DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

Xintong Wang[1] , Haihan Nan[2], Ruidong Li[1] , Huaming Wu[3]

1. Kanazawa University, Japan
2. George Mason University, USA
3. Tianjin University, China

IEEE GLOBECOM®

December 11, 2025

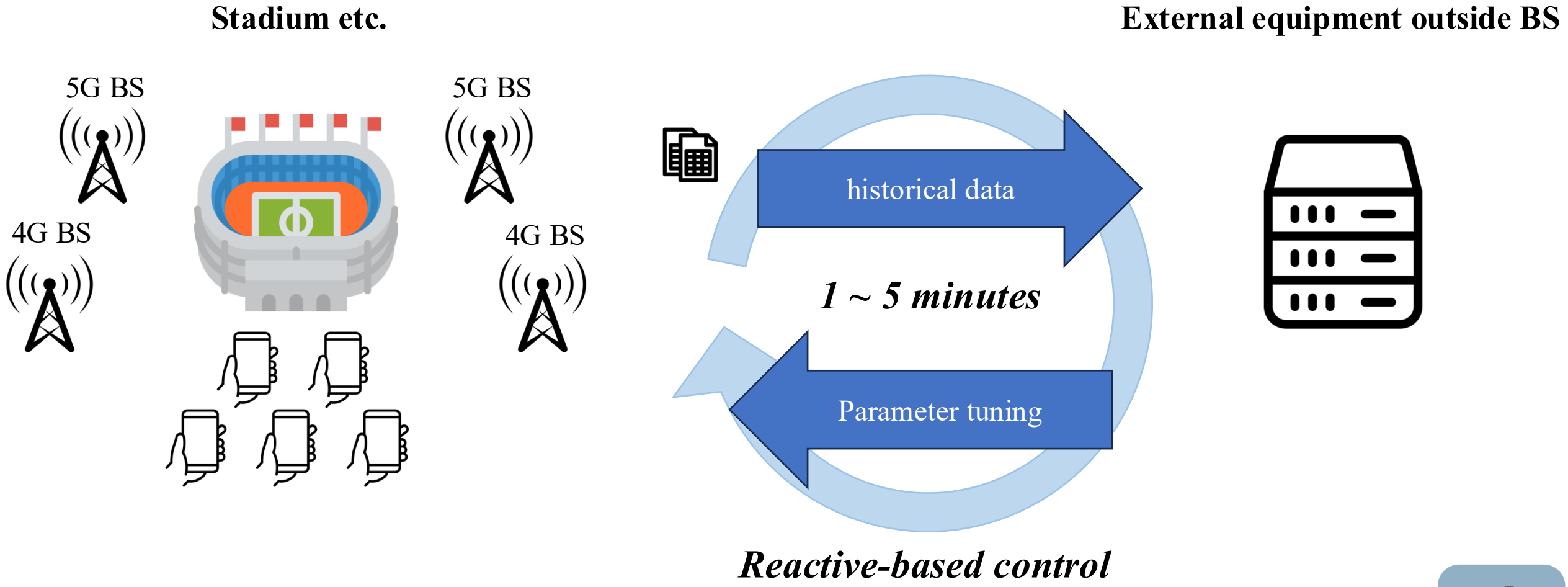
1. Background and Motivation
2. Overview of DP-LET Framework
3. Experimental results
4. Conclusion and Future work

1. **Background and Motivation**
2. Overview of DP-LET Framework
3. Experimental results
4. Conclusion and Future work

DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

High-speed automatic optimization[1]

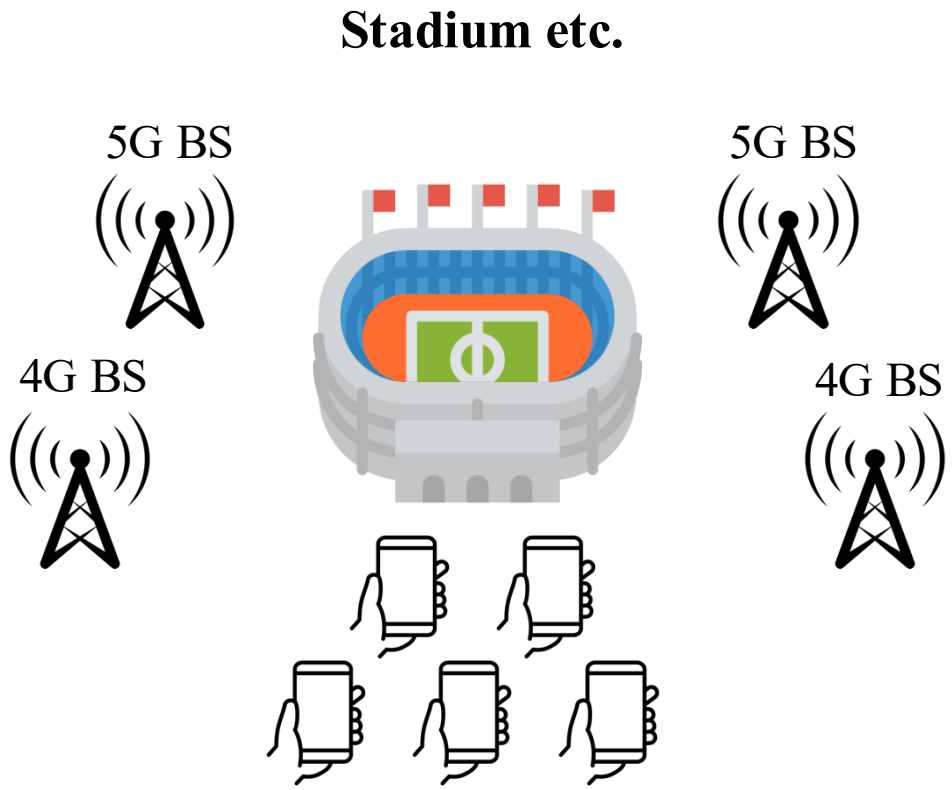


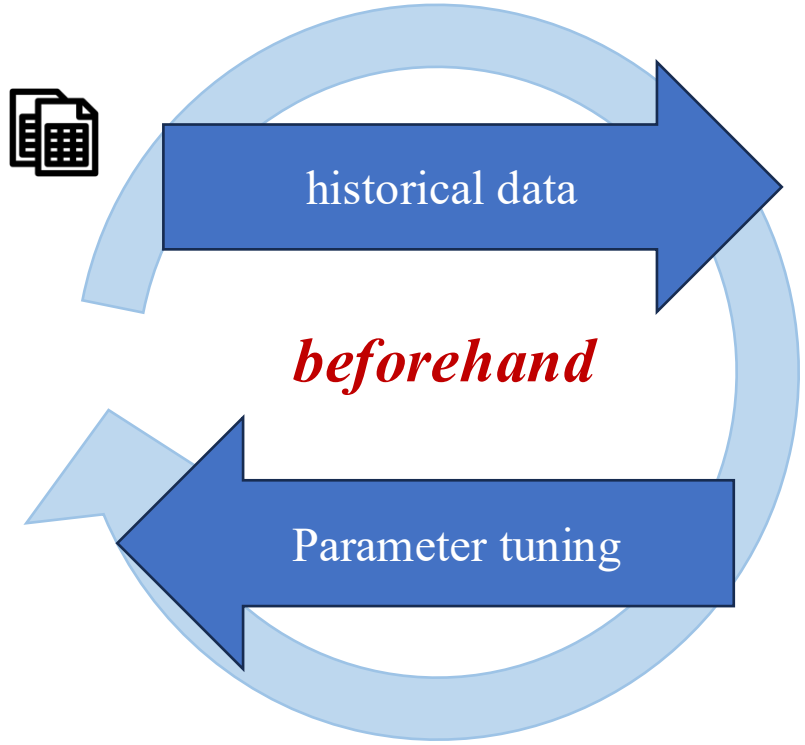
Reactive-based control

[1] SoftBank Corp. and Ericsson Japan K.K., "SoftBank Corp. and Ericsson Japan K.K. successfully demonstrate high-speed automatic optimization of 5G networks through external control of base stations," Press Release, Feb. 22, 2024. Available: https://www.softbank.jp/en/corp/news/press/sbkk/2024/20240222_01/

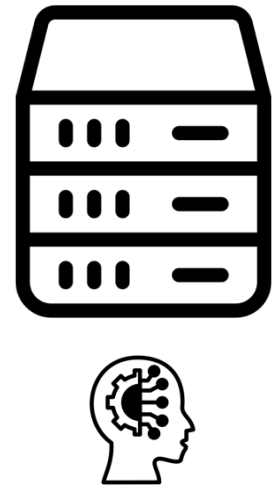
DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu



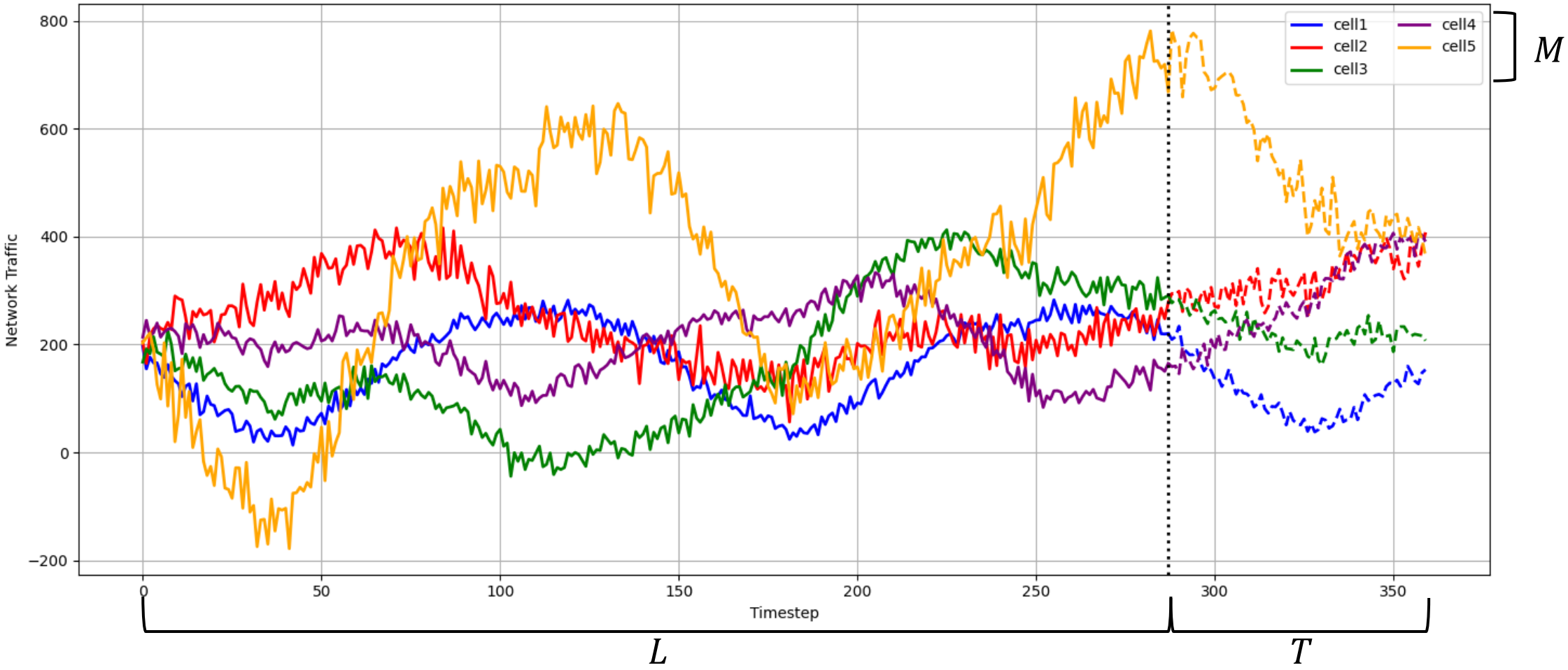


Predictive-based control



DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu



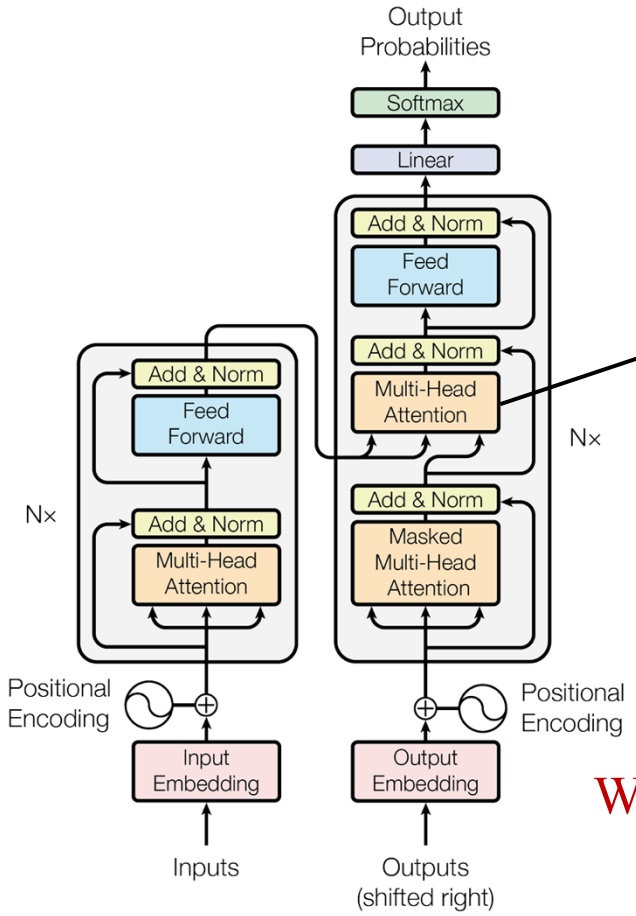
Input : Historical Traffic $X \in R^{M \times L}$

Output : Future Traffic $\hat{X} \in R^{M \times T}$

where M is the regions, L is history length, T is the horizon

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

Temporal Modeling：Transformer-based method



$$Attention\ Weight(i,j) = \frac{\exp(\frac{q_i \cdot k_j}{\sqrt{d_k}})}{\sum_{j'=1}^{n} \exp(\frac{q_i \cdot k_{j'}}{\sqrt{d_k}})}$$

q_i , k_j : linear mappings of the input series

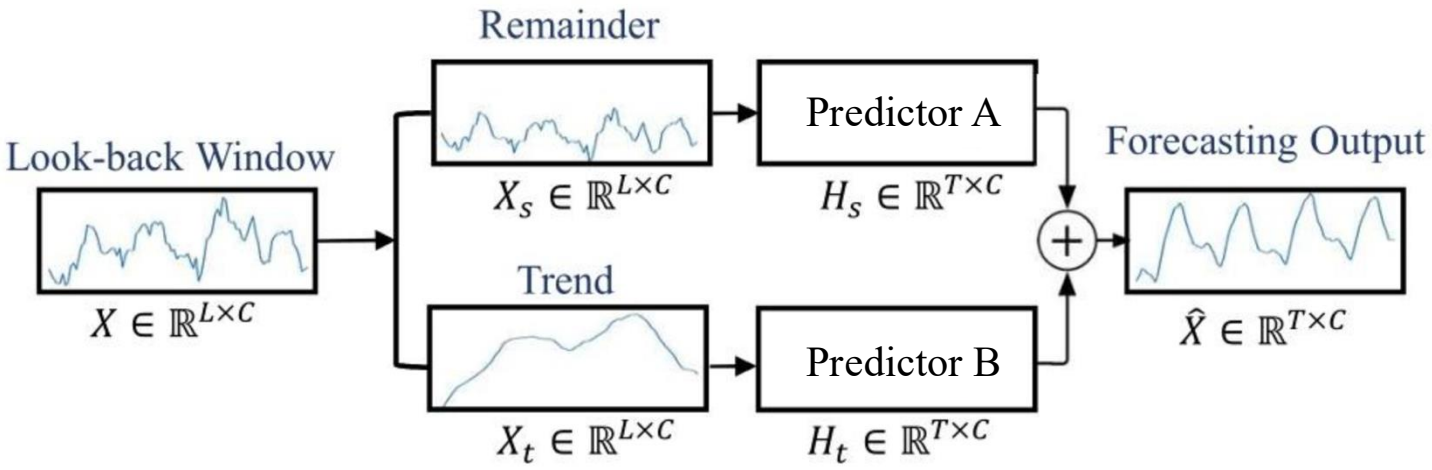
d_k : constant

n : length of the input series

When n is long, the temporal dependencies become diluted and harder to capture.

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

Temporal Modeling：Decomposition method



Making separate predictions results in almost twice as many parameters.

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

Spatial Modeling：Graph-based method

(A) REAL-WORLD BASE STATION NETWORK (NODES)



REAL-WORLD BASE STATION NETWORK (NODES)

(B) NETWORK WITH NEW NODE ADDED



NETWORK WITH NEW NODE ADDED

(C) ORIGINAL ADJACENCY MATRIX

	N1	N2	N3	N4
N1	1	0	0	0
N2	0	1	1	0
N3	0	1	1	1
N4	0	0	1	1

(D) NEW ADJACENCY MATRIX (COMPLETELY MODIFIED)

	N1	N2	N3	N4	N_{NEW}
N1	1	0	0	0	1
N2	0	1	1	1	1
N3	0	0	1	1	1
N4	1	1	0	1	1
N_{NEW}	1	1	1	1	1

Adding or removing a node requires a complete redesign of the adjacency matrix.

Existing Problems and Limitations

P1. Heavy parameter burden in decomposition-based methods

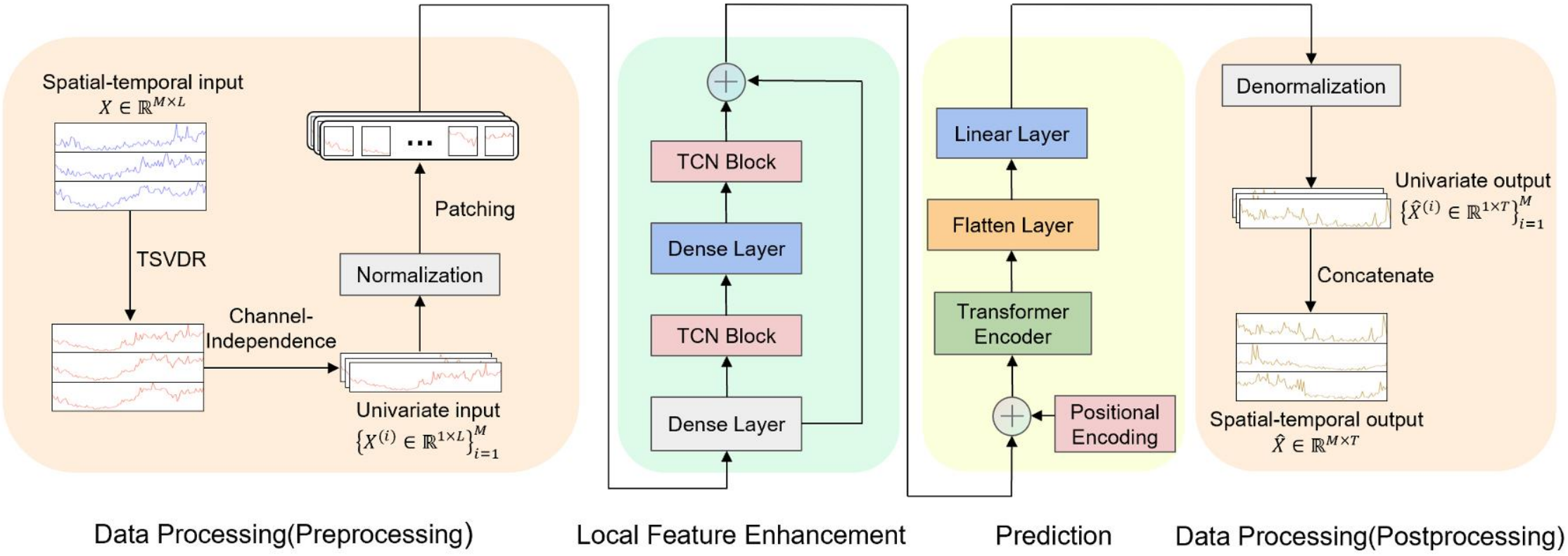
P2. Weak local modeling in Transformer-based methods

P3. Poor adaptability of graph-based methods to node changes

1. Background and Motivation
2. **Overview of DP-LET Framework**
3. Experimental results
4. Conclusion and Future work

DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

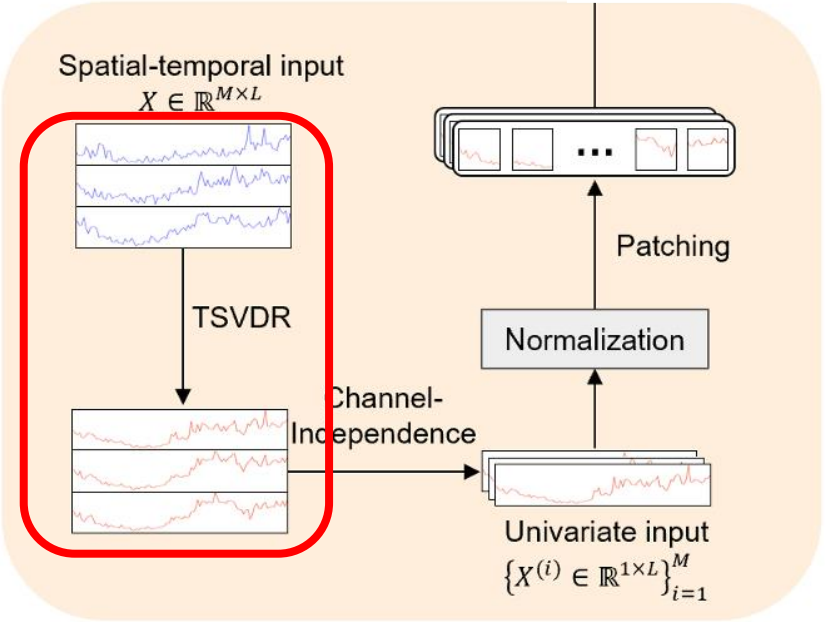
Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu



DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

P1. Heavy parameter burden in decomposition-based methods



Data Processing(Preprocessing)

Algorithm 1: Data Processing Module (Preprocessing Part)

Data: $X \in \mathbb{R}^{M \times L}$, threshold $c \in \mathbb{R}_+$, segment length ℓ, stride S

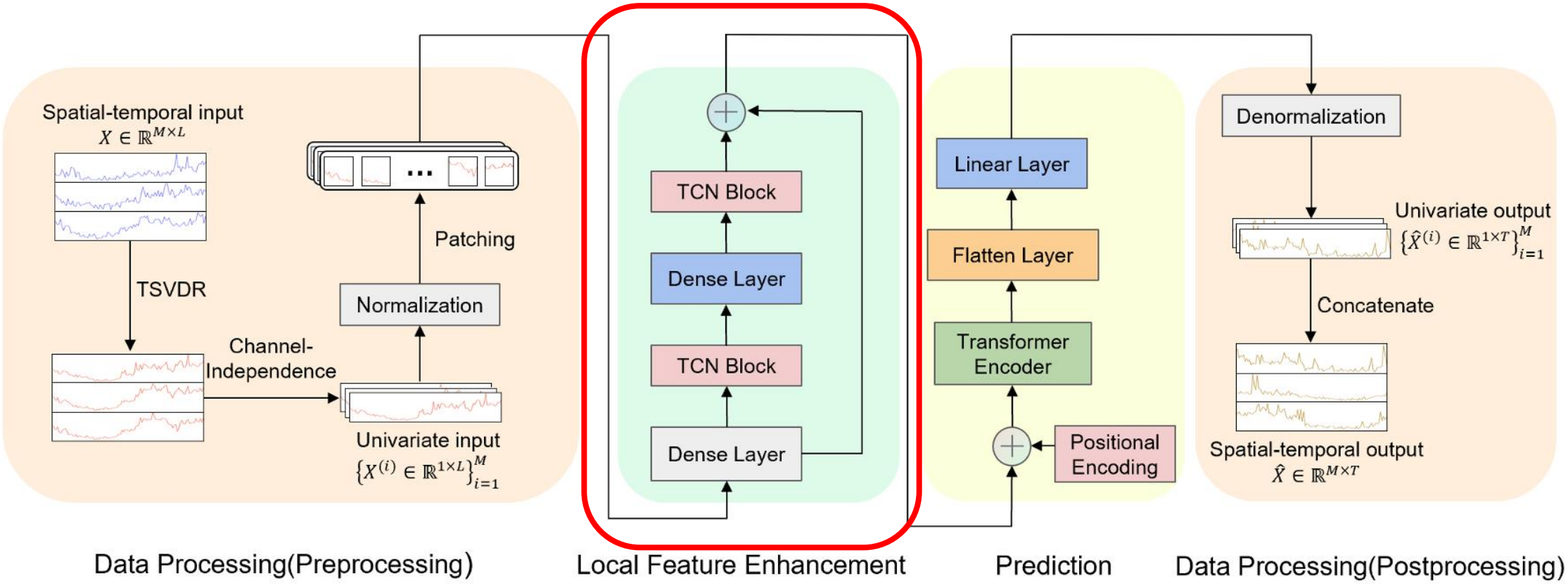
Result: $\left\{X_p^{(i)}[n] \in \mathbb{R}^{1 \times \ell}\right\}_{\substack{n=0,\ldots,N-1 \\ i=1,\ldots,M}}$, $\{\mu_i, \sigma_i\}_{i=1}^{M}$

$U\Sigma V^\top \leftarrow \mathrm{SVD}(X)$;
for $j \leftarrow 1$ ***to*** $\min(M,L)$ **do**
  if $\Sigma_{j,j} < c$ **then**
    $\Sigma_{j,j} \leftarrow 0$;
  end
end
$X' \leftarrow U\Sigma V^\top$;
for $i \leftarrow 1$ ***to*** M **do**
  $X^{(i)} \leftarrow X'[i,:]$;
  $\mu_i \leftarrow \frac{1}{L}\sum_{t=1}^{L} X^{(i)}(t)$;
  $\sigma_i \leftarrow \sqrt{\frac{1}{L}\sum_{t=1}^{L}\left(X^{(i)}(t) - \mu_i\right)^2}$;
  $X_{\mathrm{norm}}^{(i)} \leftarrow \frac{X^{(i)} - \mu_i}{\sigma_i}$;
  $X_{\mathrm{pad}}^{(i)} \leftarrow \left[X_{\mathrm{norm}}^{(i)}, 0, 0, \ldots, 0\right]$;
  $N \leftarrow \left\lfloor \frac{L-\ell}{S} \right\rfloor + 2$;
  for $n \leftarrow 0$ ***to*** $N-1$ **do**
    $X_p^{(i)}[n+1] \leftarrow X_{\mathrm{pad}}^{(i)}\left[n \cdot S + 1 : n \cdot S + \ell\right]$;
  end
end
return $\left\{X_p^{(i)}[n]\right\}_{\substack{n=0,\ldots,N-1 \\ i=1,\ldots,M}}$, $\{\mu_i, \sigma_i\}_{i=1}^{M}$;

DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

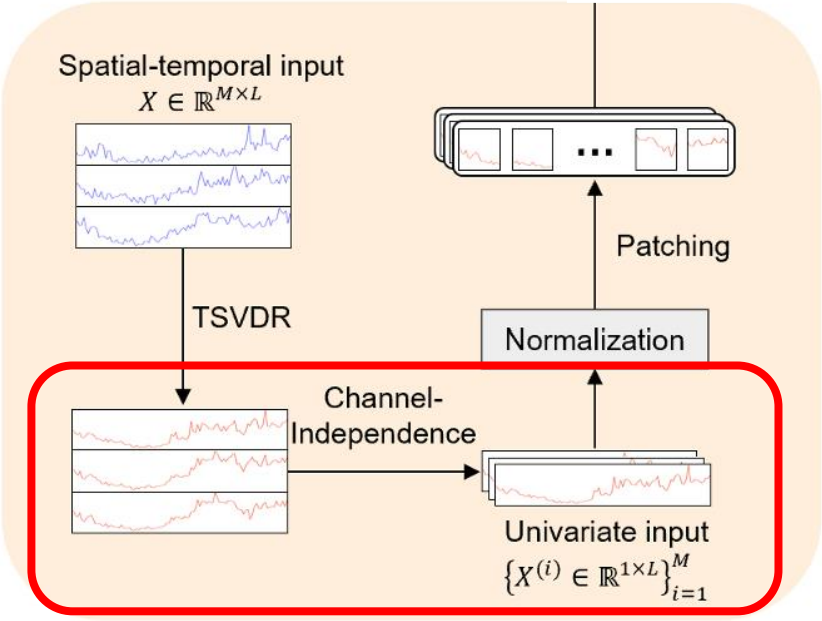
Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

P2. Weak local modeling in Transformer-based methods



Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

P3. Poor adaptability of graph-based methods to node changes



Data Processing(Preprocessing)

Algorithm 1: Data Processing Module (Preprocessing Part)

Data: $X \in \mathbb{R}^{M\times L}$, threshold $c \in \mathbb{R}_+$, segment length ℓ, stride S

Result: $\left\{X_p^{(i)}[n] \in \mathbb{R}^{1\times \ell}\right\}_{\substack{n=0,\dots,N-1\\ i=1,\dots,M}}$, $\{\mu_i, \sigma_i\}_{i=1}^{M}$

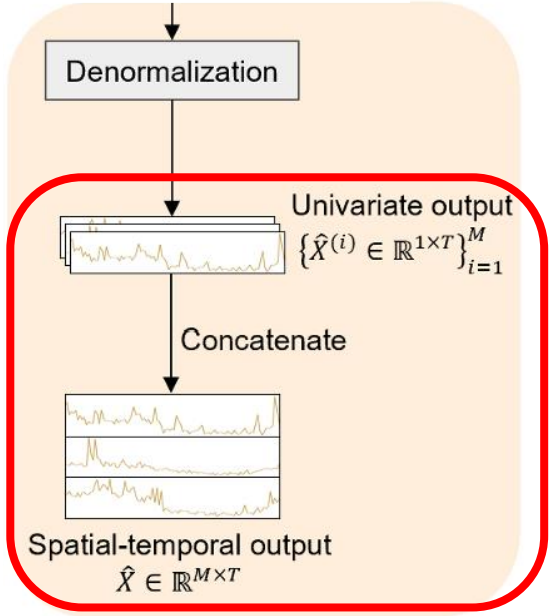
$U\Sigma V^\top \leftarrow \mathrm{SVD}(X)$;
for $j \leftarrow 1$ ***to*** $\min(M,L)$ **do**
  if $\Sigma_{j,j} < c$ **then**
    $\Sigma_{j,j} \leftarrow 0$;
  end
end
$X' \leftarrow U\Sigma V^\top$;
for $i \leftarrow 1$ ***to*** M **do**
  $X^{(i)} \leftarrow X'[i,:]$;
  $\mu_i \leftarrow \frac{1}{L}\sum_{t=1}^{L} X^{(i)}(t)$;
  $\sigma_i \leftarrow \sqrt{\frac{1}{L}\sum_{t=1}^{L}\left(X^{(i)}(t)-\mu_i\right)^2}$;
  $X_{\text{norm}}^{(i)} \leftarrow \frac{X^{(i)}-\mu_i}{\sigma_i}$;
  $X_{\text{pad}}^{(i)} \leftarrow \left[X_{\text{norm}}^{(i)}, 0, 0, \dots, 0\right]$;
  $N \leftarrow \left\lfloor \frac{L-\ell}{S} \right\rfloor + 2$;
  for $n \leftarrow 0$ ***to*** $N-1$ **do**
    $X_p^{(i)}[n+1] \leftarrow X_{\text{pad}}^{(i)}\left[n\cdot S+1 : n\cdot S+\ell\right]$;
  end
end
return $\left\{X_p^{(i)}[n]\right\}_{\substack{n=0,\dots,N-1\\ i=1,\dots,M}}$, $\{\mu_i, \sigma_i\}_{i=1}^{M}$;

P3. Poor adaptability of graph-based methods to node changes

Algorithm 2: Data Processing Module (Postprocessing Part)

Data: $\{X_{\text{linear}}^{(i)} \in \mathbb{R}^{1\times T}\}_{i=1}^{M}$, $\{\mu_i, \sigma_i\}_{i=1}^{M}$
Result: $\hat{X} \in \mathbb{R}^{M\times T}$

```
for i ← 1 to M do
    X̂^(i) ← X_linear^(i) · σ_i + μ_i;
end
X̂ ← 0^{M×T};
for i ← 1 to M do
    X̂[i,:] ← X̂^(i);
end
return X̂;
```



Data Processing(Postprocessing)

1. Background and Motivation
2. Overview of DP-LET Framework
3. **Experimental results**
4. Conclusion and Future work

Call Detail Records of Milan

Dimension	Data Size	Range	Minimum unit
Spatial	100×100	10,000 grids	Per-grid network traffic
Temporal	8,928	62 days	10 minutes

We randomly select 100 cells and conduct case studies with an input length $L = 432$ and output lengths $T = 72$ and $T = 144$.

Baseline methods

Transformer-based methods : PatchTST, iTransformer, Autoformer, FEDformer, Crossformer, Informer

Linear-based methods : DLinear

CNN-based methods : TimesNet

MLP-based methods : TimeMixer

Baseline methods

Transformer-based methods : PatchTST, iTransformer, Autoformer, FEDformer, Crossformer, Informer

Linear-based methods : DLinear

CNN-based methods : TimesNet

MLP-based methods : TimeMixer

Red : models relying on decomposition

Evaluation criteria

$$Mean\ Squared\ Error\ (MSE) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i\ )^2$$

$$Mean\ Absolute\ Error\ (MAE) = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i\ |$$

where N denotes the total number of samples, y_i represents the actual network traffic, $\hat{y}_i$ represents the predicted network traffic.

DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

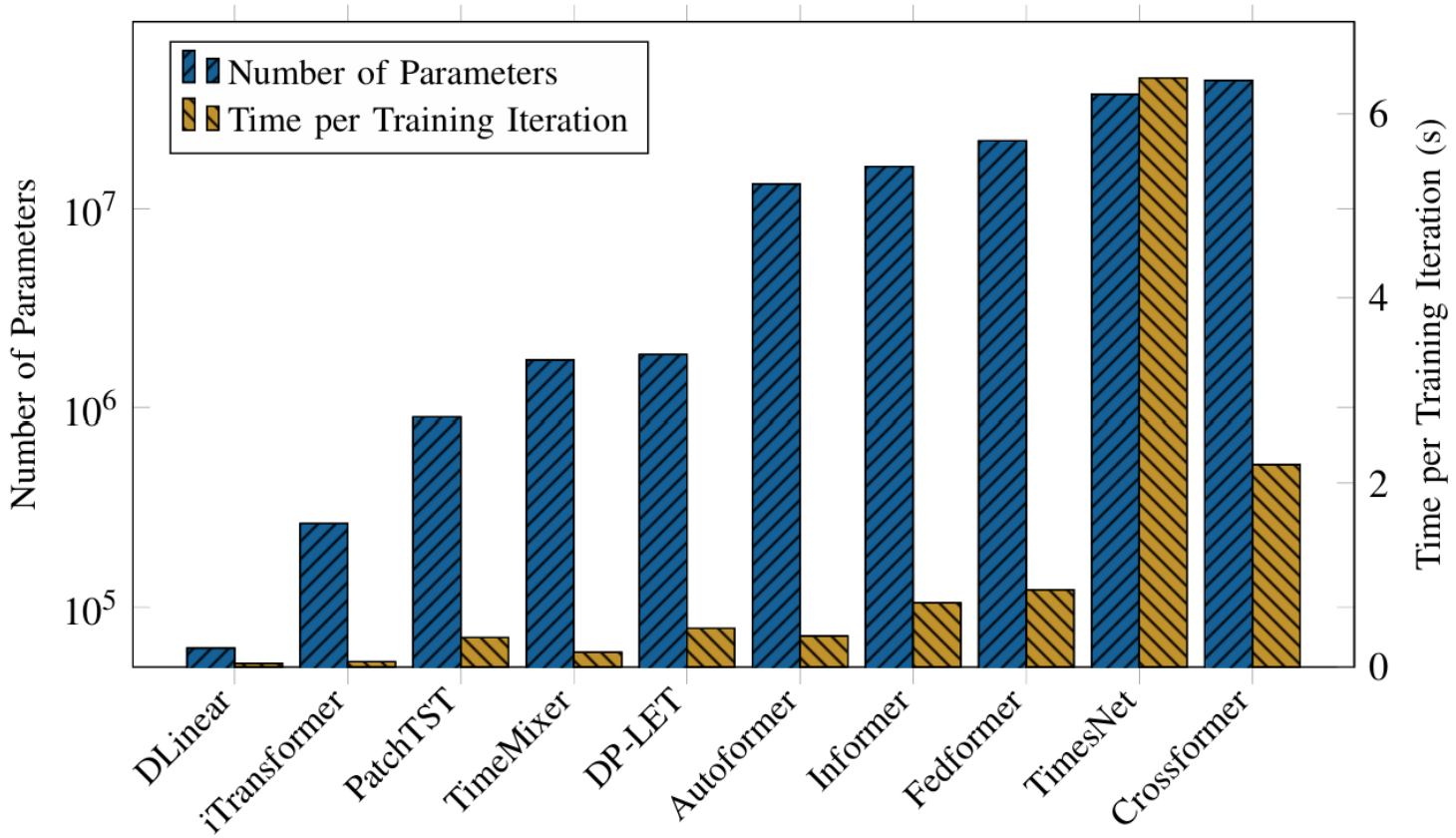
Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu

Methods	**DP-LET**		TimeMixer		PatchTST		iTransformer		TimesNet		DLinear		FEDformer		Autoformer		Crossformer		Informer	
Metrics	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
$T = 72$	**0.225**	**0.324**	0.230	0.329	0.234	0.337	0.237	0.333	0.304	0.385	0.311	0.428	0.387	0.468	0.429	0.506	0.510	0.517	0.653	0.615
$T = 144$	**0.238**	**0.333**	0.255	0.347	0.249	0.350	0.256	0.346	0.341	0.410	0.335	0.447	0.408	0.485	0.443	0.513	0.646	0.611	0.749	0.654

On average, DP-LET achieves a 31.8% reduction in MSE and a 23.1% reduction in MAE.

DP-LET: An Efficient Spatio-Temporal Network Traffic Prediction Framework

Xintong Wang, Haihan Nan, Ruidong Li and Huaming Wu



Methods	MSE	MAE
Prediction+Data Processing	0.254	0.349
Prediction+Local Feature Enhancement	0.343	0.423
Proposed Framework	**0.238**	**0.333**

Data Processing + Local Feature Enhancement + Prediction

The complete model achieves the highest overall performance.

Methods	Transformer		Non-stationary		FEDformer		Reformer	
Metrics	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Original	0.689	0.624	0.414	0.457	0.408	0.485	0.632	0.594
+ Local Feature Enhancement	0.606	0.588	0.373	0.437	0.388	0.471	0.620	0.586
Improvement (%)	12.05%	5.77%	9.90%	4.37%	4.90%	2.89%	1.90%	1.35%

Integrating the Local Feature Enhancement Module with other Transformer-based models also improves their performance, demonstrating its generality.

1. Background and Motivation
2. Overview of DP-LET Framework
3. Experimental results
4. **Conclusion and Future work**

1. DP-LET remains highly efficient while achieving, on average, a 31.8% reduction in MSE and a 23.1% reduction in MAE compared with the baselines.

2. In the future, we plan to explore more lightweight AI models suitable for edge deployment. We also aim to incorporate natural language information, enabling multimodal traffic prediction.