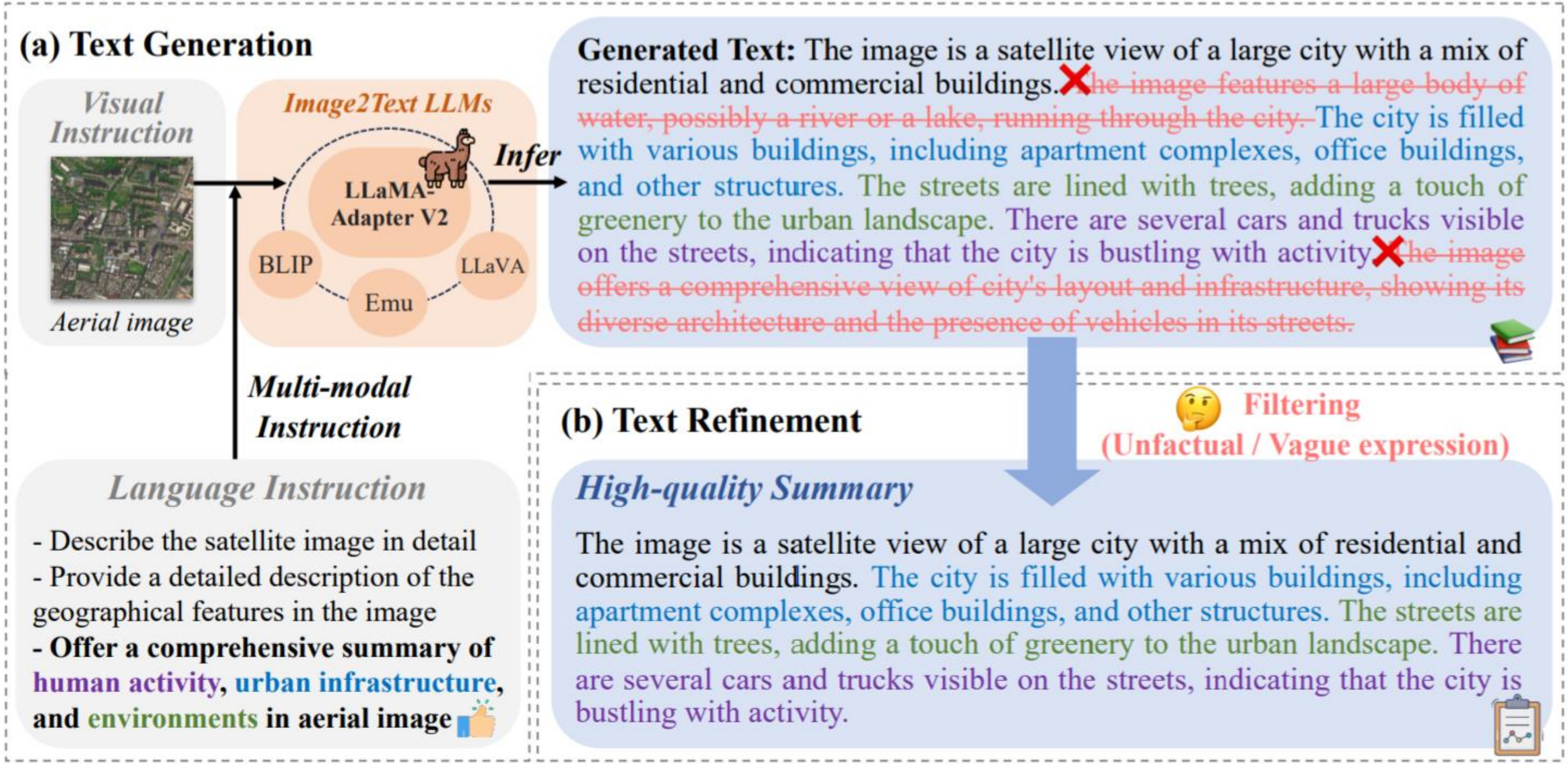
Model Architecture



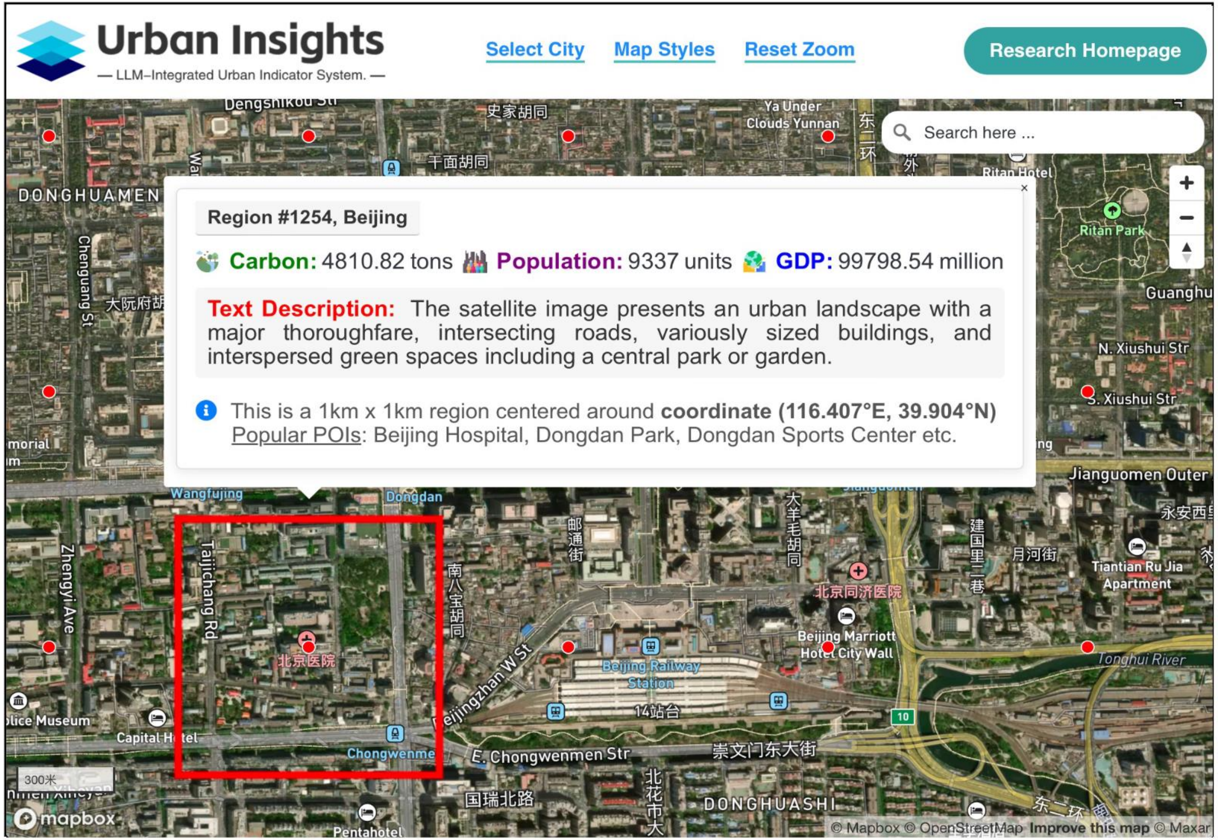
Advantages

1. Explanability
2. Generalizability

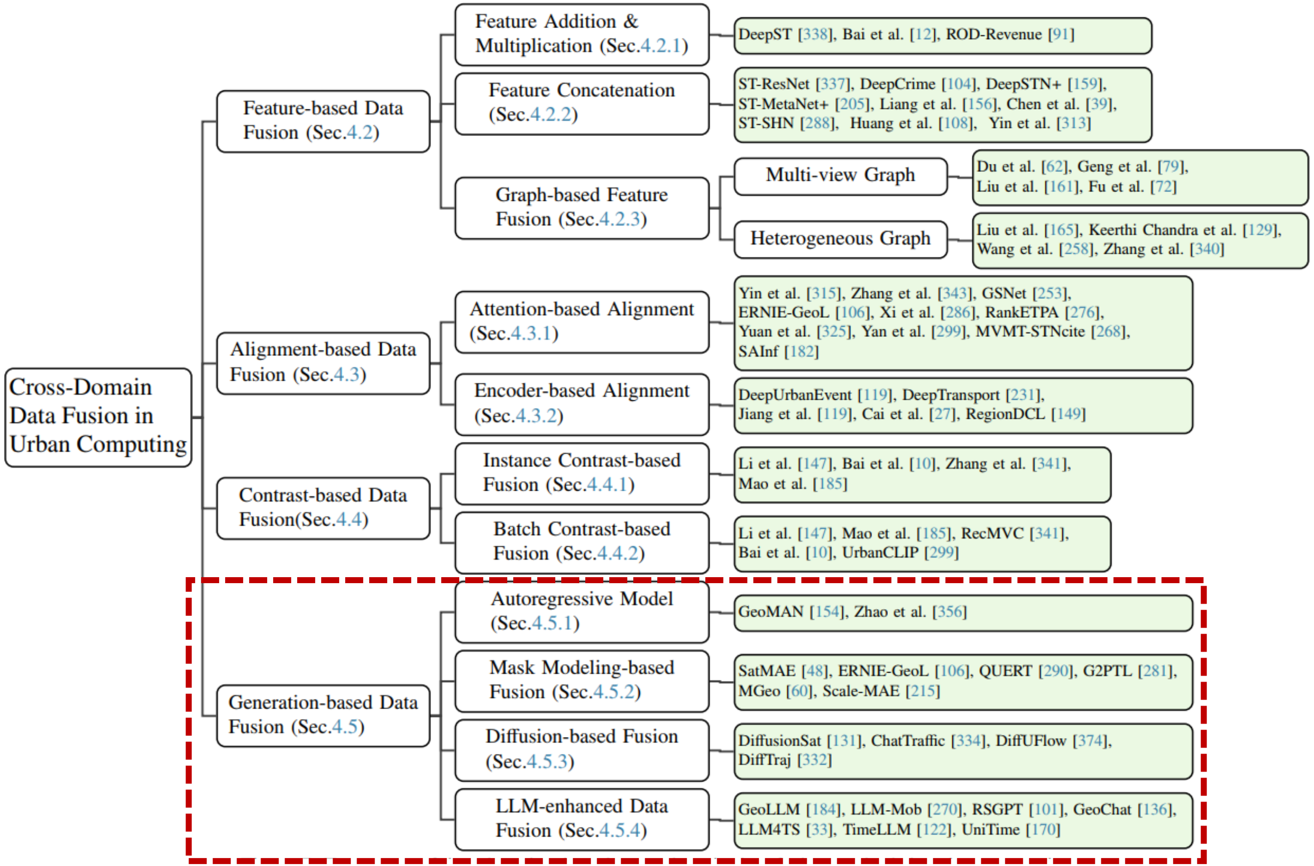
Text Generation & Refinement



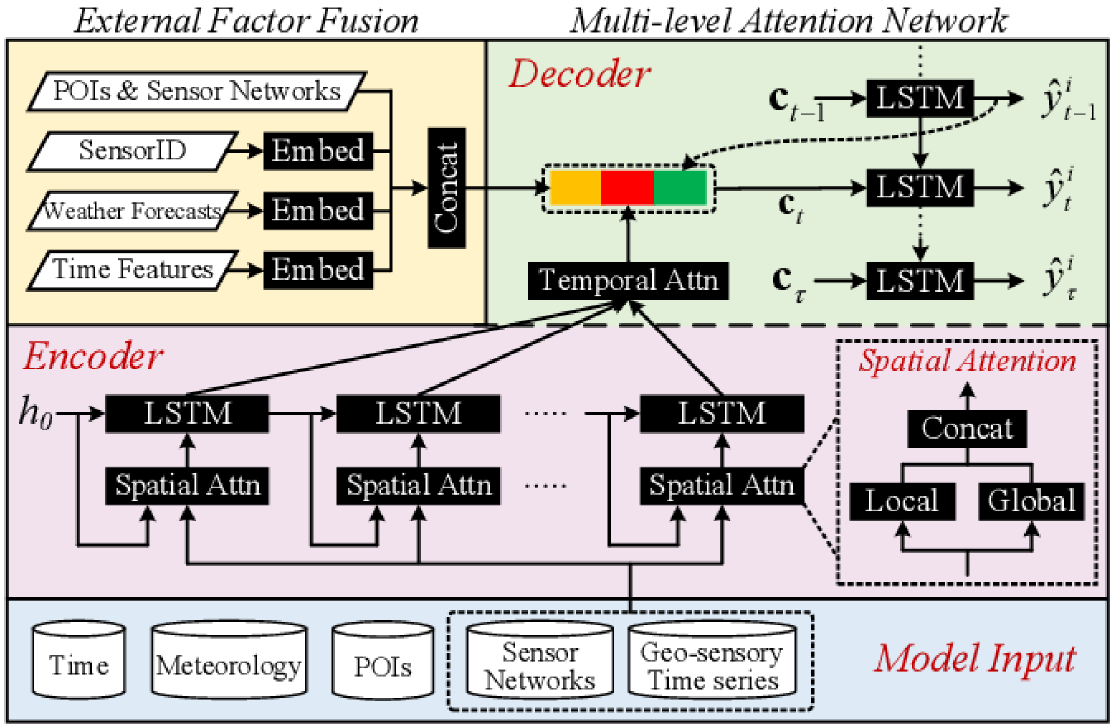
Demo


Urban Insights
— LLM–Integrated Urban Indicator System. —
Select City
Map Styles
Reset Zoom
Research Homepage
Search here ...
Region #1254, Beijing
Carbon: 4810.82 tons
Population: 9337 units
GDP: 99798.54 million
Text Description: The satellite image presents an urban landscape with a major thoroughfare, intersecting roads, variously sized buildings, and interspersed green spaces including a central park or garden.
This is a 1km x 1km region centered around coordinate (116.407°E, 39.904°N)
Popular POIs: Beijing Hospital, Dongdan Park, Dongdan Sports Center etc.
DONGHUAMEN
Wangfujing
Dongdan
Taijichang Rd
北京医院
Beijing Railway Station
Beijing Marriott Hotel City Wall
北京同济医院
E. Chongwenmen Str
崇文门东大街
DONGHUASHI
Tonghui River
Jianguomen Outer
Ritan Park
Capital Hotel
300米
mapbox
© Mapbox © OpenStreetMap Improve this map © Maxar

DL-based Data Fusion

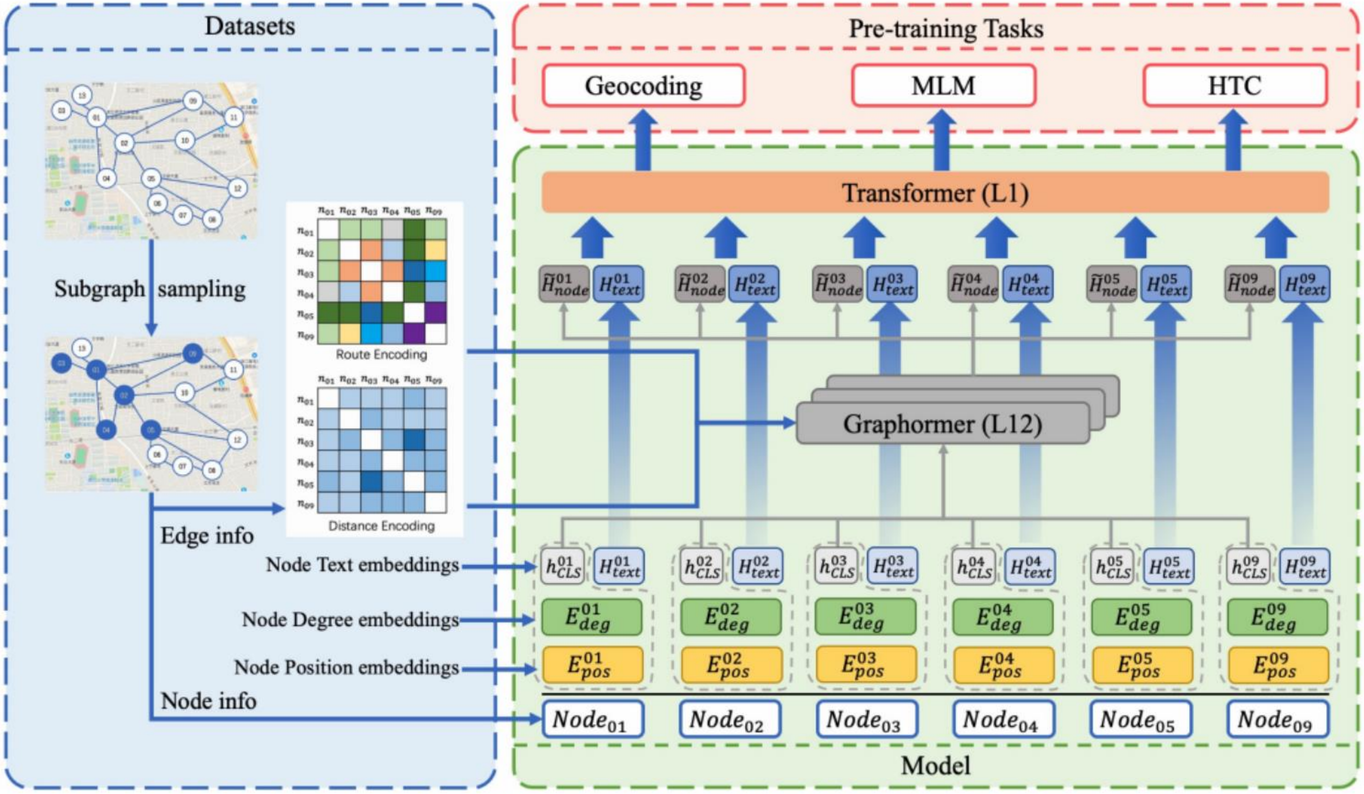


Autoregressive Models

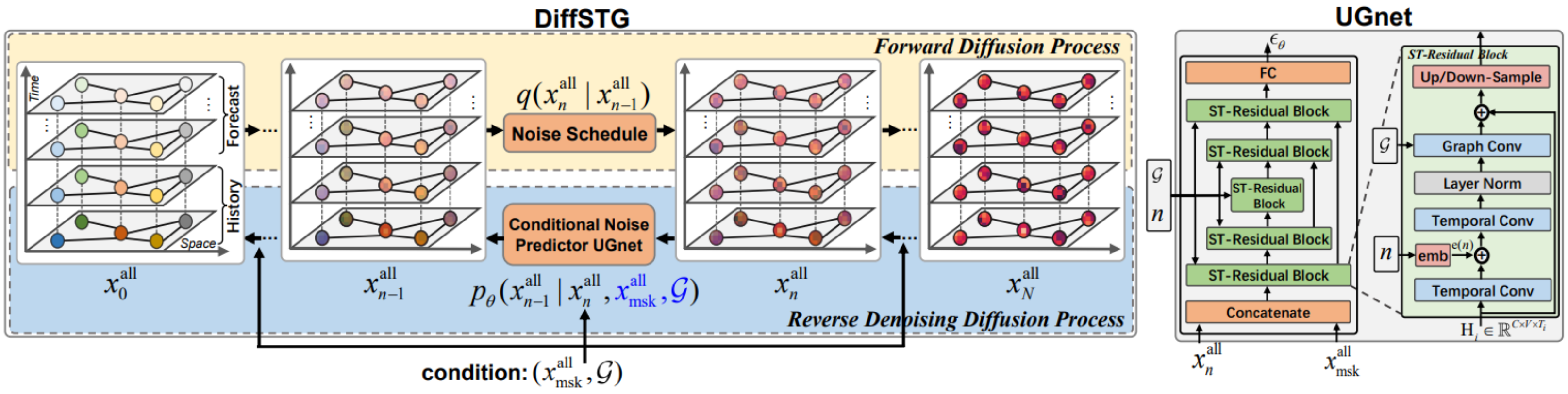


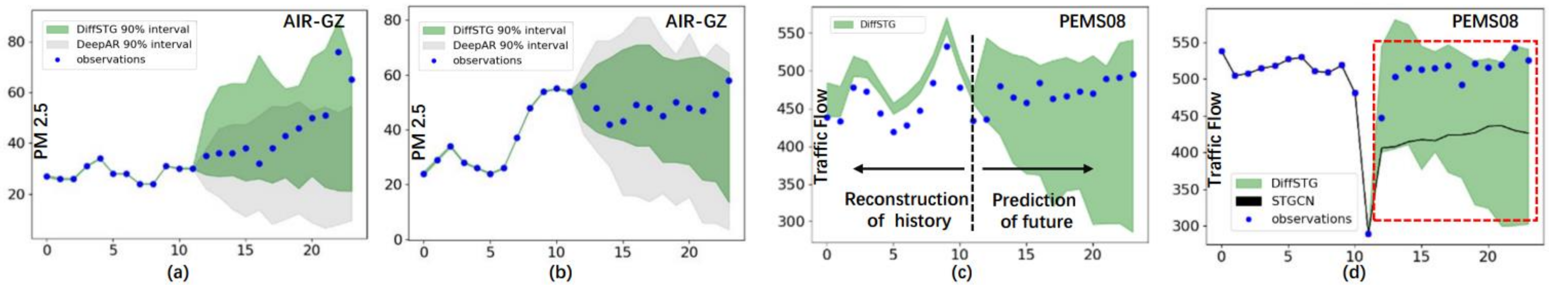
Y. Liang et al., GeoMAN: Multi-Level Attention Networks for Geo-Sensory Time Series Prediction. IJCAI 2018.

Mask Modeling-based Fusion



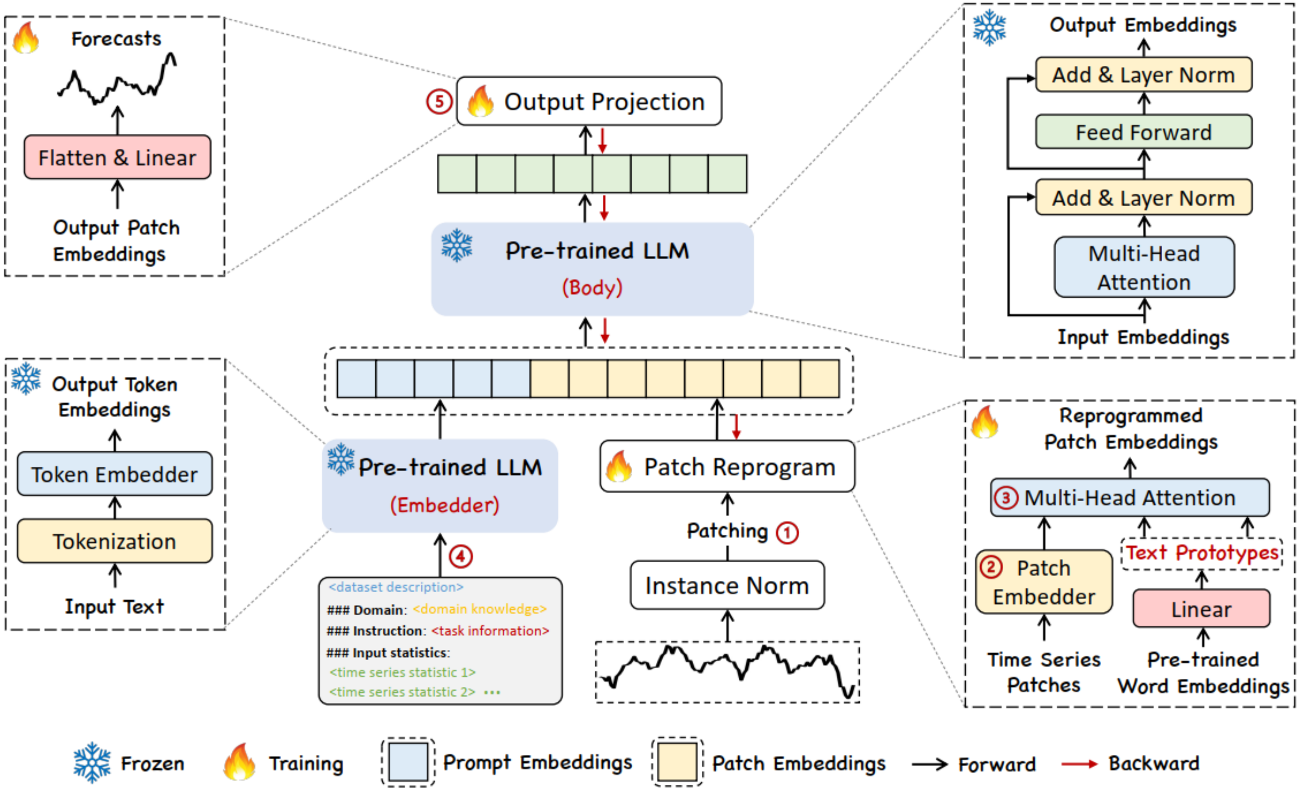
Diffusion-based Fusion





H. Wen, **Y. Liang*** et al., DiffSTG: Probabilistic Spatio-Temporal Graph Forecasting with Denoising Diffusion Models. SIGSPATIAL 2023.

LLM-based Fusion

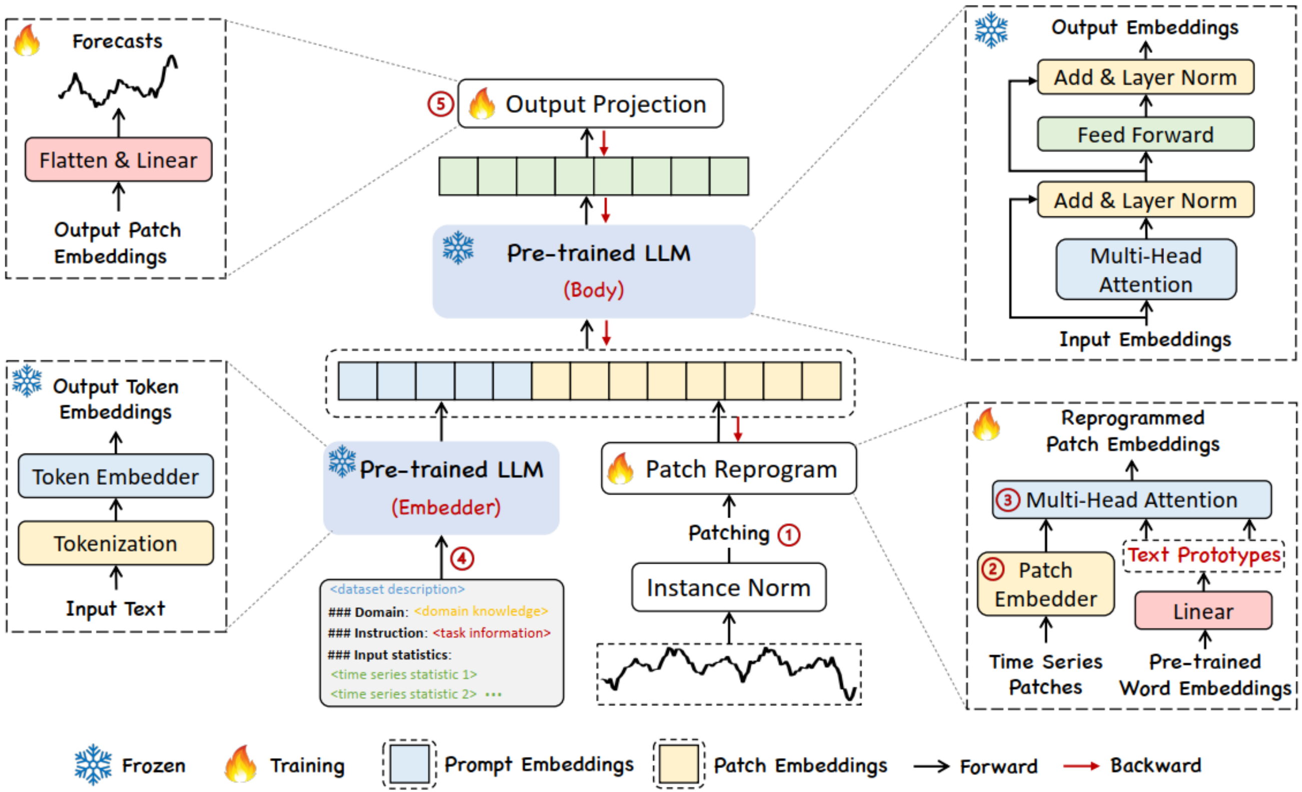


Jin et al. Time-LLM: Time Series Forecasting by Reprogramming Large Language Models. **ICLR** 2024.

Time-LLM:
LLM for Time Series Forecasting

ICLR 2024

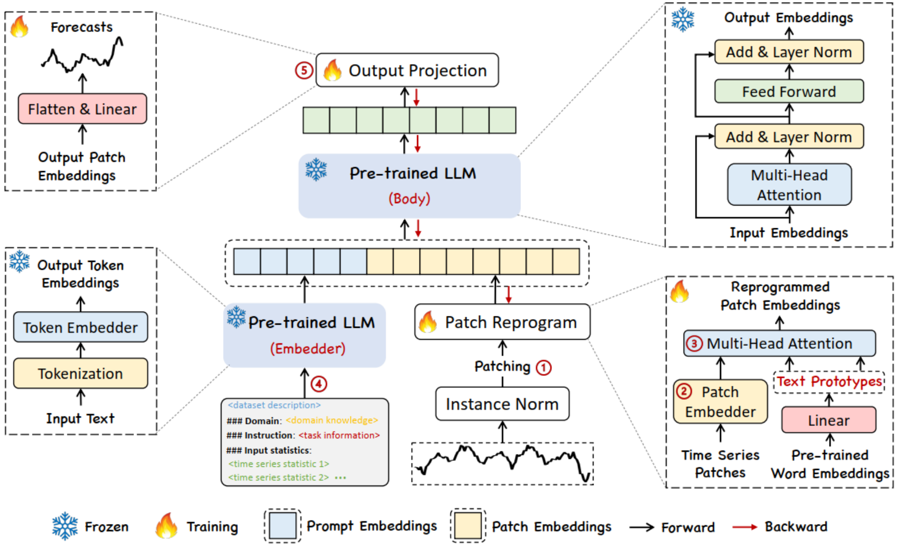
Model Framework

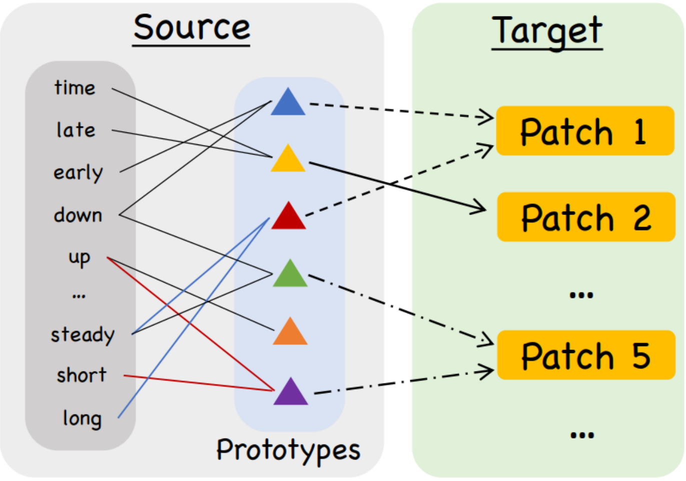


Jin et al. Time-LLM: Time Series Forecasting by Reprogramming Large Language Models. **ICLR** 2024.

Contribution 1: Patch Reprogramming

- This module aims to reprogram time series features using pre-trained word embeddings in the backbone

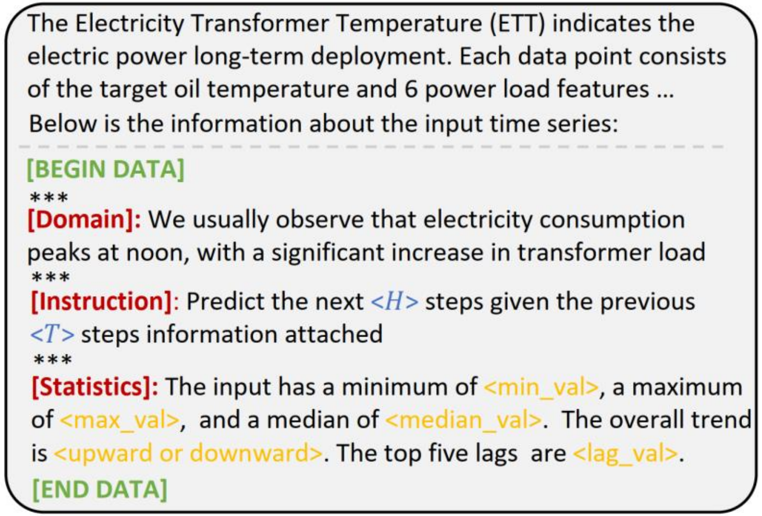


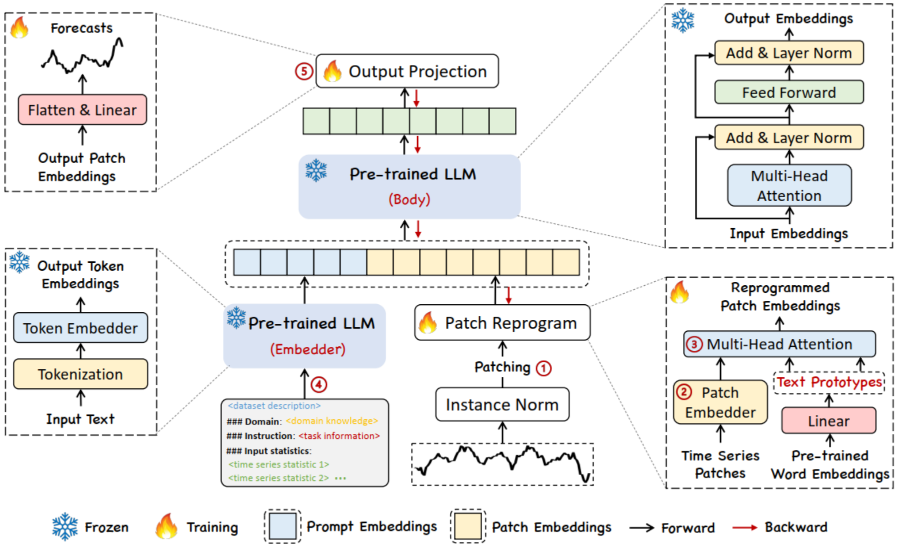


Contribution 2: Prompt-as-Profix

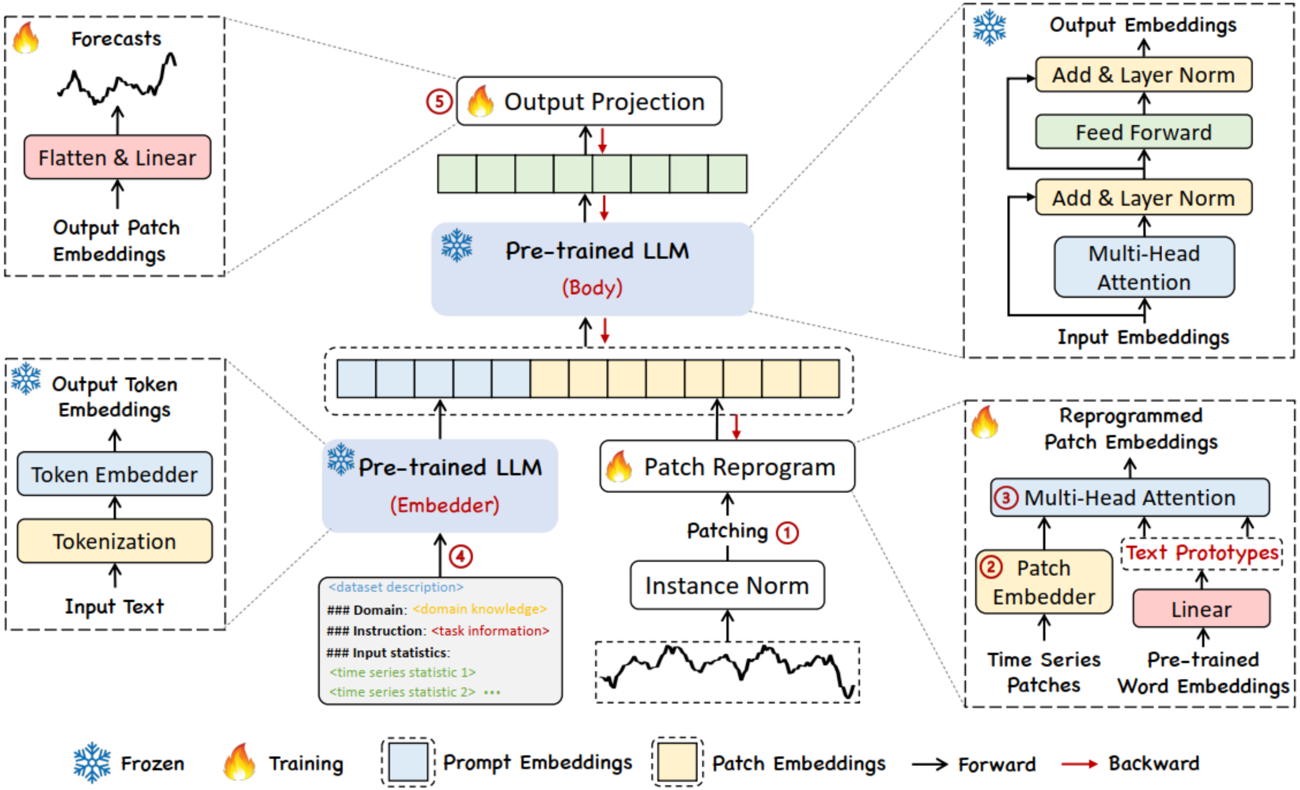
香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

- Dataset context furnishes LLM with essential background information concerning the input time series
- Task instruction serves as a crucial guide in the transformation of patch embeddings for specific tasks
- We also enrich with additional crucial statistics, such as trends and lags, to facilitate pattern recognition and reasoning





Output Projection



Results on Long-Term Forecasting

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Table 1: Long-term forecasting results. We use forecasting horizons $H \in \{96, 192, 336, 720\}$. A lower value indicates better performance. **Red**: the best, Blue: the second best.

Methods		TIME-LLM		GPT4TS		DLinear		PatchTST		TimesNet		FEDformer		Autoformer		Stationary		ETSformer		LightTS		Informer		Reformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	**0.362**	**0.392**	0.376	0.397	0.375	0.399	0.370	0.399	0.384	0.402	0.376	0.419	0.449	0.459	0.513	0.491	0.494	0.479	0.424	0.432	0.865	0.713	0.837	0.728
	192	**0.398**	0.418	0.416	0.418	0.405	**0.416**	0.413	0.421	0.436	0.429	0.420	0.448	0.500	0.482	0.534	0.504	0.538	0.504	0.475	0.462	1.008	0.792	0.923	0.766
	336	0.430	**0.427**	0.442	0.433	0.439	0.443	**0.422**	0.436	0.491	0.469	0.459	0.465	0.521	0.496	0.588	0.535	0.574	0.521	0.518	0.488	1.107	0.809	1.097	0.835
	720	**0.442**	**0.457**	0.477	0.456	0.472	0.490	0.447	0.466	0.521	0.500	0.506	0.507	0.514	0.512	0.643	0.616	0.562	0.535	0.547	0.533	1.181	0.865	1.257	0.889
	Avg	**0.408**	**0.423**	0.465	0.455	0.422	0.437	0.413	0.430	0.458	0.450	0.440	0.460	0.496	0.487	0.570	0.537	0.542	0.510	0.491	0.479	1.040	0.795	1.029	0.805
ETTh2	96	**0.268**	**0.328**	0.285	0.342	0.289	0.353	0.274	0.336	0.340	0.374	0.358	0.397	0.346	0.388	0.476	0.458	0.340	0.391	0.397	0.437	3.755	1.525	2.626	1.317
	192	**0.329**	**0.375**	0.354	0.389	0.383	0.418	0.339	0.379	0.402	0.414	0.429	0.439	0.456	0.452	0.512	0.493	0.430	0.439	0.520	0.504	5.602	1.931	11.12	2.979
	336	0.368	0.409	0.373	0.407	0.448	0.465	**0.329**	**0.380**	0.452	0.452	0.496	0.487	0.482	0.486	0.552	0.551	0.485	0.479	0.626	0.559	4.721	1.835	9.323	2.769
	720	**0.372**	**0.420**	0.406	0.441	0.605	0.551	0.379	0.422	0.462	0.468	0.463	0.474	0.515	0.511	0.562	0.560	0.500	0.497	0.863	0.672	3.647	1.625	3.874	1.697
	Avg	0.334	0.383	0.381	0.412	0.431	0.446	**0.330**	**0.379**	0.414	0.427	0.437	0.449	0.450	0.459	0.526	0.516	0.439	0.452	0.602	0.543	4.431	1.729	6.736	2.191
ETTm1	96	**0.272**	**0.334**	0.292	0.346	0.299	0.343	0.290	0.342	0.338	0.375	0.379	0.419	0.505	0.475	0.386	0.398	0.375	0.398	0.374	0.400	0.672	0.571	0.538	0.528
	192	**0.310**	**0.358**	0.332	0.372	0.335	0.365	0.332	0.369	0.374	0.387	0.426	0.441	0.553	0.496	0.459	0.444	0.408	0.410	0.400	0.407	0.795	0.669	0.658	0.592
	336	**0.352**	**0.384**	0.366	0.394	0.369	0.386	0.366	0.392	0.410	0.411	0.445	0.459	0.621	0.537	0.495	0.464	0.435	0.428	0.438	0.438	1.212	0.871	0.898	0.721
	720	**0.383**	**0.411**	0.417	0.421	0.425	0.421	0.416	0.420	0.478	0.450	0.543	0.490	0.671	0.561	0.585	0.516	0.499	0.462	0.527	0.502	1.166	0.823	1.102	0.841
	Avg	**0.329**	**0.372**	0.388	0.403	0.357	0.378	0.351	0.380	0.400	0.406	0.448	0.452	0.588	0.517	0.481	0.456	0.429	0.425	0.435	0.437	0.961	0.734	0.799	0.671
ETTm2	96	**0.161**	**0.253**	0.173	0.262	0.167	0.269	0.165	0.255	0.187	0.267	0.203	0.287	0.255	0.339	0.192	0.274	0.189	0.280	0.209	0.308	0.365	0.453	0.658	0.619
	192	**0.219**	**0.293**	0.229	0.301	0.224	0.303	0.220	0.292	0.249	0.309	0.269	0.328	0.281	0.340	0.280	0.339	0.253	0.319	0.311	0.382	0.533	0.563	1.078	0.827
	336	**0.271**	**0.329**	0.286	0.341	0.281	0.342	0.274	**0.329**	0.321	0.351	0.325	0.366	0.339	0.372	0.334	0.361	0.314	0.357	0.442	0.466	1.363	0.887	1.549	0.972
	720	**0.352**	**0.379**	0.378	0.401	0.397	0.421	0.362	0.385	0.408	0.403	0.421	0.415	0.433	0.432	0.417	0.413	0.414	0.413	0.675	0.587	3.379	1.338	2.631	1.242
	Avg	**0.251**	**0.313**	0.284	0.339	0.267	0.333	0.255	0.315	0.291	0.333	0.305	0.349	0.327	0.371	0.306	0.347	0.293	0.342	0.409	0.436	1.410	0.810	1.479	0.915
1st Count		**18**		0		1		4		0		0		0		0		0		0		0		0	

Results on Few-Shot Learning

Table 3: Few-shot learning on 10% training data. We use the same protocol and notations as in Tab. 1.

Methods		TIME-LLM		GPT4TS		DLinear		PatchTST		TimesNet		FEDformer		Autoformer		Stationary		ETSformer		LightTS		Informer		Reformer	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTh1	96	**0.448**	0.460	0.458	**0.456**	0.492	0.495	0.516	0.485	0.861	0.628	0.512	0.499	0.613	0.552	0.918	0.639	1.112	0.806	1.298	0.838	1.179	0.792	1.184	0.790
	192	**0.484**	**0.483**	0.570	0.516	0.565	0.538	0.598	0.524	0.797	0.593	0.624	0.555	0.722	0.598	0.915	0.629	1.155	0.823	1.322	0.854	1.199	0.806	1.295	0.850
	336	**0.589**	0.540	0.608	**0.535**	0.721	0.622	0.657	0.550	0.941	0.648	0.691	0.574	0.750	0.619	0.939	0.644	1.179	0.832	1.347	0.870	1.202	0.811	1.294	0.854
	720	**0.700**	**0.604**	0.725	0.591	0.986	0.743	0.762	0.610	0.877	0.641	0.728	0.614	0.721	0.616	0.887	0.645	1.273	0.874	1.534	0.947	1.217	0.825	1.223	0.838
	Avg.	**0.556**	**0.522**	0.590	0.525	0.691	0.600	0.633	0.542	0.869	0.628	0.639	0.561	0.702	0.596	0.915	0.639	1.180	0.834	1.375	0.877	1.199	0.809	1.249	0.833
ETTh2	96	**0.275**	**0.326**	0.331	0.374	0.357	0.411	0.353	0.389	0.378	0.409	0.382	0.416	0.413	0.451	0.389	0.411	0.678	0.619	2.022	1.006	3.837	1.508	3.788	1.533
	192	**0.374**	**0.373**	0.402	0.411	0.569	0.519	0.403	0.414	0.490	0.467	0.478	0.474	0.474	0.477	0.473	0.455	0.785	0.666	2.329	1.104	3.856	1.513	3.552	1.483
	336	**0.406**	**0.429**	**0.406**	0.433	0.671	0.572	0.426	0.441	0.537	0.494	0.504	0.501	0.547	0.543	0.507	0.480	0.839	0.694	2.453	1.122	3.952	1.526	3.395	1.526
	720	**0.427**	**0.449**	0.449	0.464	0.824	0.648	0.477	0.480	0.510	0.491	0.499	0.509	0.516	0.523	0.477	0.472	1.273	0.874	3.816	1.407	3.842	1.503	3.205	1.401
	Avg.	**0.370**	**0.394**	0.397	0.421	0.605	0.538	0.415	0.431	0.479	0.465	0.466	0.475	0.488	0.499	0.462	0.455	0.894	0.713	2.655	1.160	3.872	1.513	3.485	1.486
ETTm1	96	**0.346**	**0.388**	0.390	0.404	0.352	0.392	0.410	0.419	0.583	0.501	0.578	0.518	0.774	0.614	0.761	0.568	0.911	0.688	0.921	0.682	1.162	0.785	1.442	0.847
	192	**0.373**	0.416	0.429	0.423	0.382	**0.412**	0.437	0.434	0.630	0.528	0.617	0.546	0.754	0.592	0.781	0.574	0.955	0.703	0.957	0.701	1.172	0.793	1.444	0.862
	336	**0.413**	**0.426**	0.469	0.439	0.419	0.434	0.476	0.454	0.725	0.568	0.998	0.775	0.869	0.677	0.803	0.587	0.991	0.719	0.998	0.716	1.227	0.908	1.450	0.866
	720	**0.485**	0.476	0.569	0.498	0.490	**0.477**	0.681	0.556	0.769	0.549	0.693	0.579	0.810	0.630	0.844	0.581	1.062	0.747	1.007	0.719	1.207	0.797	1.366	0.850
	Avg.	**0.404**	**0.427**	0.464	0.441	0.411	0.429	0.501	0.466	0.677	0.537	0.722	0.605	0.802	0.628	0.797	0.578	0.980	0.714	0.971	0.705	1.192	0.821	1.426	0.856
ETTm2	96	**0.177**	**0.261**	0.188	0.269	0.213	0.303	0.191	0.274	0.212	0.285	0.291	0.399	0.352	0.454	0.229	0.308	0.331	0.430	0.813	0.688	3.203	1.407	4.195	1.628
	192	**0.241**	**0.314**	0.251	0.309	0.278	0.345	0.252	0.317	0.270	0.323	0.307	0.379	0.694	0.691	0.291	0.343	0.400	0.464	1.008	0.768	3.112	1.387	4.042	1.601
	336	**0.274**	**0.327**	0.307	0.346	0.338	0.385	0.306	0.353	0.323	0.353	0.543	0.559	2.408	1.407	0.348	0.376	0.469	0.498	1.031	0.775	3.255	1.421	3.963	1.585
	720	**0.417**	**0.390**	0.426	0.417	0.436	0.440	0.433	0.427	0.474	0.449	0.712	0.614	1.913	1.166	0.461	0.438	0.589	0.557	1.096	0.791	3.909	1.543	3.711	1.532
	Avg.	**0.277**	**0.323**	0.293	0.335	0.316	0.368	0.296	0.343	0.320	0.353	0.463	0.488	1.342	0.930	0.332	0.366	0.447	0.487	0.987	0.756	3.370	1.440	3.978	1.587
1st Count		**20**		3		2		0		0		0		0		0		0		0		0		0	

Results on Zero-Shot Learning

Table 5: Zero-shot learning results. **Red**: the best, Blue: the second best. Appendix D shows complete results.

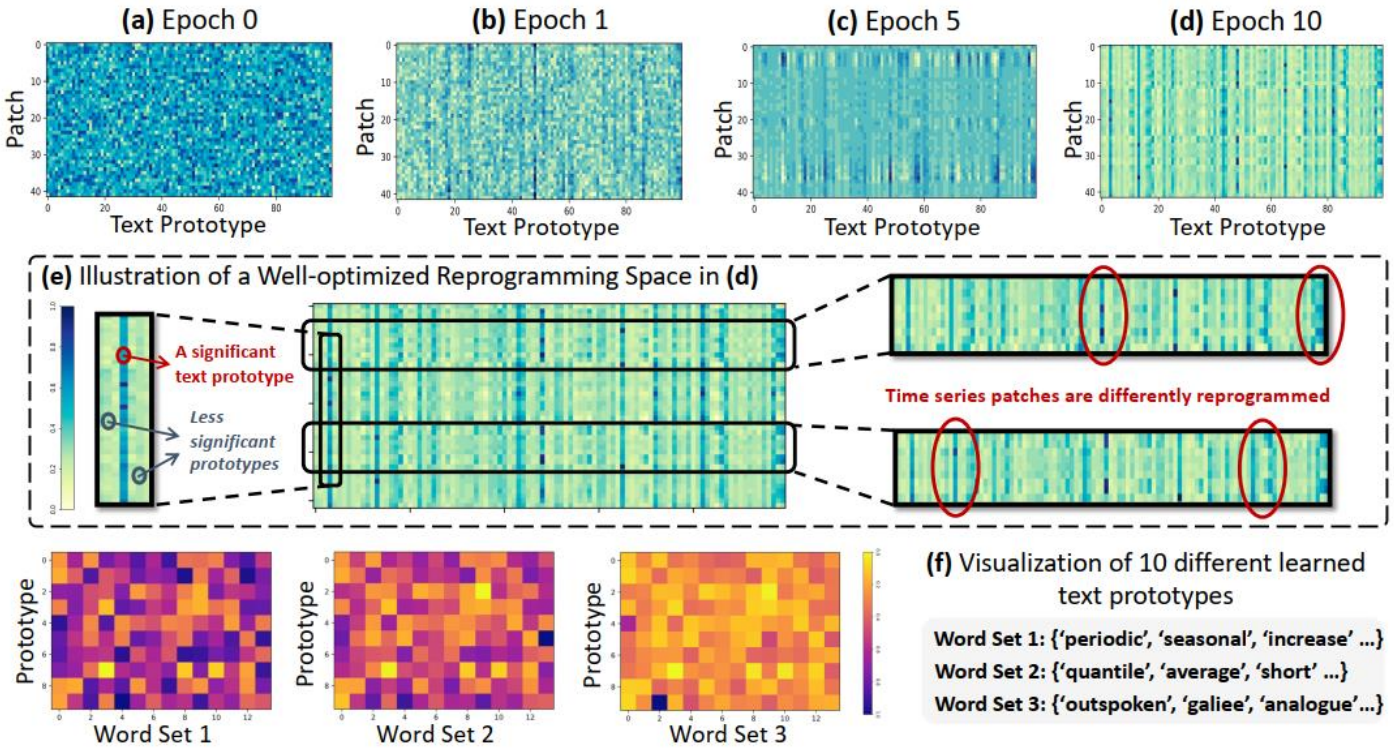
Methods	Time-LLM		GPT4TS		DLinear		PatchTST		TimesNet		Autoformer	
Metric	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
$ETTh1 \rightarrow ETTh2$	**0.353**	**0.387**	0.406	0.422	0.493	0.488	0.380	0.405	0.421	0.431	0.582	0.548
$ETTh1 \rightarrow ETTm2$	**0.273**	**0.340**	0.325	0.363	0.415	0.452	0.314	0.360	0.327	0.361	0.457	0.483
$ETTh2 \rightarrow ETTh1$	**0.479**	**0.474**	0.757	0.578	0.703	0.574	0.565	0.513	0.865	0.621	0.757	0.608
$ETTh2 \rightarrow ETTm2$	**0.272**	**0.341**	0.335	0.370	0.328	0.386	0.325	0.365	0.342	0.376	0.366	0.411
$ETTm1 \rightarrow ETTh2$	**0.381**	**0.412**	0.433	0.439	0.464	0.475	0.439	0.438	0.457	0.454	0.470	0.479
$ETTm1 \rightarrow ETTm2$	**0.268**	**0.320**	0.313	0.348	0.335	0.389	0.296	0.334	0.322	0.354	0.469	0.484
$ETTm2 \rightarrow ETTh2$	**0.354**	**0.400**	0.435	0.443	0.455	0.471	0.409	0.425	0.435	0.443	0.423	0.439
$ETTm2 \rightarrow ETTm1$	**0.414**	**0.438**	0.769	0.567	0.649	0.537	0.568	0.492	0.769	0.567	0.755	0.591

Ablation Study

香港科技大学(广州)
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)

Variant	Long-term Forecasting				Few-shot Forecasting			
	ETTh1-96	ETTh1-192	ETTm1-96	ETThm1-192	ETTh1-96	ETTh1-192	ETTm1-96	ETThm1-192
A.1 Llama (**Default**; 32)	**0.362**	**0.398**	**0.272**	**0.310**	**0.448**	**0.484**	**0.346**	**0.373**
A.2 Llama (8)	0.389	0.412	0.297	0.329	0.567	0.632	0.451	0.490
A.3 GPT-2 (12)	0.385	0.419	0.306	0.332	0.548	0.617	0.447	0.509
A.4 GPT-2 (6)	0.394	0.427	0.311	0.342	0.571	0.640	0.468	0.512
B.1 w/o Patch Reprogramming	0.410	0.412	0.310	0.342	0.498	0.570	0.445	0.487
B.2 w/o Prompt-as-Prefix	0.398	0.423	0.298	0.339	0.521	0.617	0.432	0.481
C.1 w/o Dataset Context	0.402	0.417	0.298	0.331	0.491	0.538	0.392	0.447
C.2 w/o Task Instruction	0.388	0.420	0.285	0.327	0.476	0.529	0.387	0.439
C.3 w/o Statistical Context	0.391	0.419	0.279	0.347	0.483	0.547	0.421	0.461

Visualization on Reprogramming



Presented Papers

Predicting Parking Availability in Singapore with Cross-Domain Data: A New Dataset and	/00890-learning-to-gene	IJCAI	2024	HKUST(GZ)	Qiongyan WANG
BEVFusion: Multi-Task Multi-Sensor Fusion with Unified Bird's-Eye View Representation	link	ICLR	2022	MIT	Pei Liu
UrbanCross: Enhancing Satellite Image-Text Retrieval with Cross-Domain Adaptation	Link	ACM MM	2024	HKUST(GZ)	Yongzi Yu

Q&A

Thanks!

CityMind Lab