



LLM4DyG: Can Large Language Models Solve Spatial-Temporal Problems on Dynamic Graphs

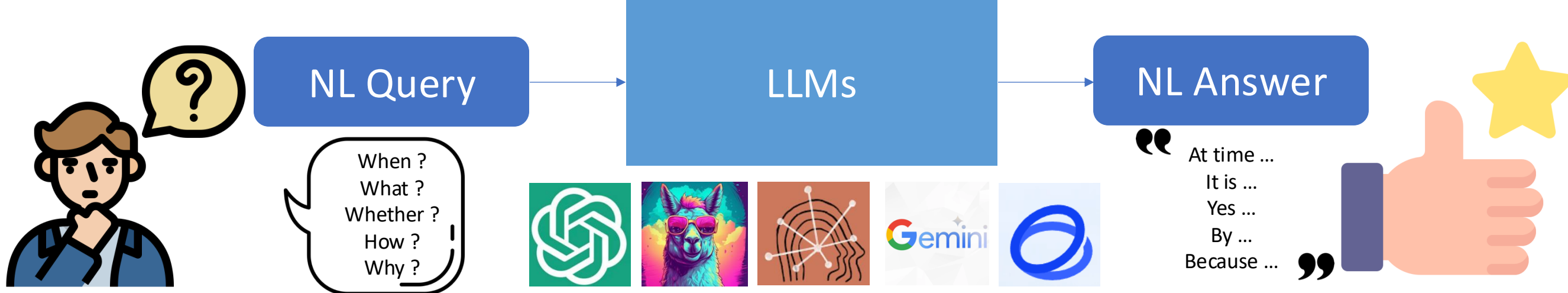
Zeyang Zhang¹, Xin Wang^{1,2}, Ziwei Zhang¹, Haoyang Li¹, Yijian Qin¹, Wenwu Zhu^{1,2}

Department of Computer Science and Technology, Tsinghua University¹ BNRist²

1

Background & Motivation

① In the era of LLMs, we handle data with natural language (NL)

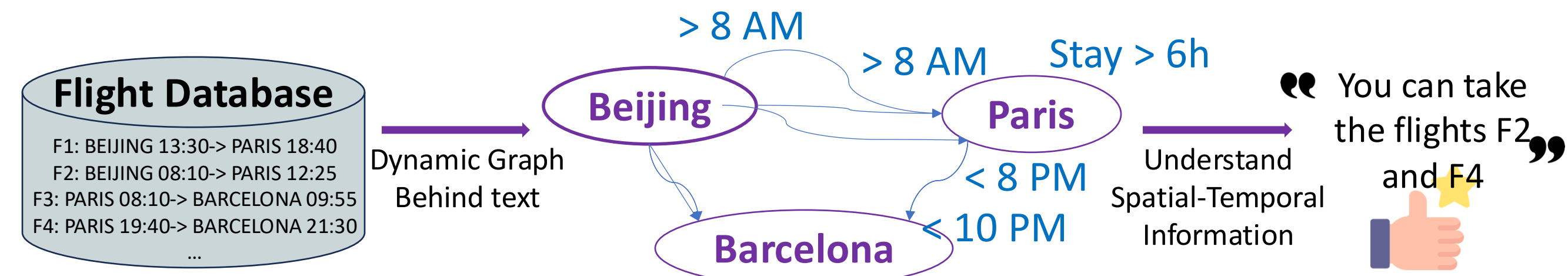


② NL Queries Can Include Implicit Dynamic Graphs (Spatio-Temporal Relations)

For example, ...

I'd like to fly from **Beijing** to **Barcelona**, with a departure time **no earlier than 8:00 AM** and an arrival time **no later than 10:00 PM**. Additionally, I'd like to have a **half-day layover** in **Paris** during the **daytime** to visit friends. Could you please suggest suitable flights?

③ LLMs have to understand spatial-temporal information to give right answers



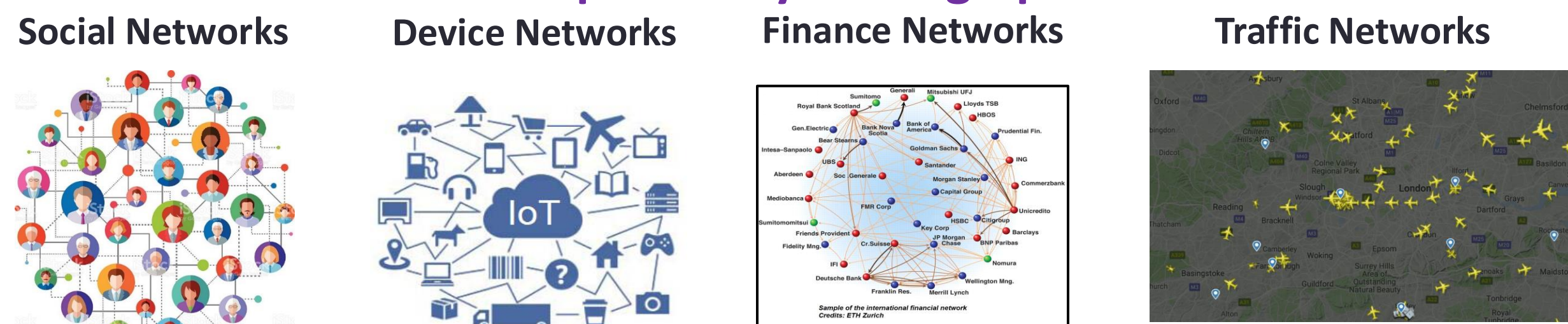
④ But, can LLMs understand the **spatial-temporal information** on dynamic graphs in natural language?

Remained Unexplored !

2

Problem & Challenge

① Existing works of LLMs for static graphs: **ignore the rich temporal information in ubiquitous dynamic graphs**



② Existing works for dynamic GNNs: **unable to tackle tasks with a unified model**



③ **The first study** to evaluate LLMs' spatial-temporal understanding abilities on dynamic graphs



④ However, we face several challenges

- How to assess **temporal and structural information** separately and simultaneously?
- How to investigate the **complex and mixed interactions** of spatial and temporal dimensions?
- How to design **prompts** to consider spatial-temporal information in natural language?

⑤ Our main contribution

- We propose LLM4DyG: **nine** specially designed tasks for LLMs with **controllable environments and data generation**.
- We **analyze the impacts of various factors**: including different data generators, data statistics, prompting techniques, and LLMs on the model performance.
- We propose **Disentangled Spatial-Temporal Thoughts (DST2)** prompt method for to enhance LLMs' spatial-temporal understanding abilities.

5

Contact Information

Zeyang Zhang,
Final-year Ph.D. in Tsinghua University
zy-zhang20@mails.tsinghua.edu.cn
<https://zzythu.com>

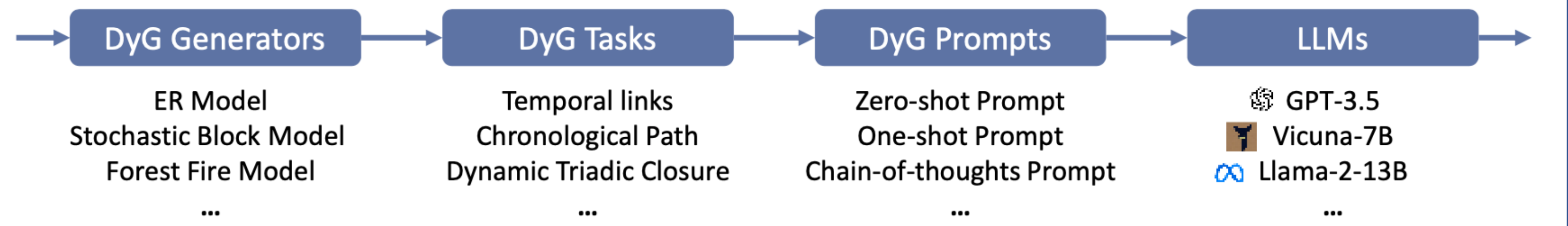
Code (Easy to Use!): <https://github.com/wondergo2017/LLM4DyG>



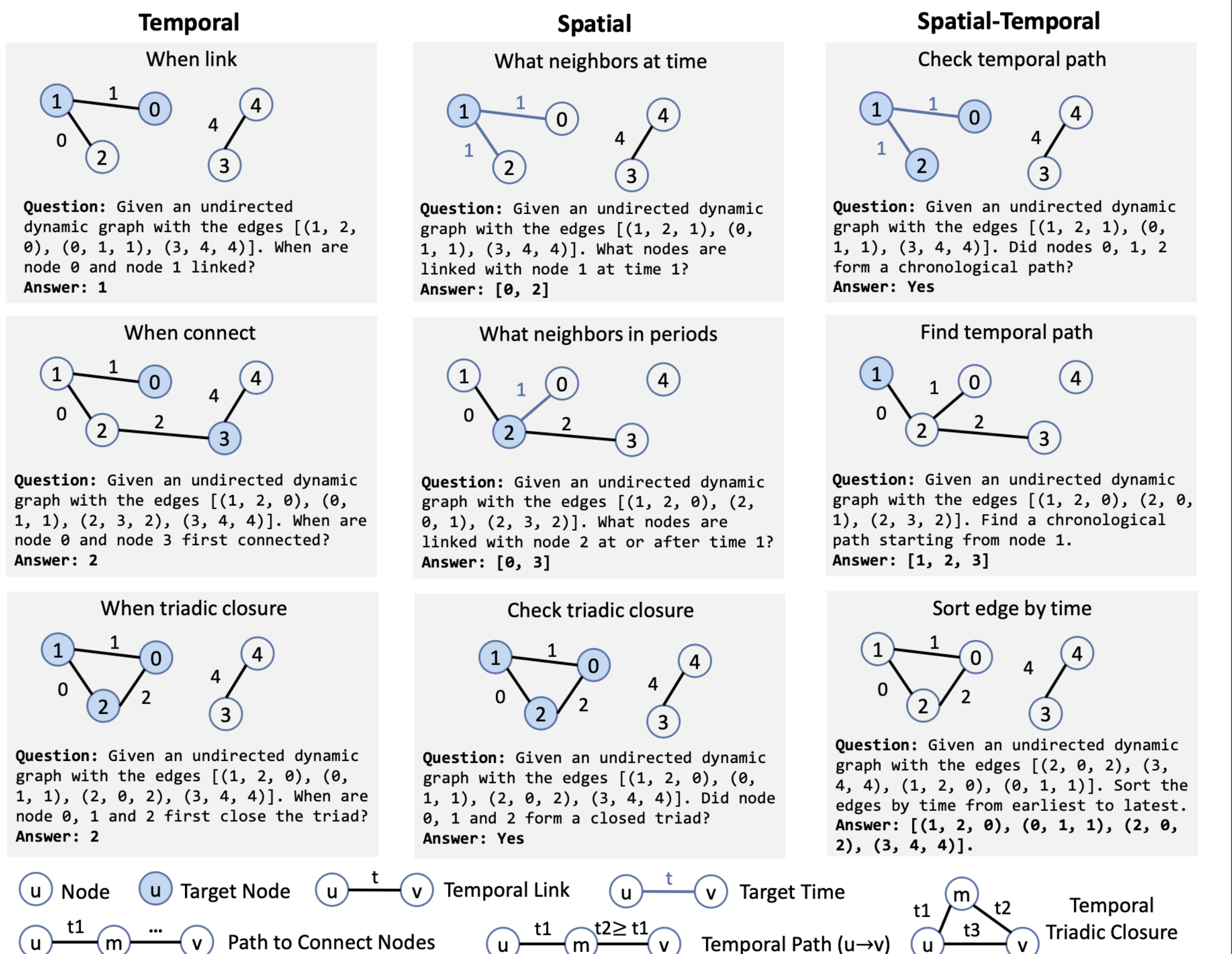
3

Framework: LLM4DyG

① LLM4DyG Benchmark & Pipeline with controlled environments and data



② Designed Tasks: **When, What, Whether** spatial-temporal patterns take place



③ Disentangled Spatial-Temporal Thoughts (DST2)

Task	Temporal			Spatial		Spatial-Temporal			
	when link	when connect	when closure	neighbor at time	neighbor in periods	check closure	check tpath	find tpath	sort edge
one-shot prompt	33.7±2.1	77.0±2.9	73.0±1.6	34.0±1.4	15.7±4.2	66.7±4.5	63.7±2.6	78.3±6.0	29.3±4.0
v1: Think (about) nodes and then time	40.0±1.6	77.0±4.1	74.0±1.4	34.0±0.8	15.0±4.2	69.3±1.7	61.0±3.3	79.0±7.5	30.0±3.6
v2: Think (about) time and then nodes	37.3±2.6	76.7±3.4	73.3±0.5	31.7±1.9	15.7±3.4	67.0±2.9	61.3±1.9	79.0±7.5	30.7±3.9
v3: Pick nodes and then time	59.3±2.1	77.0±2.4	68.0±0.8	35.0±2.9	16.7±4.7	65.0±3.7	62.3±2.9	78.0±5.4	30.0±2.9
v4: Pick time and then nodes	76.7±1.7	76.3±3.9	68.7±0.9	35.7±2.5	15.3±3.3	65.3±2.9	63.3±2.6	78.3±5.8	29.3±2.9

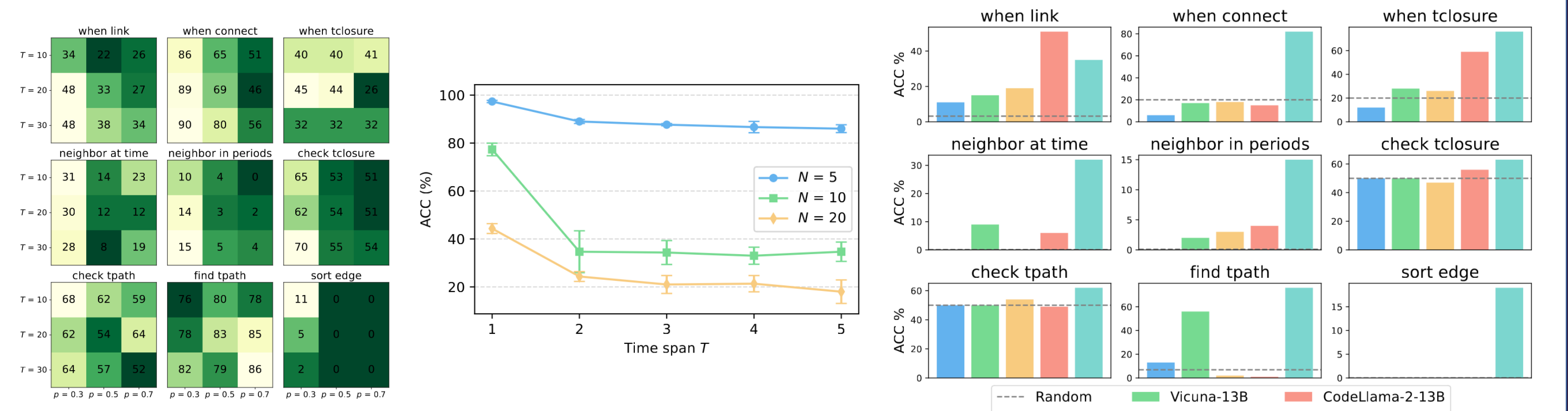
4

Results and Analyses

Task	Data	model	Temporal			Spatial		Spatial-Temporal		
			when link	when connect	when closure	neighbor at time	neighbor in periods	check closure	check tpath	find tpath
N = 5	Random	GPT-3.5	68.0±2.8	97.7±0.9	52.7±2.4	86.0±2.2	42.3±1.7	69.0±2.2	58.7±2.1	79.0±4.1
		Random	3.2	20.0	20.0	3.2	50.0	50.0	50.0	9.3
		Δ	+64.8	+77.7	+32.7	+82.8	+39.1	+19.0	+8.7	+69.7
N = 10	Random	GPT-3.5	33.7±2.1	77.0±2.9	73.0±1.6	34.0±1.4	15.7±4.2	66.7±4.5	63.7±2.6	78.3±6.0
		Random	3.2	20.0	20.0	0.1	50.0	50.0	50.0	6.7
		Δ	+30.4	+57.0	+53.0	+33.9	+15.6	+16.7	+13.7	+71.6
N = 20	Random	GPT-3.5	40.3±1.7	17.7±4.2	63.3±0.9	17.7±1.7	2.0±0.8	64.3±7.3	57.0±2.2	85.0±0.8
		Random	3.2	20.0	20.0	0.0	50.0	50.0	50.0	7.3
		Δ	+37.1	-2.3	+43.3	+17.7	+2.0	+14.3	+7.0	+77.7
Avg.	Random	GPT-3.5	47.3±1.2	64.1±0.3	63.0±1.0	45.9±3.1	20.0±0.8	66.7±2.9	59.8±0.8	80.8±0.3
		Random	3.2	20.0	20.0	1.1	50.0	50.0	50.0	7.8
		Δ	+44.1	+44.1	+43.0	+44.8	+18.9	+16.7	+9.8	+73.0

Obs.1. LLMs show preliminary spatial-temporal abilities

Obs.2. Tasks are increasing difficulty for LLMs as the graph size grows.



Obs.3. The difficulties are not sensitive to the time span but sensitive to the graph density

Obs.4. Temporal information adds additional difficulties compared to static graphs.

Task	Prompt Method	Temporal			Spatial		Spatial-Temporal		
		when link	when connect	when closure	neighbor at time	neighbor in periods	check closure	check tpath	find tpath
zero-shot	Random	2.3±0.5	73.1±0.1	68.0±0.8	36.0±0.3	4.3±0.1	70.7±1.7	66.0±0.4	56.3±0.9
	LLM4DyG	33.7±2.1	77.0±2.9	73.0±1.6	34.0±1.4	15.7±4.2	66.7±4.5	63.7±2.6	78.3±6.0
one-shot	Random	1.0±0.8	58.3±1.2	70.0±1.6	32.0±0.8	4.3±0.1	55.0±1.4	62.3±2.9	58.0±1.1
	LLM4DyG	10.3±0.1	76.0±0.4	80.0±1.6	27.7±1.9	13.0±0.5	57.7±2.1	57.7±2.1	81.3±2.6

Obs.6. Training on codes usually helps

Generation Model	ER Model	SB Model	FF Model
zero-shot	2.3±0.5	7.7±1.7	5.3±2.5
one-shot	33.7±2.1	46.0±2.9	48.0±7.1
zero-shot COT	1.0±0.8	5.7±3.1	2.0±1.6
one-shot COT	10.3±0.1	15.3±0.9	13.0±2.9

Obs.5. General prompting techniques do not consistently help