



南京大學

NANJING UNIVERSITY

RINC



基于深度学习与点过程的事件序列 预测算法研究

- 答辩人：王言 MG1937025
- 导师：申富饶 教授



目录

CONTENTS

1 研究背景

2 研究内容

- 基于动态关系建模的事件预测模型 (PGG-TPP)
- 基于软标签的事件序列预测辅助训练方法 (DBSL-Aux)

3 实际应用

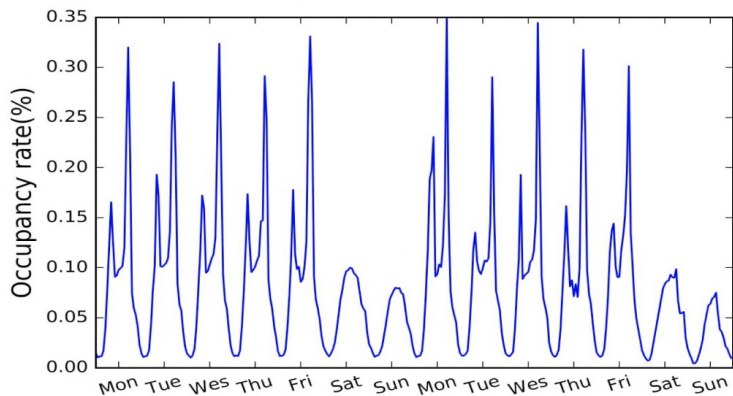
- 事件序列预测在图书管理系统中的应用

4 总结与展望

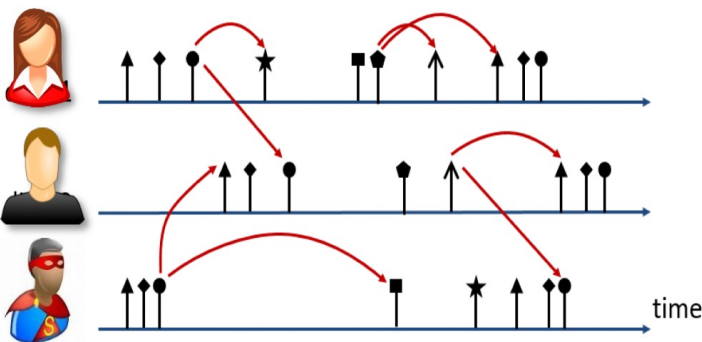


第一部分

Research Background 研究背景

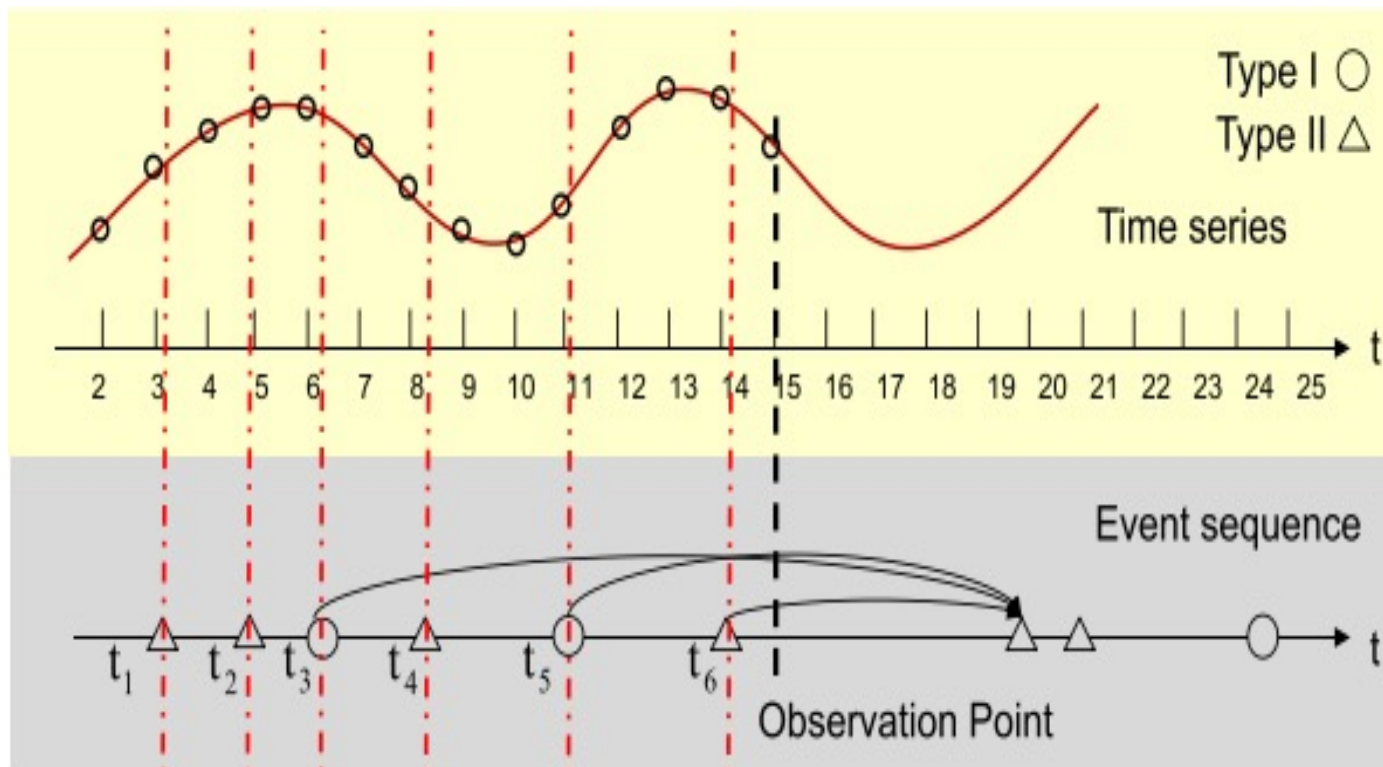


■ 一般时间序列：在时间或空间域上
均匀分布的数据点序列



■ 事件序列：在时间或空间域上非均
匀分布的数据点序列

■ 一般时间序列预测 vs 事件序列预测

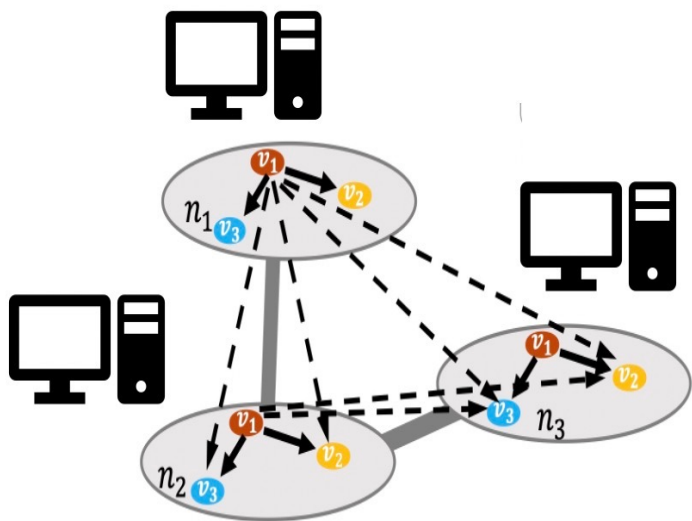


一般时间序列预测： $\{n_1, n_2, n_3 \dots n_N\} \rightarrow \{n_{N+1}\}$

事件序列预测： $\{(t_1, m_1), (t_2, m_2), (t_3, m_3) \dots (t_N, m_N)\} \rightarrow \{(t_{N+1}, m_{N+1})\}$

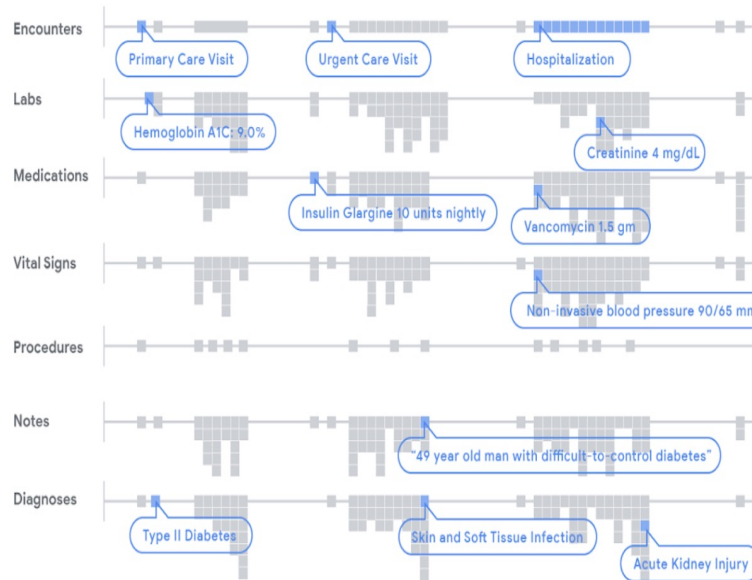


■ 事件序列预测：应用场景



■ 设备异常检测

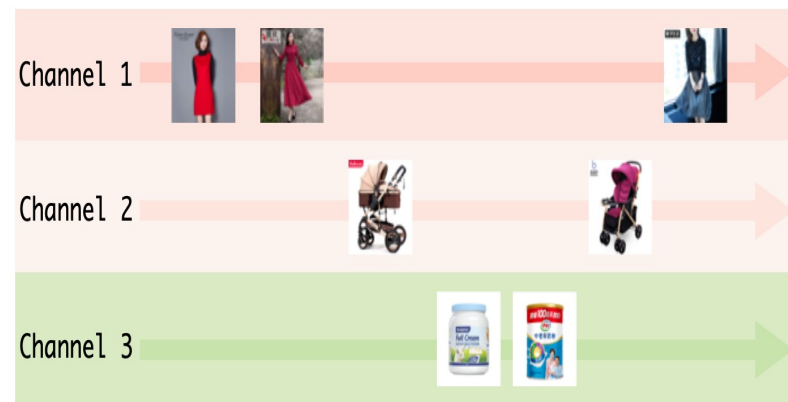
Patient Timeline



■ 疾病早期预警



User behavior sequence

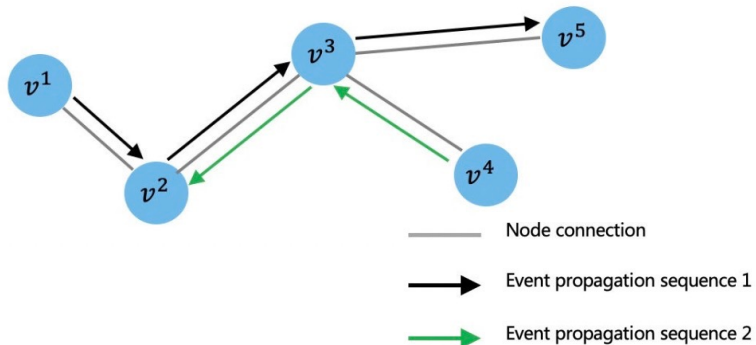


■ 用户购买行为预测

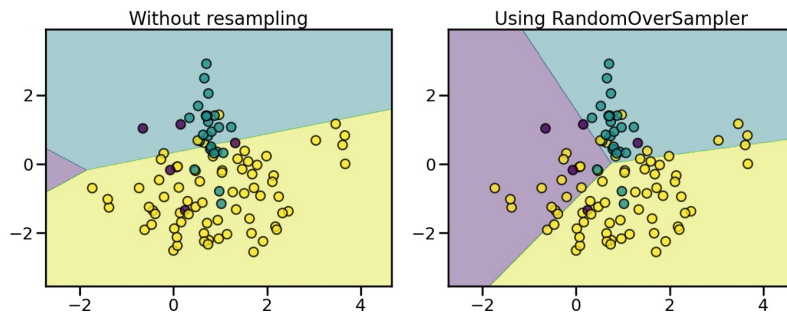


困难与挑战

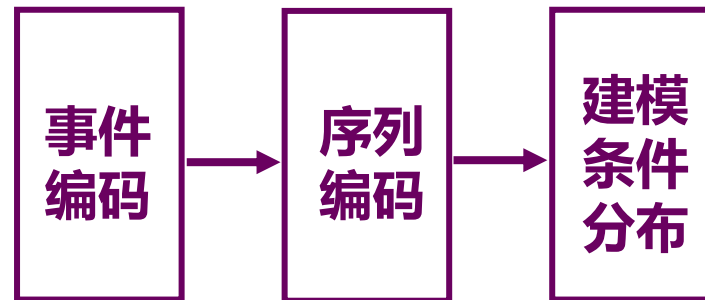
1 事件间复杂关系建模的问题



2 事件序列中的不平衡问题



点过程算法流程



● 优化事件的特征表示

● 提升模型的预测性能

● 为模型提供可解释依据





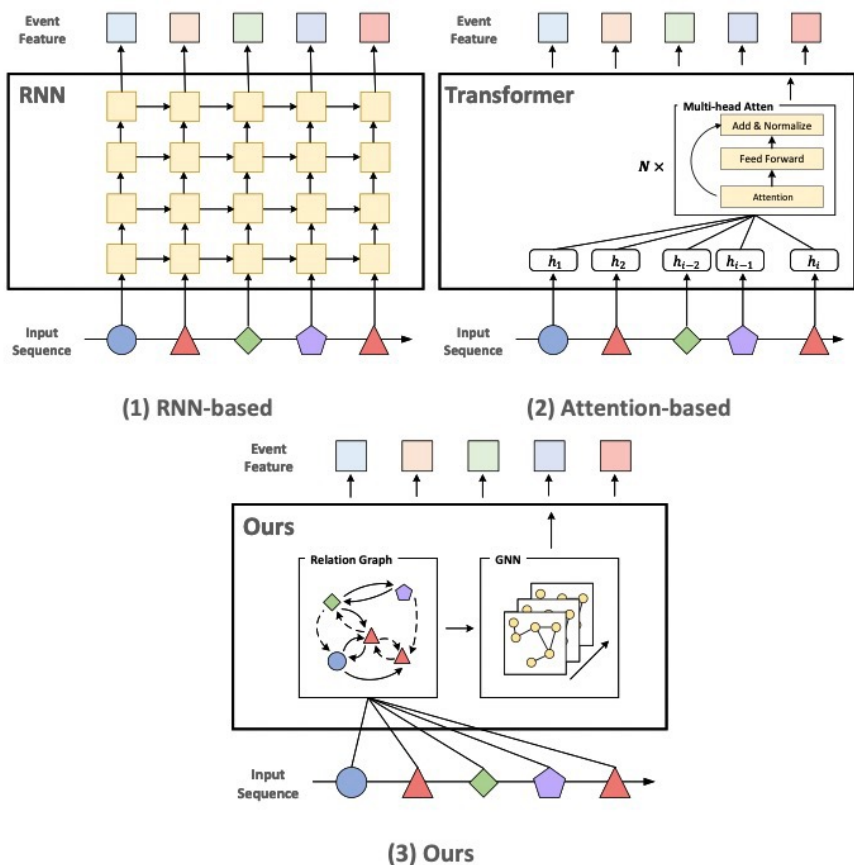
第二部分

Proposed Methods

研究内容

- 基于动态关系建模的事件预测模型 (PGG-TPP)
- 基于软标签的事件序列预测辅助训练方法 (DBSL-Aux)

基于动态关系建模的事件预测模型



- 历史序列中不同事件间影响关系相对复杂，难以被显式的定义
- 事件间影响关系种类并非单一
- 事件间的关系是随时间的推移动态变化的

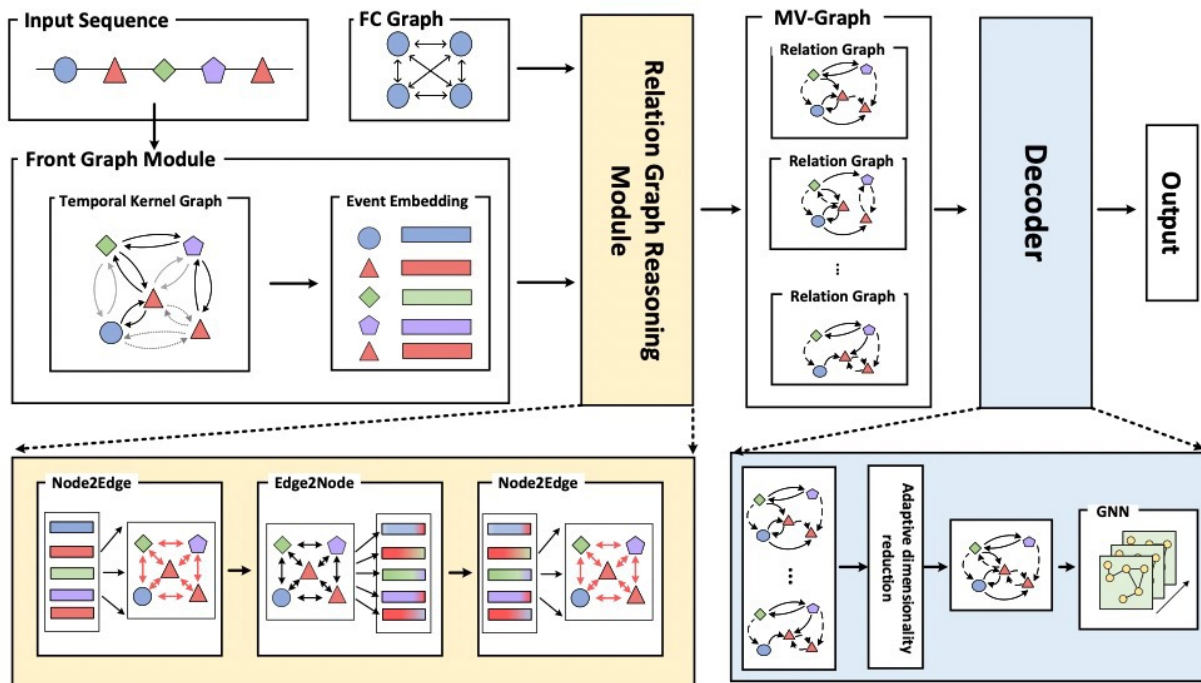


我们的解决思路

利用难度递进方式学习事件间复杂关系
利用多维图对事件间影响关系进行表示
对于事件间关系进行动态建模



基于渐进生成图的时序点过程 (PGG-TPP)



模型整体结构

Encoder

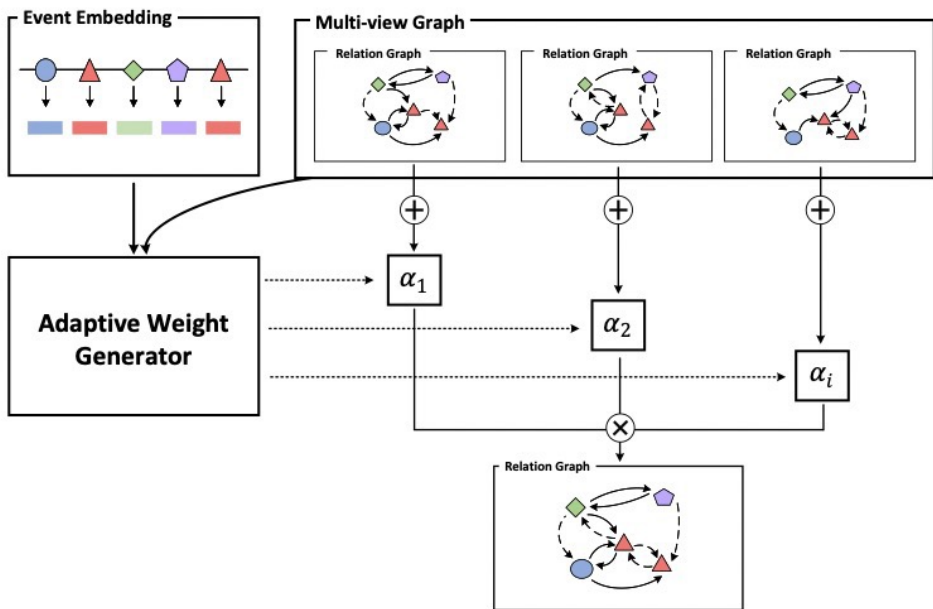
- 前置图 (Front Graph Module)
- 关系图推理 (Relation Graph Reasoning Module)

Decoder

- 自适应降维模块 (Adaptive Dimensionality Reduction)
- 图推理 (Relation Graph Reasoning Module)



基于渐进生成图的时序点过程 (PGG-TPP)



自适应降维模块

权重计算

$$\alpha_i = \begin{cases} \frac{1}{N^{RT} + \exp(-GW(\Phi_i^{MVRG}, C))}, & \text{if } 0 < i < N^{RT} \\ 1 - \sum_{j=1}^{N^{RT}-1} \alpha_j, & \text{if } i = N^{RT} \end{cases}$$

降维

$$\phi^{SDRG} = \sum_{i=1}^{N^{RT}} \alpha_i \phi_i^{MVRG}$$

GNN推理

$$HS_i^1 = GCN^1(C_i) = \sigma(\tilde{\Phi}_i^{SDRG} C_i W_1^{GRM})$$

$$HS_i^2 = GCN^2(HS_i^1) = \sigma(\tilde{\Phi}_i^{SDRG} HS_i^1 W_2^{GRM})$$



实验验证

- 数据集：2组仿真数据集，3组真实场景数据集
- Baseline：2类模型（RNN，Attention）

Dataset	Model Type	Model	ACC	RMSE
Synthetic-10	RNN	RMTTP	32.44%	5.432
	RNN	Intensity-RNN	33.34%	4.343
	RNN	NHP	33.63%	4.476
	Attention	THP	33.25%	4.922
	Attention	SAHP	31.10%	6.067
	RelationGraph	PGG-TPP	33.38%	4.405
Synthetic-100	RNN	RMTTP	29.79%	5.314
	RNN	Intensity-RNN	29.91%	5.102
	RNN	NHP	30.17%	4.885
	Attention	THP	28.73%	4.821
	Attention	SAHP	27.02%	6.832
	RelationGraph	PGG-TPP	30.67%	4.934

仿真数据集实验结果

ATM	RNN	RMTTP	76.70%	6.221
	RNN	Intensity-RNN	76.14%	3.302
	RNN	NHP	73.67%	7.031
	Attention	THP	70.71%	3.820
	Attention	SAHP	67.20%	4.591
	RelationGraph	PGG-TPP	80.11%	3.108
IPTV	RNN	RMTTP	56.67%	22.574
	RNN	Intensity-RNN	58.22%	20.218
	RNN	NHP	50.06%	18.812
	Attention	THP	72.10%	12.780
	Attention	SAHP	71.83%	13.211
	RelationGraph	PGG-TPP	74.31%	11.131
Weeplace	RNN	RMTTP	21.97%	7.320
	RNN	Intensity-RNN	23.72%	7.011
	RNN	NHP	25.17%	6.219
	Attention	THP	29.10%	6.695
	Attention	SAHP	28.65%	6.889
	RelationGraph	PGG-TPP	30.38%	6.445

真实场景数据集实验结果



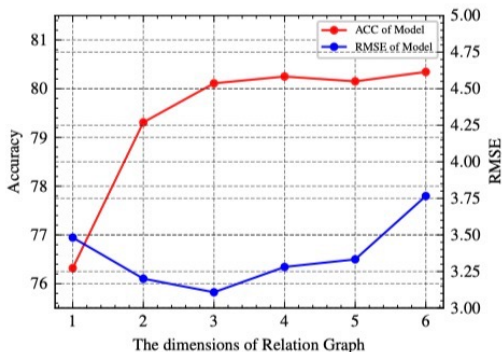
实验验证

Model	ACC	RMSE
Only-TKG	78.52%	3.472
Only-MVRG	77.84%	3.550
Ours	80.11%	3.108

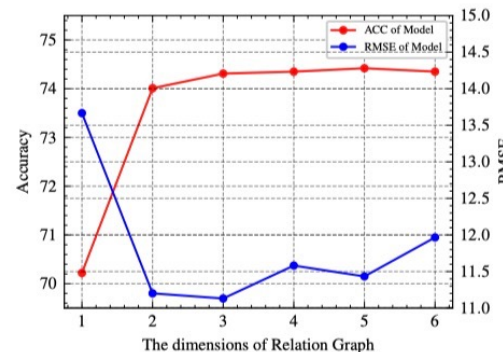
不同学习策略对模型性能影响

Model	ACC	RMSE
$N_s = 1$	77.33%	3.394
$N_s = 2$	80.11%	3.108
$N_s = 3$	80.08%	3.096
$N_s = 4$	80.13%	3.135

不同关系推理步数对模型性能影响

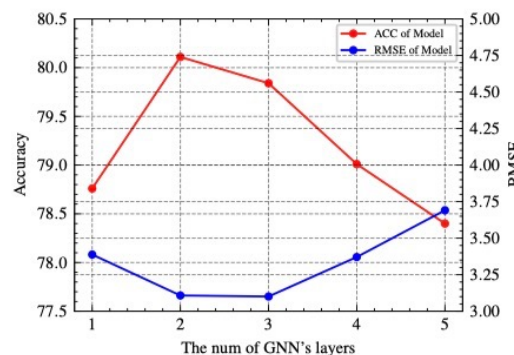


(a) ATM

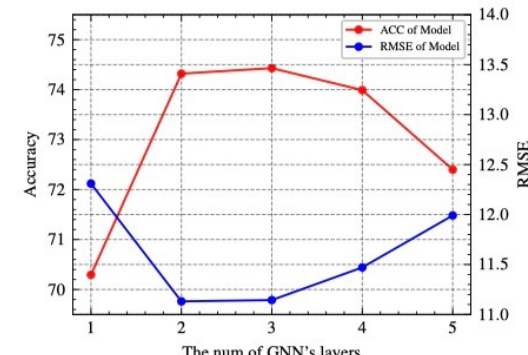


(b) IPTV

图 3-5: 关系图结构使用不同维度下模型的性能对比 (以 ACC 和 RMSE 度量)



(a) ATM

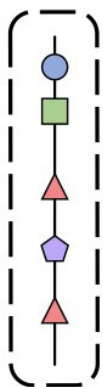


(b) IPTV

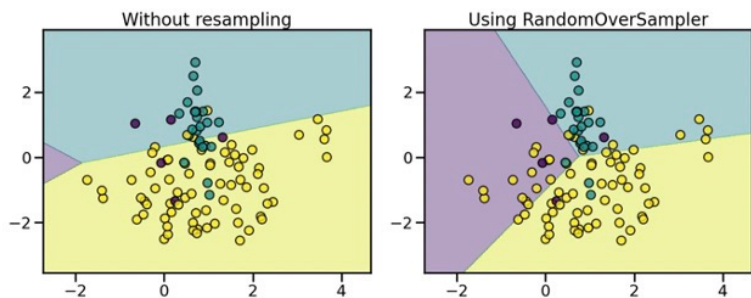
图 3-6: 不同图神经网络层数下模型的性能对比 (以 ACC 和 RMSE 度量)



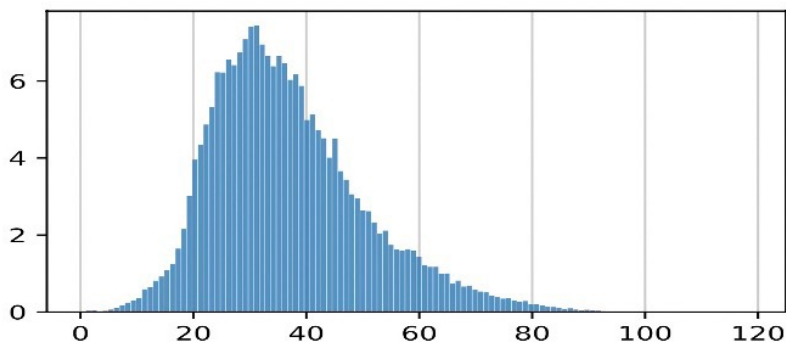
基于软标签的事件序列预测辅助训练方法



Input



离散空间中的不平衡



连续空间中的不平衡

事件序列预测问题中的不平衡

- 离散标记信息空间中的不平衡
- 连续时序空间中的不平衡



我们的解决思路

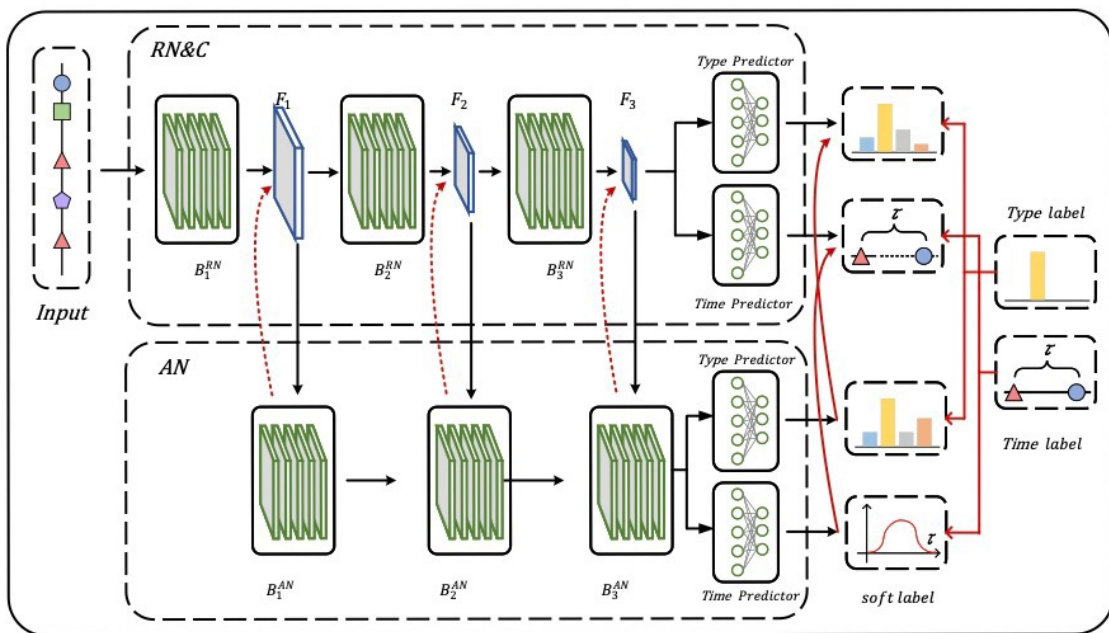
设计Soft Label优化的解耦学习框架

设计辅助网络进行Soft Label的生成

设计针对连续空间的辅助软标签



基于双重平衡软标签的辅助训练方法 (DBSL-Aux)



模型整体结构

模型基本构架

- Self-Attention结构
- 基础网络 + 辅助网络

表示学习阶段

- 时序预测器 (Future Time predictor)
- 标记信息预测器 (Future Mark predictor)
- 时序软标签生成器 (Time Soft label generator)
- 标记软标签生成器 (Mark Soft label generator)

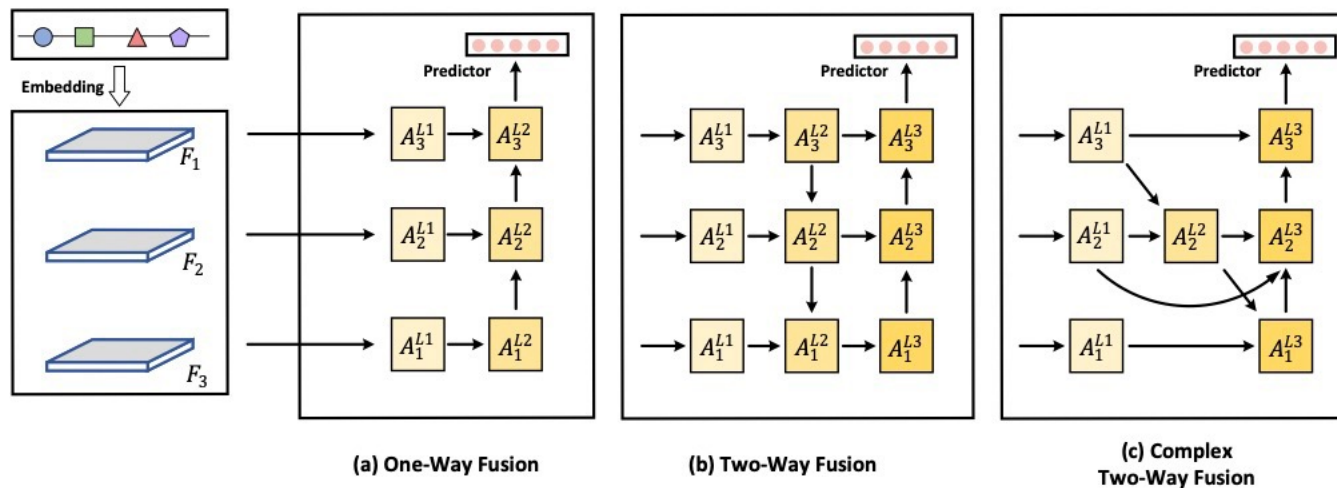
预测器学习阶段

- 重加权学习方式

基于双重平衡软标签的辅助训练方法 (DBSL-Aux)

辅助网络结构

- 单向融合结构
- 双向融合结构
- 交叉融合结构



模型学习

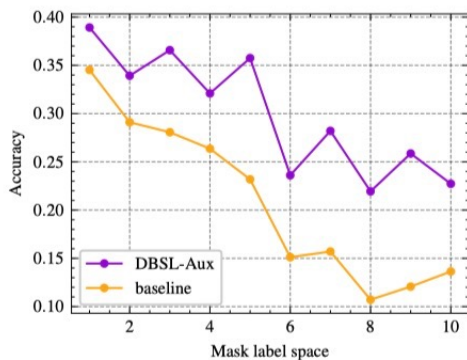
- 标签对于基础网络的监督
- 标签对于辅助网络的监督
- 辅助网络对于基础网络的监督

$$loss = l_{gt2b}(X) + \beta_{gt2a} l_{gt2a}(X) + \beta_{a2b} l_{a2b}(X)$$

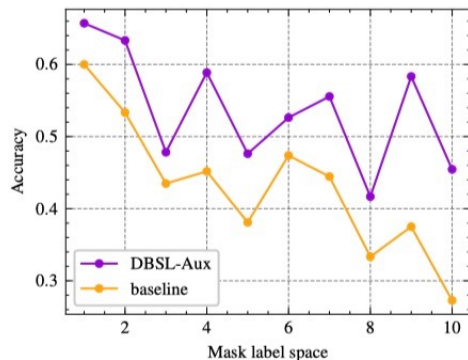


实验验证

- 数据集：3组真实场景数据集
- Baseline：2类模型 (RNN , Attention)



(a) StackOverflow



(b) MIMIC-II

模型提升对于少样本类别预测准确率

Dataset	Model Type	Model	ACC	Macro-F1	RMSE
Retweets	RNN	RMTTP	52.12%	0.4320	37.23
	RNN	NHP	54.97%	0.4872	35.21
	MLP	FullyNN-TTP	56.03%	0.5144	35.76
	Attention	THP	58.83%	0.5379	33.24
	Attention	SAHP	55.28%	0.5021	34.45
	Attention	DBSL-Aux	59.27%	0.5414	33.81
StackOverflow	RNN	RMTTP	43.10%	0.2428	7.455
	RNN	NHP	44.21%	0.2547	6.084
	MLP	FullyNN-TTP	44.17%	0.2503	6.342
	Attention	THP	45.77%	0.2813	4.284
	Attention	SAHP	43.19%	0.2441	3.807
	Attention	DBSL-Aux	46.28%	0.3006	4.162
MIMIC-II	RNN	RMTTP	81.10%	0.4826	6.71
	RNN	NHP	83.66%	0.5275	3.38
	MLP	FullyNN-TTP	82.46%	0.4597	4.34
	Attention	THP	85.12%	0.5405	1.250
	Attention	SAHP	83.93%	0.5054	2.123
	Attention	DBSL-Aux	85.75%	0.5535	1.192

真实场景数据集实验结果



实验验证

Model	ACC	Macro-F1	RMSE
baseline	45.22%	0.2768	4.540
Label Smoothing	44.93%	0.2702	4.451
OLS	45.30%	0.2759	4.201
Ours	46.28%	0.3006	4.162

不同软标签生成策略对模型性能影响

Model	ACC	Macro-F1	RMSE
baseline	45.22%	0.2768	4.540
baseline + soft-mark-label	45.98%	0.2955	4.522
baseline + soft-time-label	44.86%	0.2751	4.173
baseline + feature-level	45.25%	0.2773	4.536
Ours	46.28%	0.3006	4.162

不同辅助监督方法对模型性能影响

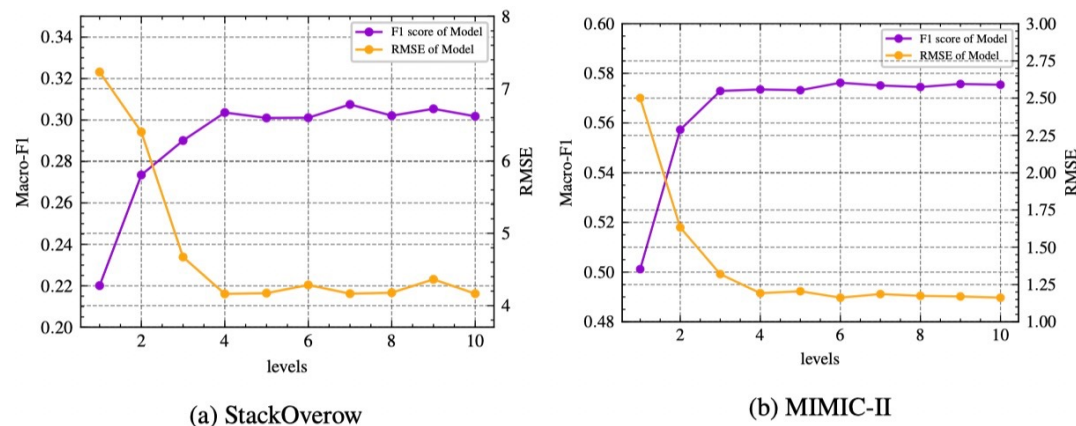


图 4-7: 不同特征提取模块数量下模型的性能对比 (以 Macro-F1 和 RMSE 度量)

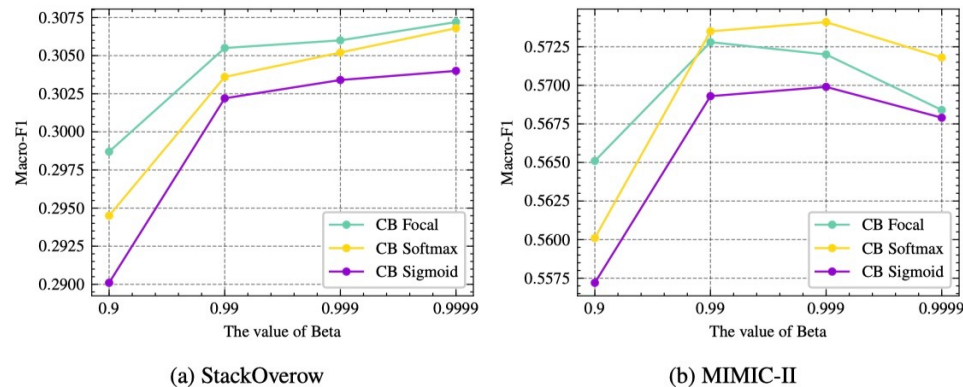


图 4-8: γ 不同取值下模型的性能对比 (以 Macro-F1 度量)



第三部分

Applications 实际应用

- 事件序列预测在图书管理系统中的应用



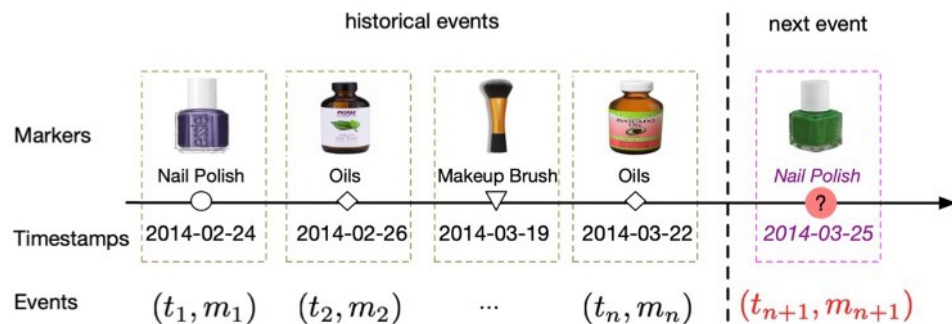
图书管理系统中的算法需求

■ 基础推荐算法

- 提前针对已登陆用户的精准推荐

■ 辅助推荐算法

- 针对未登陆游客用户的辅助推荐



利用事件序列预测模型构造辅助推荐算法

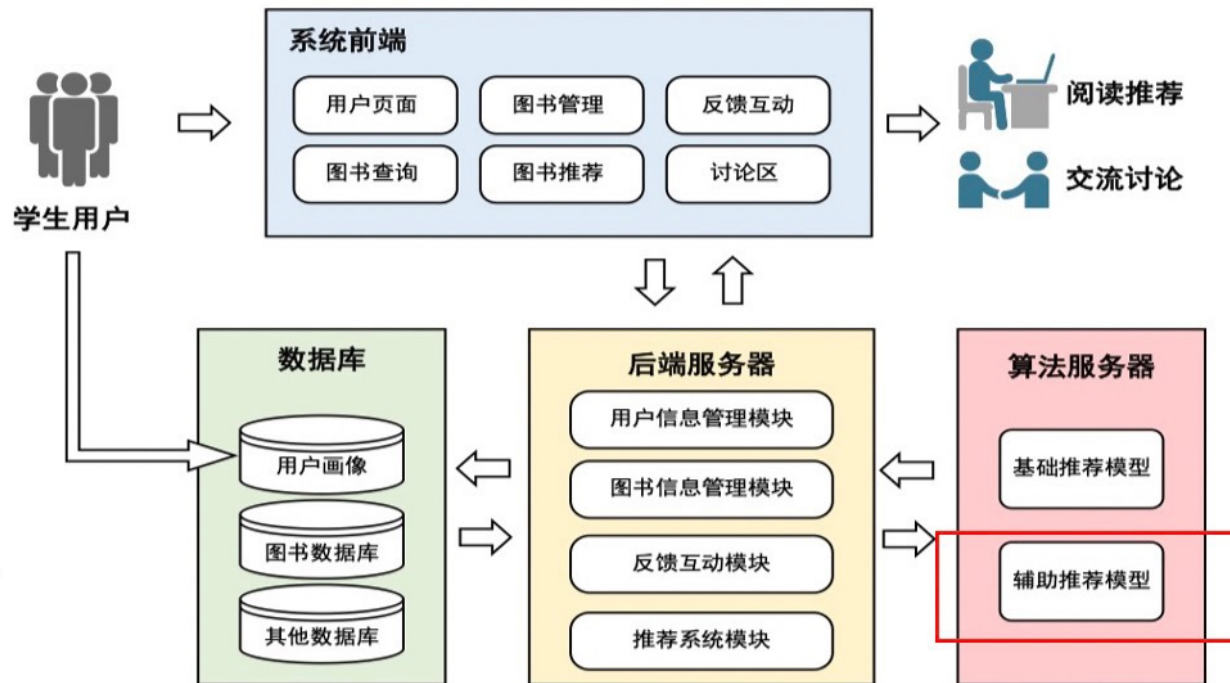
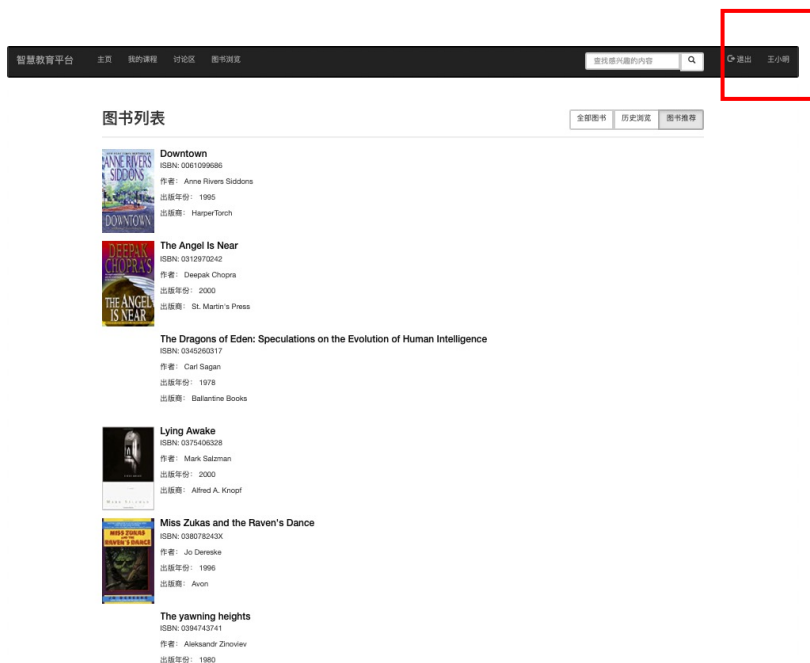


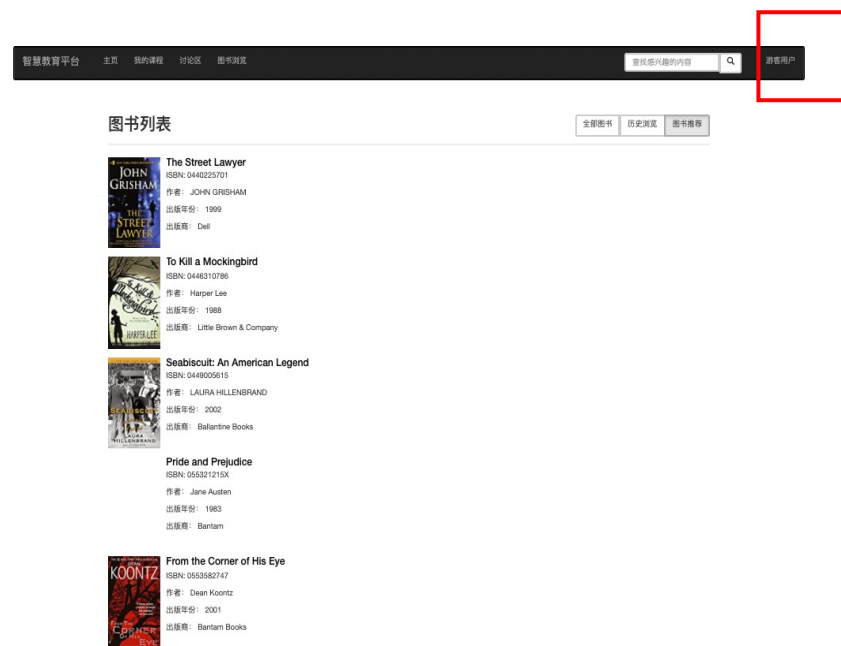
图 5-1: 系统整体架构图



图书管理系统中的推荐算法：效果展示



针对已登陆用户的推荐



针对未登陆用户的推荐



第四部分

Summary 总结与展望



全文工作总结

基于动态关系建模的事件预测模型

- 提出PGG-TPP模型，利用难度递增方式学习事件间复杂关系
- 利用多维图对事件间影响关系进行表示
- 对于事件间关系进行动态建模

基于软标签的事件序列预测辅助训练方法

- 定义事件序列预测任务中的不平衡问题
- 提出DBSL-Aux模型
- 设计Soft Label优化的解耦学习框架

事件序列预测在图书管理系统中的应用

- 将针对未登陆用户的推荐功能抽象为事件序列预测任务
- 利用Flask框架将提出的算法进行部署



未来工作展望

优化事件间信息传递过程

在生成事件间关系图之后，我们仅使用基础的图神经网络进行之后的节点信息传递，但是如何利用关系图结构进行更好的信息传递依然需要进一步探索

更有效的时序编码

对于事件序列中的时序信息的特征学习，我们仅使用相对位置信息，如何利用更完善的位置编码信息优化时序特征需要进一步探索

拓展应用场景

将本文所提出的事件序列预测算法应用于更多实际场景中



南京大學
NANJING UNIVERSITY



谢谢！

