Unifying Text Semantics and Graph Structures for Temporal Text-attributed Graphs with Large Language Models

NEURAL INFORMATION PROCESSING SYSTEMS 2025

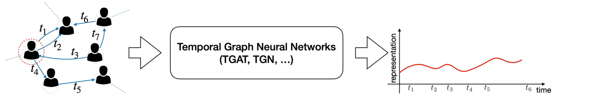
Siwei Zhang[1] · Yun Xiong[1*] · Yateng Tang[2] · Jiarong Xu[1] · Xi Chen[1] · Zehao Gu[1] · Xuehao Zheng[2] · Zi'an Jia[1] · Jiawei Zhang[3]
[1]Fudan University, [2]Tencent Weixin Group, [3]University of California, Davis

Paper Link

Motivation

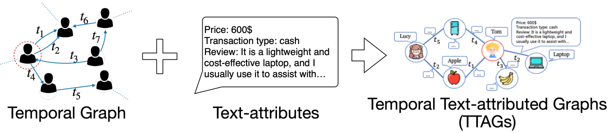
Temporal Graph Neural Networks (TGNNs)

- ❑ Learning time-aware node representations within temporal graphs



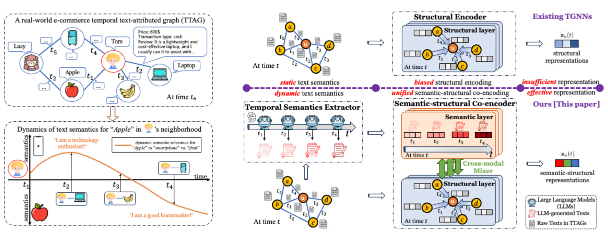
Temporal Text-attributed Graphs (TTAGs)

- ❑ Nodes and edges are associated with text attributes



Limitations of Existing TGNNs for TTAG Modeling

- ❑ **Neglect of semantic dynamics**: failing to capture the dynamic semantic shift of nodes within TTAGs
- ❑ **Ineffective encoding for semantic–structural reinforcement**: solely emphasizing graph structures and overly reliant on structural dynamics



(a) An illustrative example of semantic dynamics in TTAGs.

(b) Comparison between existing TGNNs and our framework.

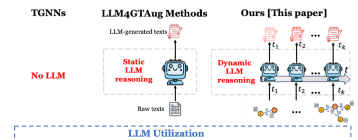
Limitations of LLM4GTAug Methods for TTAG Modeling

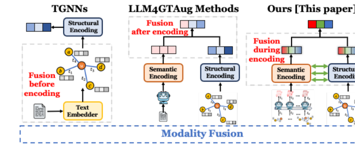
- ❑ **Inability of dynamic reasoning**: cannot provide semantic dynamics within the dynamic reasoning capability

Can we advance LLMs for TTAG modeling as a promising solution to the aforementioned limitations of existing TGNNs?

Key Innovation

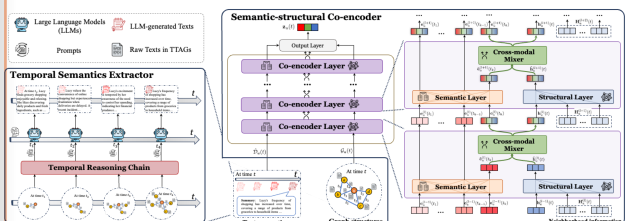
	TGNNs	LLM4GTAug	CROSS
Structural Dynamics	✗	✓	✓
Semantic Dynamics	✗	✗	✓
LLM Utilization	static	no LLM	**dynamic**
Modality Fusion	shallow at output (after encoding)	shallow at input (before encoding)	**hierarchical** (during encoding)





Framework

CROSS: a flexible LLM-based framework that unifies text semantics and graph structures for TTAG modeling



Temporal Semantics Extractor (for extracting semantic dynamics)

- ❑ **Temporal Reasoning Chain:** sampling reasoning timestamps for LLM reasoning $\hat{\mathcal{T}}_u = \left\{ \hat{t}_i \,\middle|\, \hat{t}_i = \begin{cases} t_{i \cdot \lceil \frac{n}{m} \rceil}, & i = 1, \ldots, m-1, \\ t_n, & i = m \end{cases} \in \mathcal{T}_u \right\}$
- ❑ **LLM Summarization:** extracting temporal semantics in natural language modality with LLMs

$$\hat{d}_u(\hat{t}) = \mathtt{LLM}\left(d_u, \hat{t}, \{r_*\}_{* \in \mathcal{H}_u(\hat{t})}; \textsc{Prompt}\right)$$

Semantic-structural Co-encoder (to ensure unimodal reinforcement)

- ❑ **Semantic Layer:** encoding semantic information with Transformer

$$\tilde{\mathbf{e}}_u^{(l)}(t_1), \ldots, \tilde{\mathbf{e}}_u^{(l)}(t_k) = \mathtt{TRM}^{(l)}\left(\mathbf{e}_u^{(l-1)}(t_1), \ldots, \mathbf{e}_u^{(l-1)}(t_k)\right)$$

- ❑ **Structural Layer:** aggregating structural information with TGNNs

$$\tilde{\mathbf{h}}_u^{(l)}(t) = \mathtt{MLP}^{(l)}\left(\mathbf{h}_u^{(l-1)}(t) \parallel \mathtt{AGG}\left(\mathbf{H}_u^{(l)}(t)\right)\right)$$

- ❑ **Cross-modal Mixer:** performing hierarchical (layer-wise) modality fusion throughout the whole encoding process

$$\mathbf{e}_u^{(l)}(t_k);\ \mathbf{h}_u^{(l)}(t) = \mathtt{Mixer}^{(l)}\left(\tilde{\mathbf{e}}_u^{(l)}(t_k) \parallel \tilde{\mathbf{h}}_u^{(l)}(t)\right)$$

Experiments

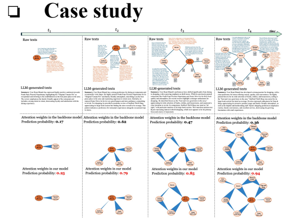
- ❑ **Realizing SOTA performance for temporal link prediction**

	Transductive setting					Inductive setting				
	Enron	GDELT	ICEWS1819	Googlemap_CT	Industrial	Enron	GDELT	ICEWS1819	Googlemap_CT	Industrial
JODIE	66.69 ± 2.0	48.81 ± 1.1	71.47 ± 4.0	56.72 ± 0.7	49.75 ± 1.2	53.41 ± 2.5	37.90 ± 2.4	57.75 ± 7.1	55.21 ± 1.0	30.38 ± 0.3
DyRep	58.85 ± 7.9	45.61 ± 2.8	63.13 ± 3.6	49.04 ± 2.2	36.49 ± 0.9	42.95 ± 8.3	42.23 ± 3.1	50.42 ± 2.8	47.44 ± 2.5	25.47 ± 1.8
TCL	71.16 ± 0.7	59.49 ± 0.5	87.58 ± 0.3	68.98 ± 0.4	50.87 ± 0.5	55.28 ± 1.5	47.13 ± 1.0	77.06 ± 0.2	66.26 ± 0.2	33.42 ± 0.5
CAWN	74.56 ± 0.6	57.00 ± 0.2	82.93 ± 0.1	65.34 ± 0.2	63.58 ± 0.8	61.58 ± 2.0	43.56 ± 0.6	70.65 ± 0.1	62.11 ± 0.2	53.42 ± 0.5
PINT	74.82 ± 2.8	52.71 ± 2.5	83.81 ± 0.9	72.94 ± 0.7	53.51 ± 0.6	56.38 ± 3.9	31.82 ± 4.1	63.16 ± 2.8	70.02 ± 0.6	39.72 ± 0.6
GraphMixer	62.68 ± 1.3	53.33 ± 0.4	80.69 ± 0.3	53.11 ± 0.2	50.50 ± 0.5	43.75 ± 1.5	41.18 ± 0.3	67.09 ± 0.5	51.36 ± 0.2	34.06 ± 0.7
FreeDyG	81.52 ± 1.8	68.27 ± 0.7	86.31 ± 0.6	78.82 ± 1.2	75.91 ± 0.7	70.38 ± 0.1	52.71 ± 0.3	74.16 ± 0.4	66.01 ± 2.8	56.48 ± 0.6
LKD4DyTAG	73.18 ± 0.3	57.28 ± 1.9	80.62 ± 4.2	77.11 ± 0.5	77.73 ± 0.7	67.45 ± 1.9	45.75 ± 2.0	73.81 ± 0.3	60.73 ± 1.0	57.91 ± 1.0
LLM_{zero}	24.18	7.99	33.68	30.30	11.27	17.73	10.08	32.26	38.21	2.62
LLM_{one}	46.27	28.91	50.82	48.79	30.29	48.14	28.91	44.69	43.83	20.28
TGAT	66.06 ± 0.1	56.73 ± .04	85.81 ± 0.2	63.13 ± 0.5	46.74 ± 3.9	47.80 ± 0.8	42.01 ± 0.5	74.10 ± 0.2	60.96 ± 0.2	30.04 ± 3.0
TGAT+	95.58 ± 0.7	**81.63 ± 1.7**	93.05 ± 1.6	99.91 ± 0.0	86.97 ± 2.8	81.52 ± 2.0	64.56 ± 1.8	82.25 ± 2.0	91.59 ± 0.0	62.22 ± 2.1
Avg. ↑ 26.59	↑ 29.52	**↑ 24.90**	↑ 7.24	↑ 36.78	↑ 40.23	↑ 33.72	↑ 22.55	↑ 8.15	↑ 30.63	↑ 32.18
TGN	73.05 ± 1.7	54.28 ± 1.6	84.79 ± 0.6	71.35 ± 0.5	54.46 ± 3.0	54.98 ± 2.3	37.48 ± 2.8	69.69 ± 0.8	67.88 ± 0.2	38.28 ± 4.1
TGN+	**95.84 ± 0.4**	77.95 ± 2.8	94.74 ± 5.7	**99.92 ± 0.0**	94.26 ± 0.8	82.38 ± 1.2	56.65 ± 3.8	84.01 ± 9.2	**92.68 ± 0.1**	**83.23 ± 2.6**
Avg. ↑ 25.45	**↑ 22.79**	↑ 23.67	↑ 9.95	**↑ 28.57**	↑ 39.80	↑ 27.40	↑ 19.17	↑ 14.32	**↑ 24.80**	**↑ 44.02**
DyGFormer	79.93 ± 0.1	61.35 ± 0.3	87.51 ± 0.3	54.82 ± 2.7	74.45 ± 0.7	66.86 ± 0.1	50.61 ± 0.2	78.14 ± 0.3	52.98 ± 2.5	54.20 ± 0.4
DyGFormer+	95.31 ± 2.8	81.28 ± 4.4	**95.77 ± 0.3**	99.82 ± 0.0	**94.78 ± 1.4**	**86.01 ± 4.9**	**66.37 ± 4.4**	**87.80 ± 0.6**	91.74 ± 0.1	82.30 ± 1.2
Avg. ↑ 22.03	↑ 15.38	↑ 19.93	**↑ 8.26**	↑ 44.99	**↑ 20.33**	**↑ 19.15**	**↑ 15.76**	**↑ 9.66**	↑ 38.76	↑ 28.10

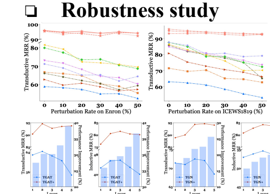
- ❑ **Ablation study for the texts and encoding mechanisms**

	Datasets	Methods	Semantic Encoding			Structural Encoding			Semantic-structural Co-encoding		
			$Text_{raw}$	$Text_{LLM}$	*imprv.*	$Text_{raw}$	$Text_{LLM}$	*imprv.*	$Text_{raw}$	$Text_{LLM}$ (ours)	*imprv.*
Transductive	Enron	TGAT	49.73 ± 0.8	63.07 ± 0.5	↑ 13.34	66.06 ± 0.1	63.65 ± 1.7	↓ 2.41	70.27 ± 0.2	**95.58 ± 0.7**	↑ 25.31
		TGN	49.73 ± 0.8	63.07 ± 0.5	↑ 13.34	73.05 ± 1.7	72.36 ± 4.0	↓ 0.69	74.28 ± 0.9	**95.84 ± 0.4**	↑ 21.56
		DyGFormer	49.73 ± 0.8	63.07 ± 0.5	↑ 13.34	79.93 ± 0.1	80.46 ± 0.5	↑ 0.53	80.91 ± 0.1	**95.31 ± 2.8**	↑ 14.40
	ICEWS1819	TGAT	77.45 ± 0.5	85.04 ± 1.5	↑ 7.59	85.81 ± 0.2	86.12 ± 0.1	↑ 0.31	87.33 ± 1.0	**93.05 ± 1.6**	↑ 5.72
		TGN	77.45 ± 0.5	85.04 ± 1.5	↑ 7.59	84.79 ± 0.6	85.69 ± 0.4	↑ 0.90	85.96 ± 0.8	**94.74 ± 5.7**	↑ 8.78
		DyGFormer	77.45 ± 0.5	85.04 ± 1.5	↑ 7.59	87.51 ± 0.3	88.11 ± 0.5	↑ 0.60	86.72 ± 0.4	**95.77 ± 0.3**	↑ 9.05
Inductive	Enron	TGAT	31.94 ± 0.7	45.24 ± 1.1	↑ 13.30	47.80 ± 0.8	45.01 ± 1.3	↓ 2.79	53.26 ± 1.0	**81.52 ± 2.0**	↑ 28.26
		TGN	31.94 ± 0.7	45.24 ± 1.1	↑ 13.30	54.98 ± 2.3	53.93 ± 4.0	↓ 1.05	58.92 ± 1.4	**82.38 ± 1.2**	↑ 23.46
		DyGFormer	31.94 ± 0.7	45.24 ± 1.1	↑ 13.30	66.86 ± 0.1	67.64 ± 1.4	↑ 0.78	68.27 ± 0.1	**86.01 ± 4.9**	↑ 17.74
	ICEWS1819	TGAT	60.63 ± 0.8	71.45 ± 0.6	↑ 10.82	74.10 ± 0.2	74.12 ± 0.2	↑ 0.02	75.19 ± 0.2	**82.25 ± 2.0**	↑ 7.06
		TGN	60.63 ± 0.8	71.45 ± 0.6	↑ 10.82	69.69 ± 0.8	70.39 ± 1.2	↑ 0.70	70.01 ± 0.6	**84.01 ± 9.2**	↑ 13.99
		DyGFormer	60.63 ± 0.8	71.45 ± 0.6	↑ 10.82	78.14 ± 0.3	77.70 ± 0.8	↓ 0.44	79.79 ± 0.4	**87.80 ± 0.6**	↑ 8.01

- ❑ **Case study**



- ❑ **Robustness study**



(a) TGAT as the backbone. (b) TGN as the backbone.

Please refer to our paper for more interesting experimental results