天津大学 Tianjin University

Discovering Drift Phenomena in Evolving Landscape (DELTA 2024)

CeDFormer: Community Enhanced Transformer for Dynamic Network Embedding

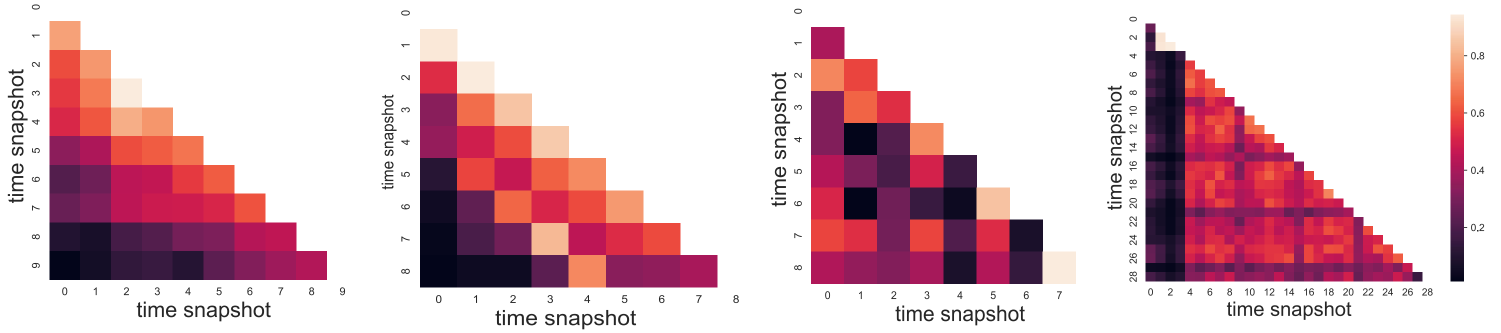
Jiaqi Guo, Tianpeng Li, Minglai Shao, Wenjun Wang, Lin Pan, Xue Chen, Yueheng Sun

August 26, 2024

Background

Dynamic Graph Data May Lack Strict Chain-Like Dependency

- complex, non-smooth, non-uniform, and highly frequent changes over multiple snapshots



Time Snapshot Dependency Analysis

Transformer-Based Models on Sequential Data

Advantages

Parallelization
Long-Range Dependencies
Flexibility
Interpretability

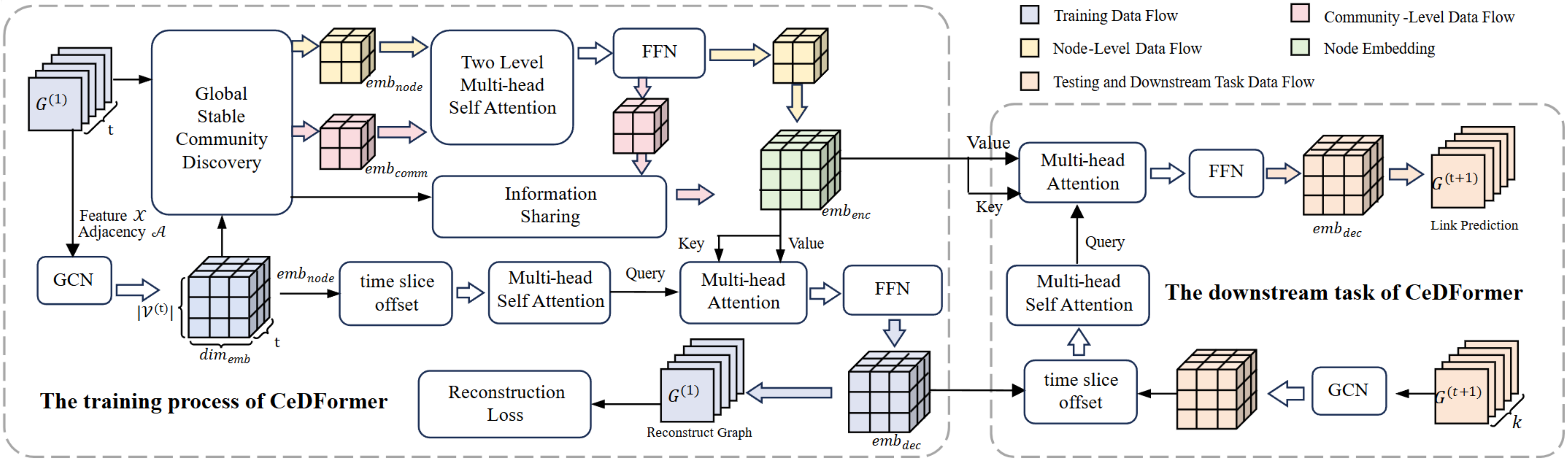
Disadvantages

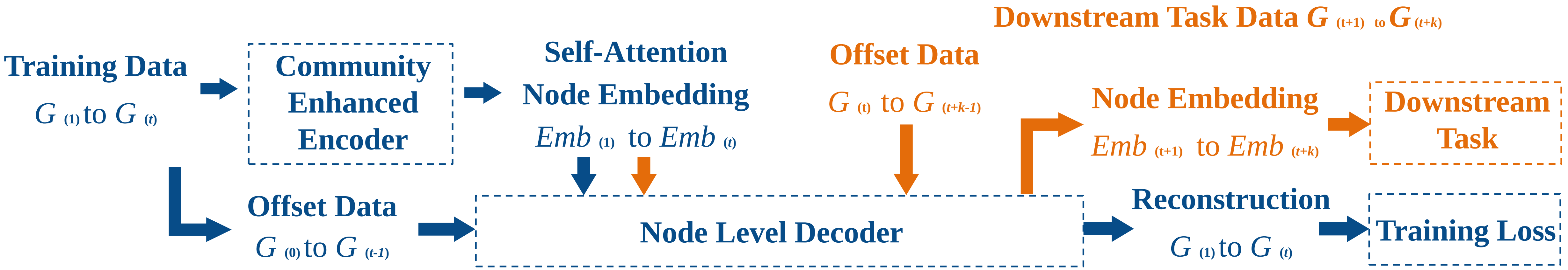
Computational Complexity
Training Costs

Using Transformer for Dynamic Network Embedding ➡ **CeDFormer** ⬅ **Optimization strategy using parameter sharing within the community**

Model Overview

Tianjin University





Community Enhanced Encoder

Tianjin University

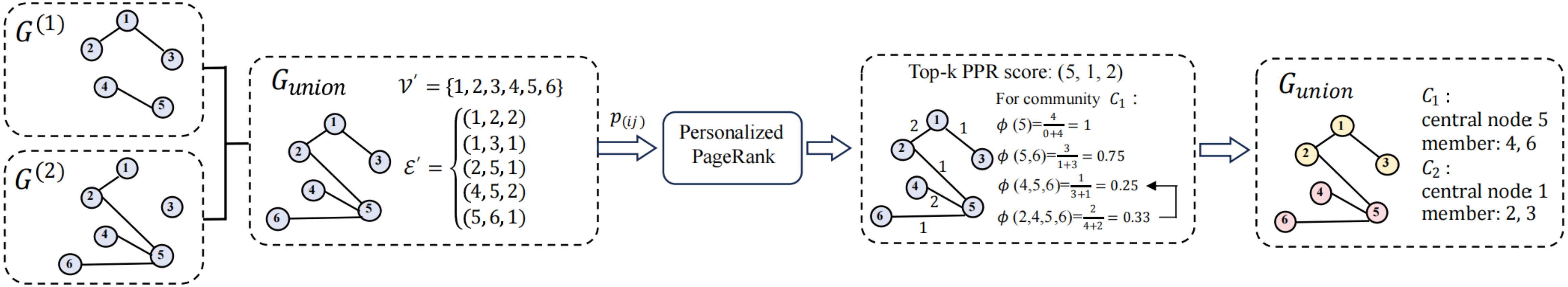
Global Stable Community Discovery:

- aggregate the dynamic graph $G = \{G_{(1)}, G_{(2)}, \ldots, G_{(T)}\}$ into a joint temporal graph G_{union}
- find the central points of community on G_{union} through Personalized PageRank (PPR)

$$p(ij) = \frac{\mathcal{W}'_{ij}}{\sum_{k \in \mathcal{N}(i)} \mathcal{W}'_{ik}}$$

- discover all members of various communities through conductance

$$\phi(C) = \frac{\sum_{i \in C, k \in \overline{C}, (i,k) \in \mathcal{E}'} \mathcal{W}'_{ik}}{\sum_{i,j \in C, (i,j) \in \mathcal{E}'} \mathcal{W}'_{ij} + \sum_{i \in C, k \in \overline{C}, (i,k) \in \mathcal{E}'} \mathcal{W}'_{ik}}$$



Community Enhanced Encoder

Embedding Strategy

$$\mathcal{F}_t = GCN(X^{(t)}, A^{(t)})$$

training data

$G_{(1)}$ to $G_{(t)}$

time-level node feature

$t \times n \times dim_{emb}$

$$\begin{cases} PE(t, 2i) = \sin\left(t/1000^{2i/dim_{emb}}\right) \\ PE(t, 2i+1) = \cos\left(t/1000^{2i/dim_{emb}}\right) \end{cases}$$

positional encoding

$t \times dim_{emb}$

Matrix transposition

embeddings of all nodes at all steps

$n \times t \times dim_{emb}$

Internal parameter sharing within the community

Global Stable Community

Two Level Multi-head Attention

$Emb_{\text{isolated nodes}}$ → $Rep_{\text{isolated nodes}}$

Emb_{comm} → Rep_{comm}

attention

If there are c communities in total, with an average of m members per community and a total of v nodes, the overall optimization rate for the entire encoder part can be calculated as $\frac{c*(m-1)}{v}$

Node Level Decoder

Tianjin University

BOS in dynamic graph

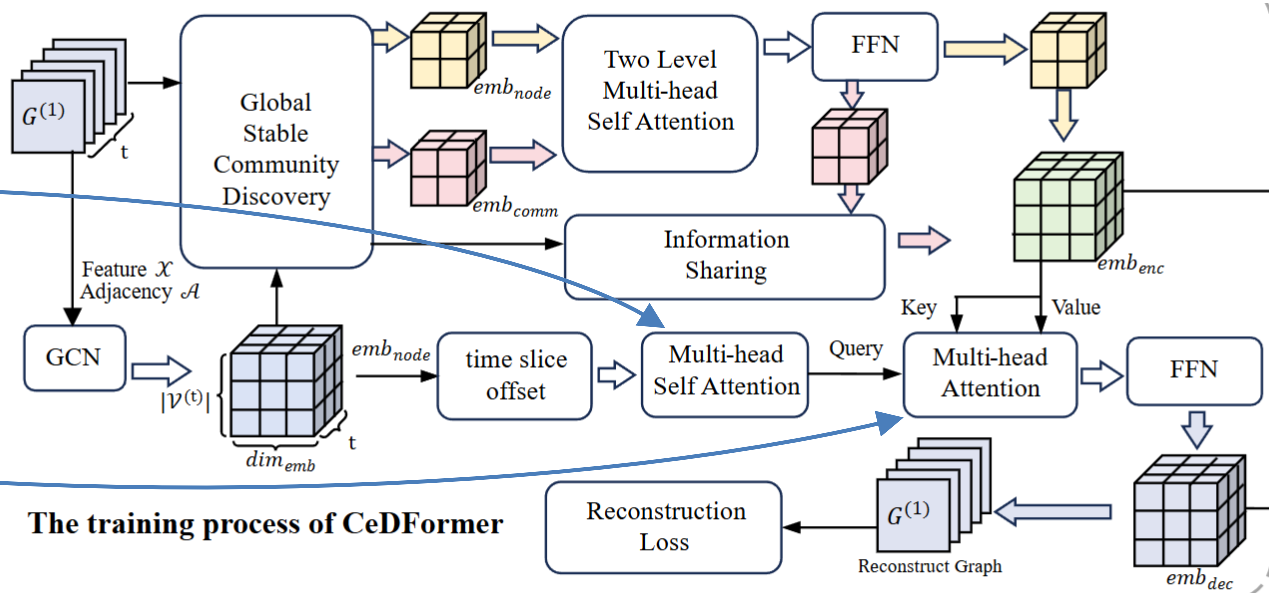
"Beginning of Sentence" (BOS) → initial graph embedding

NLP → **Dynamic Graph**

- unlike the boss in NLP, the initial graph embedding is not fixed but can be learned, similar to the concept in **meta learning**, where each training dataset will receive a unique initial graph embedding

Two Attention Layers

1. Masked self attention
2. Calculate attention with encoder

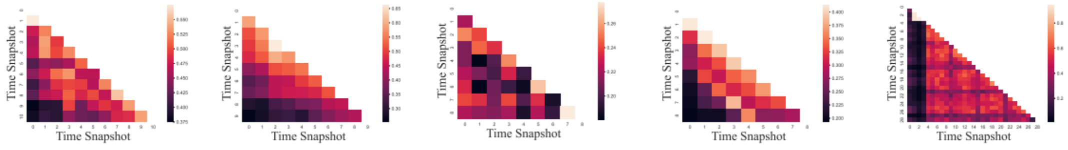


The training process of CeDFormer

Experiment

Time Snapshot Dependency Analysis

Time Snapshot Dependency Analysis



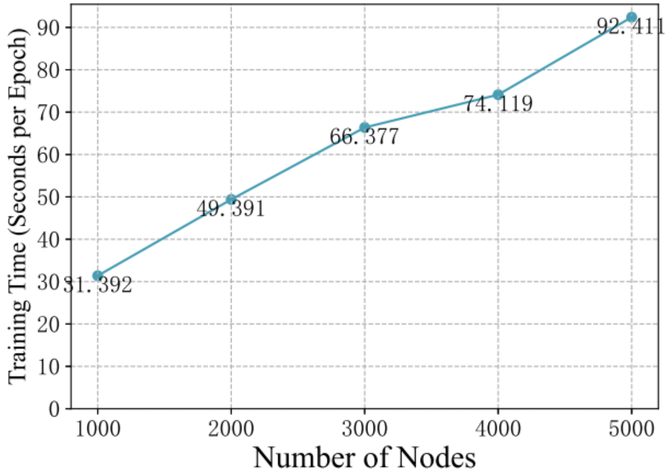
Enron DBLP highSchool FaceBook Email

Node Quantity Robustness Experiment

AUC and AP scores of link prediction on the BitcoinOTC Dataset.

Model	bitcoinotc		bitcoinotc-4k		bitcoinotc-3k		bitcoinotc-2k		bitcoinotc-1k	
	AUC	AP	AUC	AP	AUC	AP	AUC	AP	AUC	AP
VGRNN	81.34	88.62	82.02	89.31	81.19	89.05	76.95	85.99	63.07	71.40
DySAT	83.06	81.18	86.32	85.20	84.19	81.29	87.37	85.90	68.54	74.31
CeDFormer	89.23	88.63	93.48	94.05	93.22	95.03	88.42	85.21	73.01	75.20

The training time per epoch as a function of the number of nodes



Link Prediction

The average training time for one epoch on different datasets, in seconds/epoch.

Model	Enron	DBLP	highSchool	FaceBook	Email
DySAT	2.240	2.268	2.740	4.622	8.620
Dyformer	1.123	2.474	4.728	6.822	17.774
HGWaveNet	67.740	95.316	127.335	262.109	683.212
CeDFormer	2.267	3.319	2.945	5.972	25.719

AUC and AP scores of link prediction

Metrics	Model	Enron	DBLP	highSchool	FaceBook	Email	bitcoinOTC
AUC	VGRNN	93.42±0.70	85.80±0.78	89.66±0.32	89.79±0.34	91.92±1.18	81.34±3.25
	DySAT	88.81±1.10	86.74±1.51	91.51±0.48	88.88±0.89	90.42±0.91	83.06±2.71
	Dyformer	90.35±0.45	77.74±0.63	85.01±0.29	83.74±0.76	91.08±0.44	83.62±1.01
	DGCN	85.27±0.85	72.03±0.54	66.18±0.53	68.65±0.29	95.12±0.30	84.35±4.88
	HGWaveNet	94.45±0.26	**88.95±0.47**	91.64±0.22	86.98±0.66	92.49±0.36	80.06±1.27
	CeDformer	**94.67±0.51**	87.64±0.92	**94.33±0.80**	**90.05±0.50**	**95.83±1.01**	**89.23±1.97**
AP	VGRNN	**94.50±0.66**	88.64±0.58	88.71±0.65	89.14±0.41	93.34±0.70	88.62±2.37
	DySAT	87.30±1.54	89.64±1.00	89.17±1.12	88.61±0.96	89.19±1.17	81.18±2.59
	Dyformer	90.78±0.44	78.94±0.51	84.25±0.34	81.16±0.72	92.39±0.49	84.54±1.74
	DGCN	84.51±0.79	73.14±0.69	66.09±0.56	68.78±0.34	95.32±0.31	81.42±3.74
	HGWaveNet	94.37±0.32	**91.72±0.38**	90.86±0.25	85.96±0.71	94.02±0.32	81.15±1.24
	CeDformer	93.87±0.39	90.52±0.48	**94.10±0.83**	**89.84±0.56**	**96.65±0.42**	**88.63±2.70**

AUC and AP scores of new link prediction

Metrics	Model	Enron	DBLP	highSchool	FaceBook	Email	bitcoinOTC
AUC	VGRNN	87.41±0.96	75.87±1.65	88.09±0.28	86.76±0.54	90.37±1.28	80.22±2.78
	DySAT	82.58±3.00	76.57±2.30	90.42±0.78	85.97±1.21	81.71±3.97	83.09±2.61
	Dyformer	87.35±0.52	74.76±0.66	84.23±0.33	81.58±0.69	90.62±0.68	81.31±1.69
	DGCN	81.87±0.70	62.48±1.54	62.08±0.28	64.98±0.43	93.21±0.40	83.50±1.83
	HGWaveNet	89.38±0.36	**83.73±0.55**	89.63±0.24	83.97±0.61	92.05±0.38	79.97±0.99
	CeDformer	**90.46±0.59**	80.41±0.88	**91.35±0.78**	**88.37±0.46**	**94.36±0.49**	**88.72±1.55**
AP	VGRNN	88.76±0.70	78.38±1.21	86.81±0.54	85.30±0.66	92.22±0.77	84.85±3.45
	DySAT	83.07±2.89	78.81±2.13	88.76±1.56	84.78±1.79	74.96±3.13	82.24±2.94
	Dyformer	88.00±0.46	73.96±0.73	84.25±0.31	80.66±0.55	91.24±0.76	80.02±2.06
	DGCN	82.03±0.57	64.72±1.52	62.15±0.37	65.65±0.53	93.36±0.43	81.38±1.15
	HGWaveNet	87.71±0.33	**86.89±0.47**	88.48±0.27	82.30±0.67	93.29±0.38	79.86±1.25
	CeDformer	**89.74±0.71**	83.39±0.64	**90.09±0.72**	**87.91±0.45**	**95.75±0.53**	**87.11±2.01**

Thanks for your listening