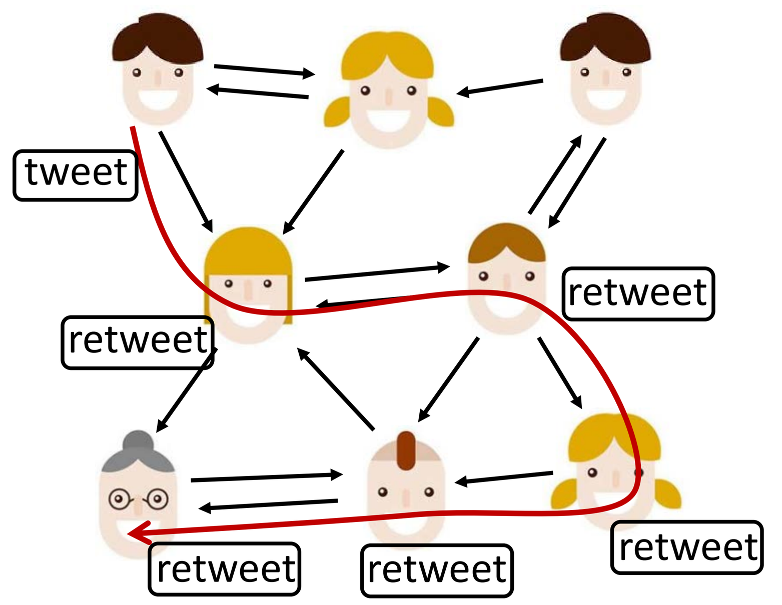
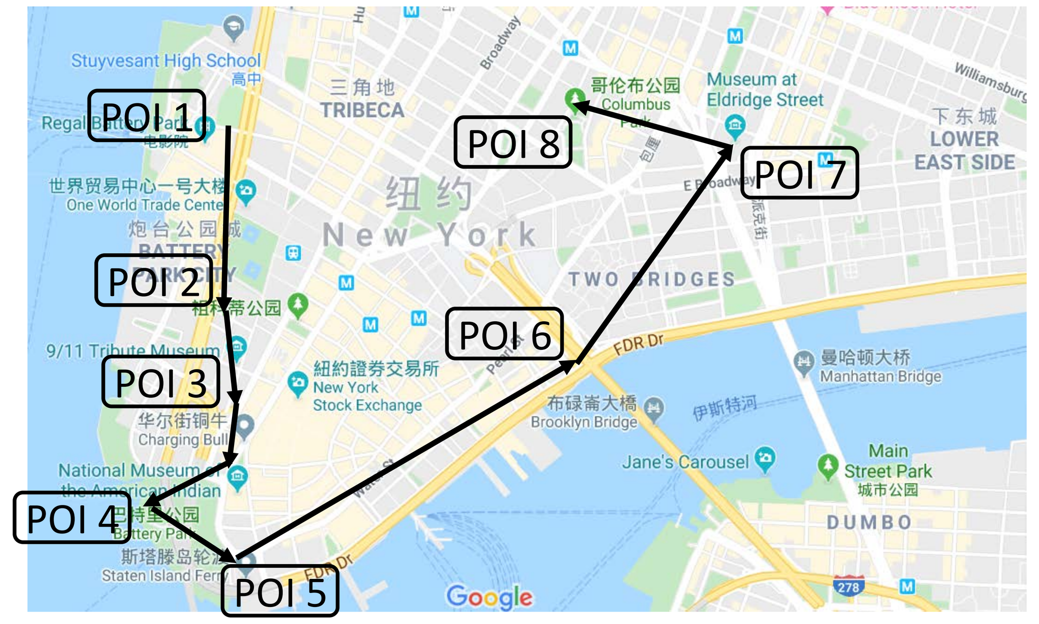
Learning Latent Process from High-Dimensional Event Sequences via Efficient Sampling

Qitian Wu, Zixuan Zhang, Xiaofeng Gao, Junchi Yan, Guihai Chen

Department of Computer Science and Engineering, Shanghai Jiao Tong University, China

Motivation & Background





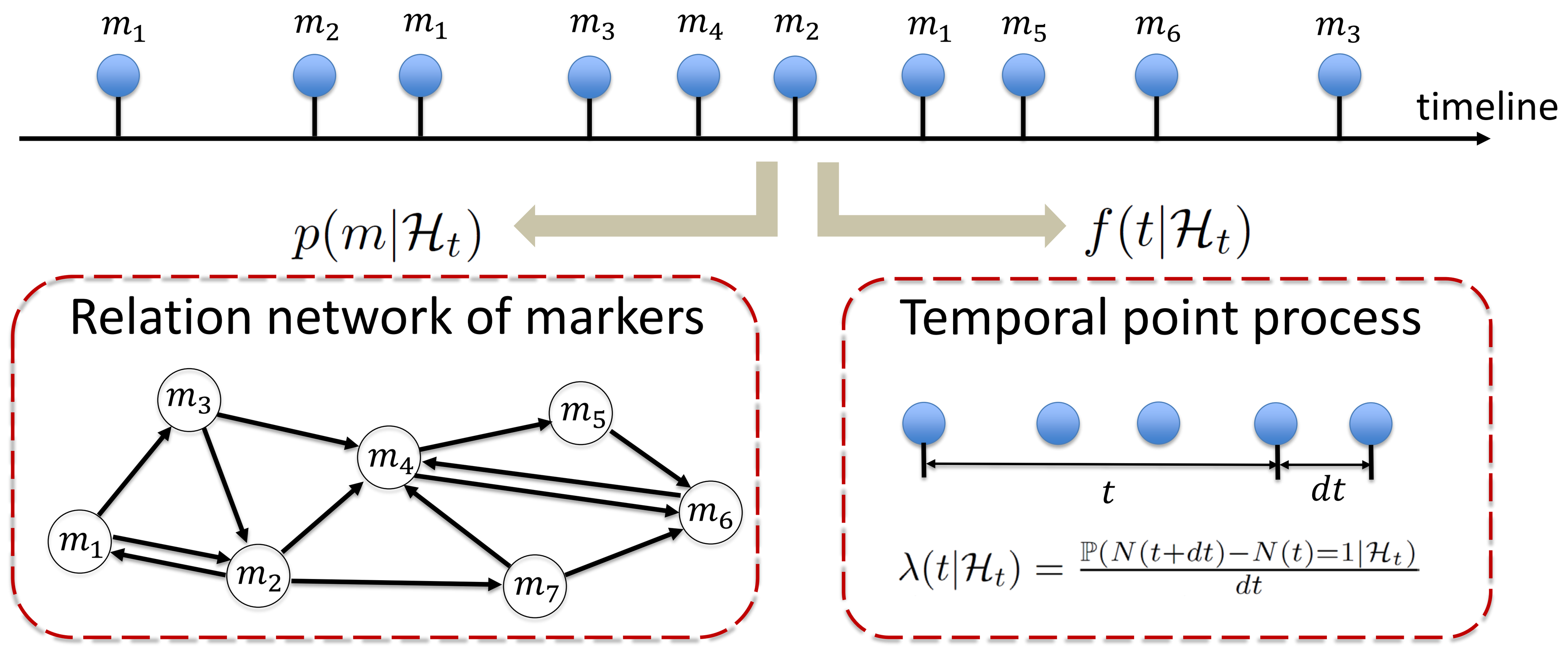
High-dimensional event sequences are ubiquitous:

- Information cascade in large-scale social networks
- Point-of-interest visiting route in a large city
- Markers contain plenty of combinational features

Main challenges for high-dimensional event sequence modeling:

- Unknown networks among high-dimension markers
- Multiply subsequences of interdependent events
- Hard to measure the discrepancy

Problem Formulation



Methodology

- hidden relation network ➡ graph attention network

$$p(m_j \in \mathcal{N}_i) = \frac{\exp(\mathbf{w}_C^\top[\mathbf{d}_j||\mathbf{d}_i])}{\sum_{u=1}^{M}\exp(\mathbf{w}_C^\top[\mathbf{d}_u||\mathbf{d}_i])}$$

- temporal point process ➡ attentive intensity model

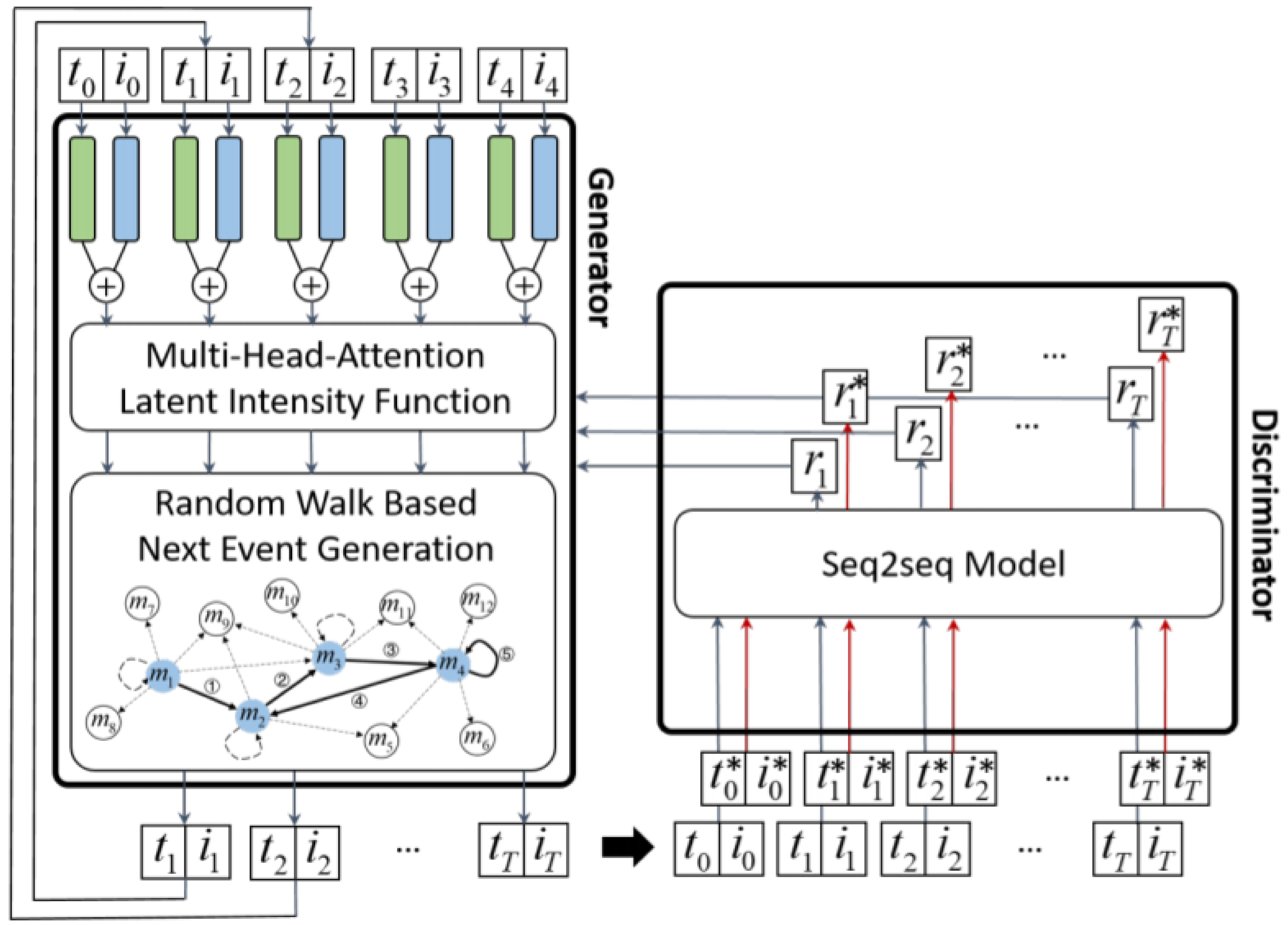
$$\mathbf{h}_n = MultiHeadAttn(\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_k), n = 0, 1, \cdots, k,$$

- Random walk approach for marker generation

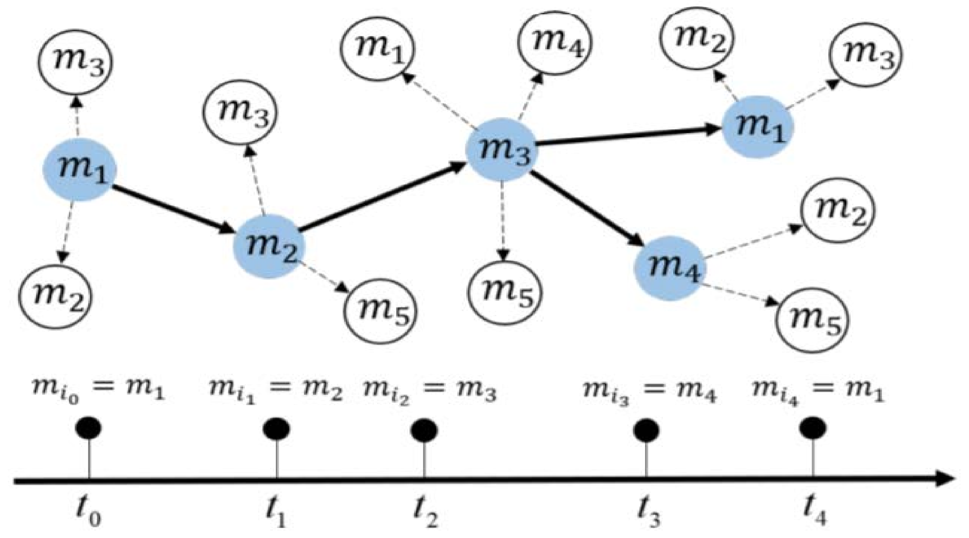
$$p(m_j \in \overline{\mathcal{N}}_{i_n} | m_j \in \mathcal{N}_{i_n}) = \frac{\exp(\mathbf{w}_N^\top[\mathbf{h}_n||\mathbf{d}_j] + b_N)}{\sum_{m_u \in \mathcal{N}_{i_n}}\exp(\mathbf{w}_N^\top[\mathbf{h}_n||\mathbf{d}_u] + b_N)}$$

- Adversarial generative imitation learning

$$\min_{\pi} -H(\pi) + \max_{r} \mathbb{E}_{\pi_E}(r(\mathcal{S}^*)) - \mathbb{E}_{\pi}(r(\mathcal{S}))$$



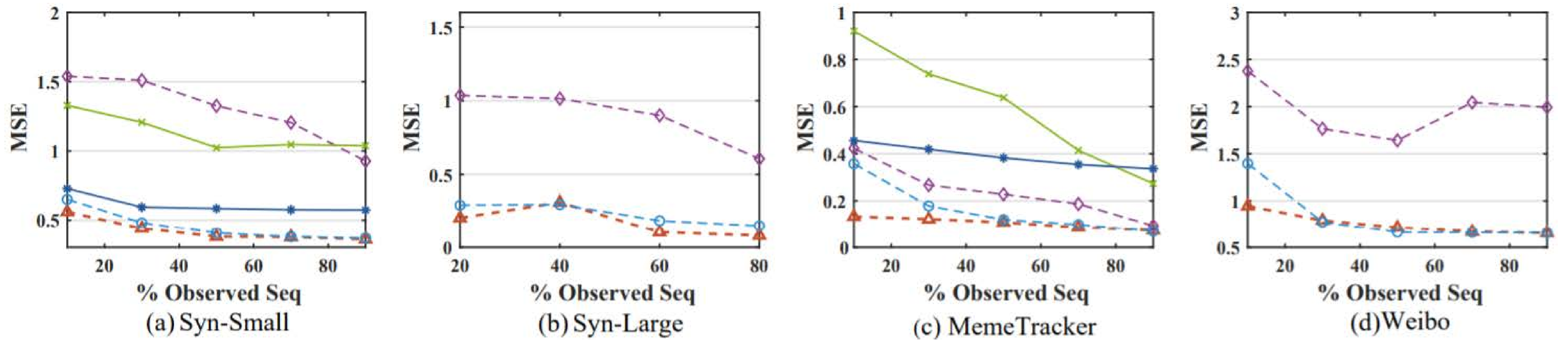
- Sampling for the marker and time of new event



estimate descendants → update intensity → draw new marker → estimate arrival time

Experiments

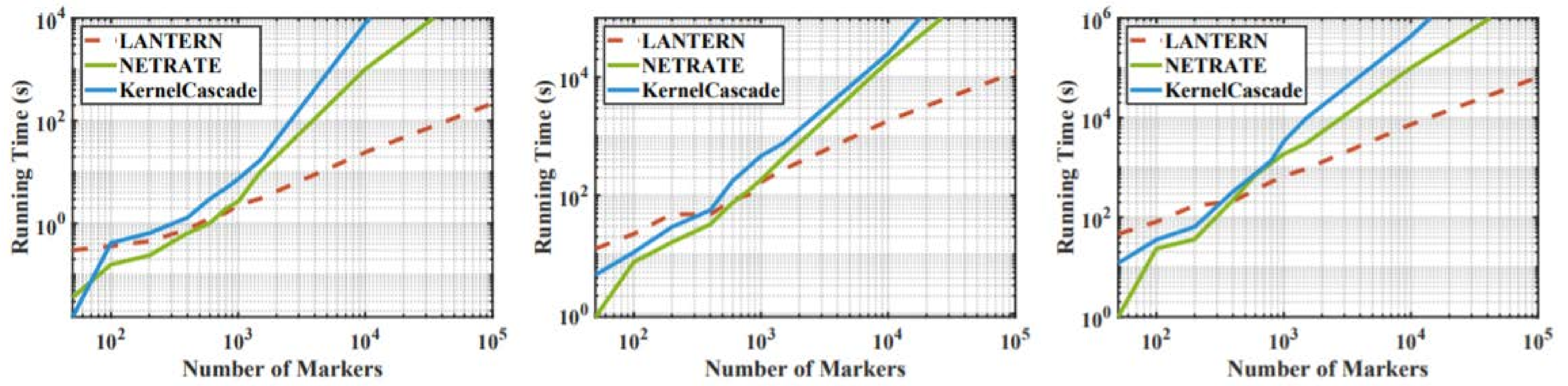
Prediction of next event's time and marker



Reconstruction of marker relation network

Methods	Syn-Small			Syn-Large			MemeTracker			Weibo		
	PRE	REC	F1	PRE	REC	F1	PRE	REC	F1	PRE	REC	F1
NETRATE	0.4983	0.3986	0.4429	-	-	-	0.5638	0.4510	0.5011	-	-	-
KernelCascade	0.4975	0.3980	0.4422	-	-	-	0.5560	0.4448	0.4942	-	-	-
LTN-PR (K_1)	0.6486	0.3892	0.4865	0.5573	0.3344	0.4180	0.5200	0.3120	0.3899	0.3628	0.2984	0.3275
LTN-PR (K_2)	0.6298	0.5038	0.5598	0.5637	0.4510	0.5011	0.5273	0.4218	0.4687	0.3062	0.3148	0.3104
LTN-PR (K_3)	0.6328	0.6328	**0.6328**	0.5604	0.5604	0.5800	0.6370	0.5092	0.5662	0.2780	0.3427	0.3069

Scalability to million-level markers



(a) Sequence length = 5. (b) Sequence length = 25. (c) Sequence length = 50.