



网络内容认知与行为对抗技术

报告人：彭浩

penghao@buaa.edu.cn

<https://penghao-bdsc.github.io/>





彭浩，北京航空航天大学，网络空间安全学院，副教授，博士生导师，互联网舆情与安全治理实验室主任，服务Springer Nature出版集团旗下JMLC期刊（SCI, JCR Q1）Associate Editor编委；

近五年发表IEEE TPAMI、TKDE、TC、ACM TOIS、WWW、SIGIR、NeurIPS、AAAI、IJCAI等学术论文80余篇，主持国家自然科学基金青年基金、重点研发子课题、北京市自然科学基金面上等项目课题10余项

主要研究方向：网络舆情事件监测、用户行为异常检测、网络群体行为对抗；

科研获奖：ESI热点论文1篇（ACM TOIS 2021）、ESI高被引论文3篇（IEEE TKDE 2019、TITS 2019、Information Science 2020），ACM CIKM2022 Best Paper Runner Up、国际旗舰学术会议Paper Digest Most Influential Papers 6篇（WWW2018、CIKM2019、SIGIR2020、CIKM2020、WWW2021），获2018年中国电子学会技术发明一等奖、2020年中国人工智能学会优秀博士学位论文、2022年中国仿真学会自然科学二等奖、2022年河北省科技进步一等奖等



CONTENTS

研究背景

RESEARCH
BACKGROUNDS

1

研究思路

RESEARCH IDEAS

2

研究工作

RESEARCH WORKS

3

4

系统应用

SYSTEM APPLICATIONS

5

未来工作

FUTURE WORK



互联网内容安全与治理-突发性事件



当前全球互联网每日新增网络消息超100亿条



美国当地时间3月23日上午10时，
TikTok首席执行官周受资出席美国国会众议院能源
和商务委员会听证会



互联网内容安全与治理-突发性事件

[韩总统就踩踏事故致歉 承诺改善制度防悲剧重演-新闻中心-...](#)

11小时前 内容提要 当地时间7日,韩国总统尹锡悦就梨泰院踩踏事故表示,对遭遇悲剧的遗属和同感悲伤与痛苦的国民深感歉疚。韩国国务总理韩德洙当天也就踩踏事故向国民和受害外国...

北方网

[154人遇难!韩国梨泰院踩踏事件真相曝光,太揪心了! 事故 人...](#)

8天前 10月29日晚,整个韩国陷入了至暗时刻。这一晚,韩国首尔梨泰院发生极为悲惨的踩踏事故。截止到目前,已造成了154人死亡,149人受伤。令人心痛的是,在这154名遇难者中,有4...

搜狐网

[韩国梨泰院踩踏事故调查进展:6人因涉嫌公务过失致死伤罪被...](#)

2天前 央视新闻客户端消息,据韩国警察厅特殊调查本部11月7日通报,包括龙山警察署署长李林宰、踩踏事故发生当晚首尔警察厅负责情况管理的值班总警柳美珍、龙山消防署署长崔成范...

上游新闻

[痛心!韩国踩踏事故已致154人死亡!超10万人聚集,悲剧如何发...](#)

8天前 据央视新闻报道,韩国首尔警察厅当地时间30日晚间通报,截至当天18时,又有一名女性在梨泰院踩踏事故中因受伤过重而死亡。至30日晚,梨泰院踩踏事故中的死亡人数已升至154人...

搜狐网

[韩国首尔发生踩踏事故,已致151人死亡 手机搜狐网](#)

2022年10月30日 当地时间10月29日晚,韩国首尔龙山区梨泰院一带发生大规模踩踏事故。目前,梨泰院踩踏事故已致151人死亡。中国驻韩国大使馆向总台记者确认:目前已有三名中国公...

搜狐网

[惨烈!韩国踩踏事故现场曝光:别让生命成为狂欢的代价 - 蓝...](#)

韩国首尔梨泰院发生严重踩踏事故。到目前为止,已有153人死亡,100多人受伤 图片来源:网络 新闻在全世界媒体上迅速刷屏。事故起源于10月29日的万圣节活动。由于新冠肺炎疫情...

www.scsdsl.cn/2193...html



2022年10月31日,英国《独立报》把韩国首尔梨泰院踩踏事故发生地点错误写成“香港”

互联网内容安全与治理-突发性事件

www.xinhuanet.com › world ▼

热点问答：英国货车惨案凸显欧洲非法移民问题-新华网

2019年10月29日 - 英国埃塞克斯郡近日发生的**货车惨案**震惊世界。藏身**货车**中的39人不幸死亡令人痛心，此次**惨案**暴露出来的欧洲非法移民、人口贩卖以及边境管控等 ...

www.xinhuanet.com › world ▼

英国货车惨案司机承认参与协助偷渡-新华网

2019年11月26日 - 英国**货车惨案**司机承认参与协助偷渡---英国埃塞克斯郡警方25日发表声明说，上月发生在该郡的货车惨案涉事司机承认参与协助非法移民偷渡。

www.nbd.com.cn › articles ▼

痛心！外媒公布英国货车惨案遇难者照片和信息最小年仅15岁 ...

2019年11月9日 - ... 英国埃塞克斯郡警方11月8日公布了埃塞克斯郡集装箱**货车**案39名遇难者的身份信息。他们均为越南公民，年龄大多在20岁至40岁之间，遇难 ...

world.people.com.cn › 国际 ▼

英国货车惨案全部遇难者遗体 and 骨灰运回越南--国际--人民网

2019年12月1日 - 11月30日，在越南首都河内，工作人员运送遇难者遗体。新华社/越通社新华社河内12月1日电越南外交部日前说，**英国货车惨案**全部遇难者遗体 and ...

www.bbc.com › zhongwen › simp ▼

英国集装箱惨案：39名死者“均为中国公民” - BBC News 中文

2019年10月25日 - 警方把集装箱卡车移送附近仓库开展遗体辨认工作。**英国**警方证实，埃塞克斯郡集装箱藏尸案中的39名死者“均为中国公民”，其中8 ...

dangjian.com › 国际观察 ▼

英国货车惨案司机认罪---党建网

2020年4月9日 - 新华社伦敦4月8日电（记者张代蕾杨海若）**英国**埃塞克斯郡警方8日发表声明说，去年10月份发生在该郡的**货车惨案**涉事司机当天承认过失杀人。

BBC NEWS 中文

主页 国际 两岸 英国 科技 财经 视频材料 BBC英伦网

英国集装箱惨案：39名死者“均为中国公民”

2019年10月24日
最近更新：2019年10月25日



警方把集装箱卡车移送附近仓库开展遗体辨认工作。

英国警方证实，埃塞克斯郡集装箱藏尸案中的39名死者“均为中国公民”，其中8女31男。

CNN World Africa Americas Asia Australia China Europe India Middle East United Kingdom

Why would people from China, the world's second-biggest economy, risk their lives to enter the UK?

By Aimee Lewis, CNN
Updated 1705 GMT (0305 HKT) October 25, 2019

为什么中国已经成为世界上第二大经济体，却还有人冒着生命危险偷渡英国



Paid Content by @hubrain

Trump voter shifts stance on impeachment inquiry: Here's why
President Trump will stop at nothing to win a second term

REUTERS Business Markets World Politics TV More

INDUSTRIALS OCTOBER 24, 2019 / 5:46 PM / UPDATED 38 MINUTES AGO

The 39 people found dead in truck near London were Chinese nationals- ITV

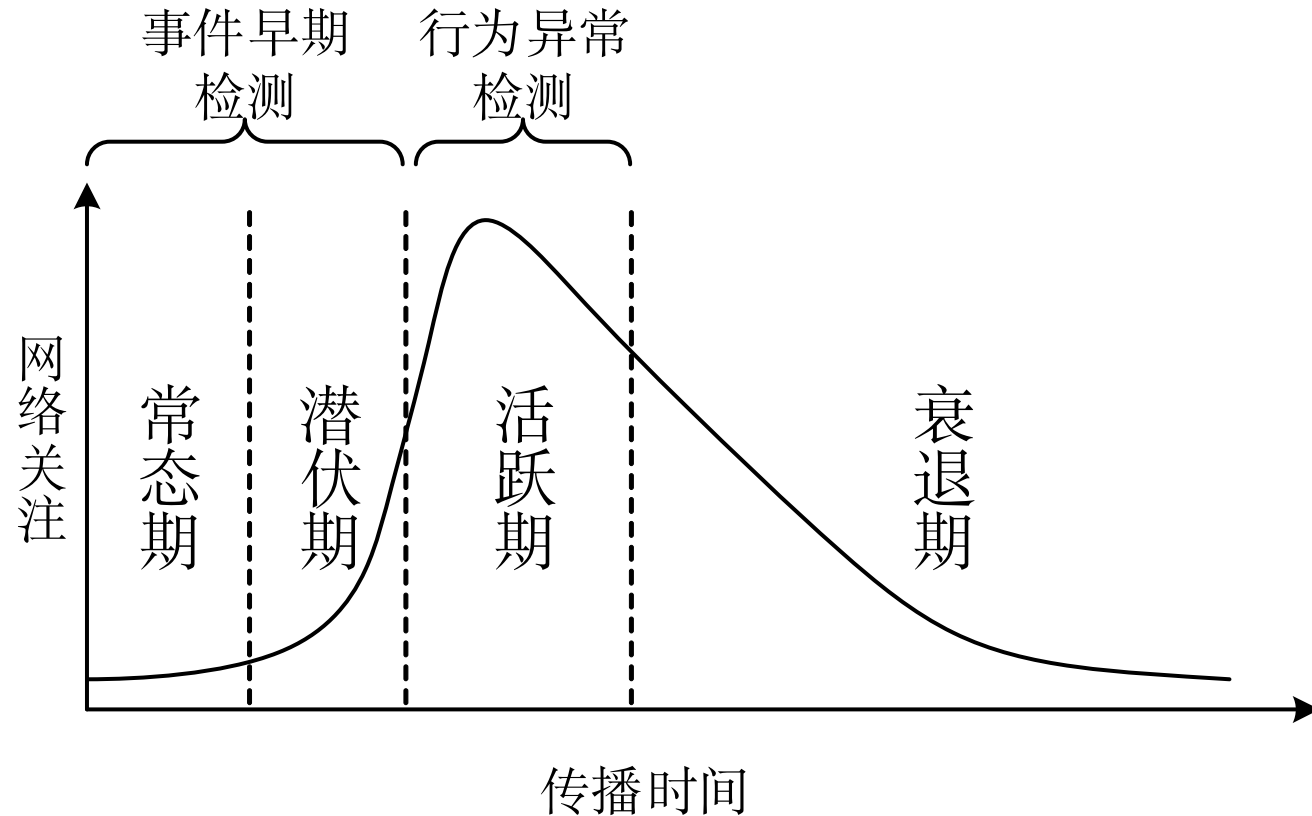
1 MIN READ

2019年10月25日，英国BBC报道、法国路透社等西方媒体错误报道英国集装箱惨案39名死者“均为中国公民”

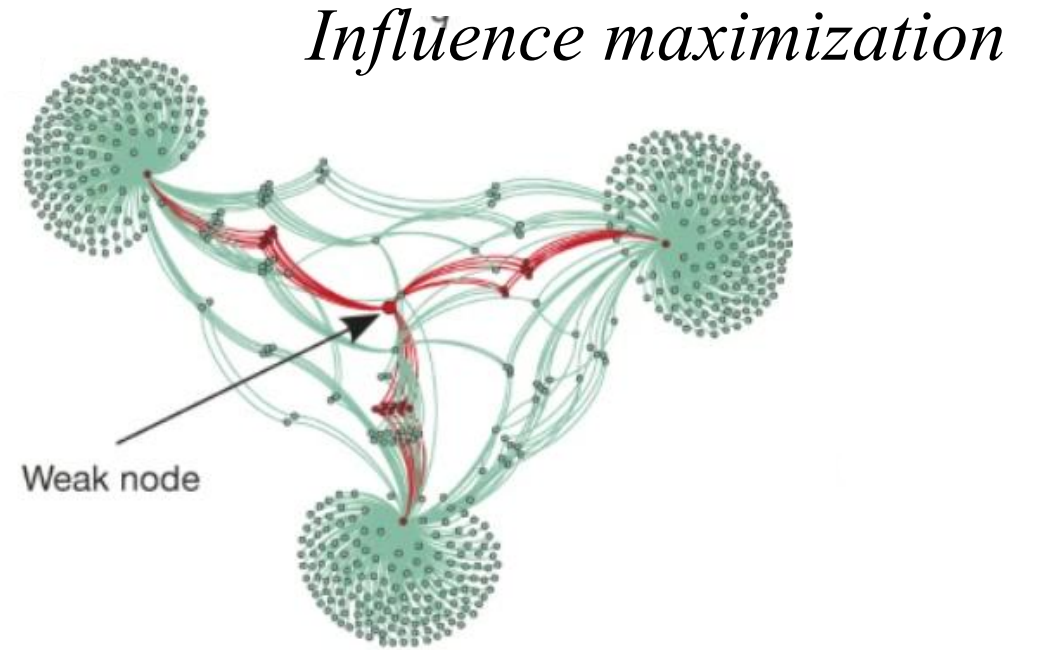


战争、疫情、冲突等突发事件时刻都在发生，威胁政府、民众的公共安全

互联网内容安全与治理-传播分析



社会事件网络舆情的政府干预机制. 中国科学院院刊, 2015



Influence maximization in complex networks through optimal percolation. Nature, 2015

互联网内容安全与治理-事件分析

www.xinhuanet.com › world ▼

热点问答：英国货车惨案凸显欧洲非法移民问题-新华网

2019年10月29日 - 英国埃塞克斯郡近日发生的货车惨案震惊世界。藏身货车中的39人不幸死亡令人痛心，此次惨案暴露出来的欧洲非法移民、人口贩卖以及边境管控等 ...

www.xinhuanet.com › world ▼

英国货车惨案司机承认参与协助偷渡-新华网

2019年11月26日 - 英国货车惨案司机承认参与协助偷渡——英国埃塞克斯郡警方25日发表声明说，上月发生在该郡的货车惨案涉事司机承认参与协助非法移民偷渡。

www.nbd.com.cn › articles ▼

痛心！外媒公布英国货车惨案遇难者照片和信息最小年仅15岁 ...

2019年11月9日 - ... 英国埃塞克斯郡警方11月8日公布了埃塞克斯郡集装箱货车案39名遇难者的身份信息。他们均为越南公民，年龄大多在20岁至40岁之间，遇难 ...

world.people.com.cn › 国际 ▼

英国货车惨案全部遇难者遗体 and 骨灰运回越南--国际--人民网

2019年12月1日 - 11月30日，在越南首都河内，工作人员运送遇难者遗体。新华社/越通社新华社河内12月1日电越南外交部日前说，英国货车惨案全部遇难者遗体 and ...

www.bbc.com › zhongwen › simp ▼

英国集装箱惨案：39名死者“均为中国公民” - BBC News 中文

2019年10月25日 - 警方把集装箱卡车移送附近仓库开展遗体辨认工作。英国警方证实，埃塞克斯郡集装箱藏尸案中的39名死者“均为中国公民”，其中8 ...

dangjian.com › 国际观察 ▼

英国货车惨案司机认罪---党建网

2020年4月9日 - 新华社伦敦4月8日电（记者张代蕾杨海若）英国埃塞克斯郡警方8日发表声明说，去年10月份发生在该郡的货车惨案涉事司机当天承认过失杀人。

事件梳理

英国当地时间10月23日，埃塞克斯郡的 Waterglade 工业园一辆货车内发现39具尸体

英国当地时间10月23日，警察局副局长 Pippa Mills 发表声明：“尚无法确认身份”

英国当地时间10月23日，BBC、镜报等英媒报道，39具遗体“为中国公民”

10月24日，中国驻英国大使馆回应

10月25日，CN

问题1: 如何从海量互联网社交数据中获知有价值的社会信息？

英国时间10月2

文章称遇难者中疑似有越南人。

越南总理阮福春下令调查此“英国货车”事件

英国时间11月1日，英国警方：认为“遇难者为越南人”

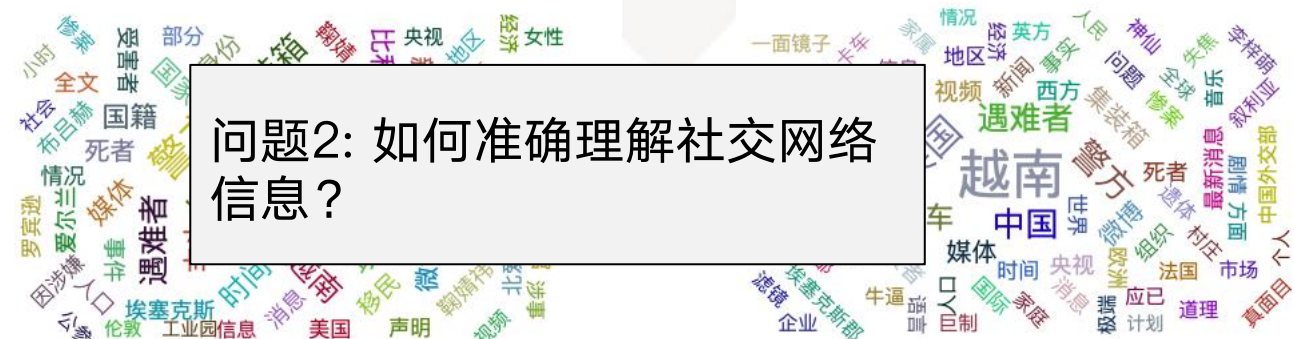
蛇头为偷渡设置“VIP路线”

2020年4月9日，英国埃塞克斯郡惨案司机已伏法认罪

阶段一：2019.10.23~2019.10.25

国内舆论变化

阶段二：2019.10.26~2019.10.29



主题类型：非法移民、人口贩卖，边境管控



互联网内容安全与治理



互联网媒体行为大数据

问题1: 如何从海量社会网络媒体大数据中获知高价值的社会信息？

问题2: 如何准确理解和组织社会网络舆情事件？

问题3: 网络舆情的发展与发酵，同网络水军行为有什么关系？

问题4: 面对群体性网络水军行为，怎么发现？怎么对抗？

核心研究挑战1

社会网络舆情监测需要克服噪音大、开放域、动态性、小样本、多语言等问题

核心研究挑战2

社会网络事件分析需要克服内容理解、多模态融合、样本分布不均衡等问题

核心研究挑战3

网络水军检测需要克服用户行为对抗和内容伪装、平台差异、数据海量等问题



阶段成果

1

异构图神经网络小样本事件检测

实现融合外部知识的多源异构社交数据的建模、特征表示和事件检测。

2

增强的网络社会舆情事件早期检测

实现多关系图建模、多智能体控制消息特征聚合表示、多尺度舆情事件检测。

3

质量感知自提升社会事件检测

利用知识迁移和质量自感知导向的优化实现对未知事件的精准检测。

4

基于时间感知图的社会事件检测

实现基于时间感知聚合器及深度学习的社交事件挖掘。

5

在线事件检测与分析系统

实现海量在线社会网络事件检测与分析系统，事件监测、事件分析、跨语言事件监测。



阶段成果

递归强化邻域选择多关系图神经网络

6

强化多关系图学习异常行为检测

7

联邦的跨平台、跨模型、跨语言水军检测

强化聚合的多视图张量图神经网络

8

9

结构熵优化增强的图结构扰动学习表征

10

结构熵优化的群体行为对抗

11

在线网络水军监测与分析系统
实现海量在线社交网络用户水军行为检测与分析。

12



研究工作：异构图神经网络事件检测

现状

Huge Social Messages



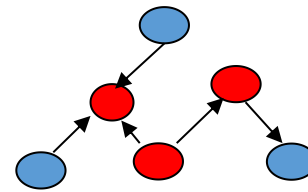
keyword or template
matching



Q: narrow fields

Traditional Methods

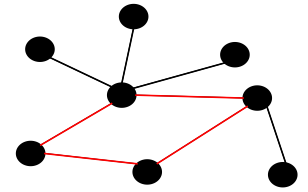
burst words



Q: low

accuracy

Abnormal subgraphs



Q: high complexity

开放域事件检测难题

思路

特征
融合
学习

Semantics of words and phrases

Social users and behavior

External knowledge information



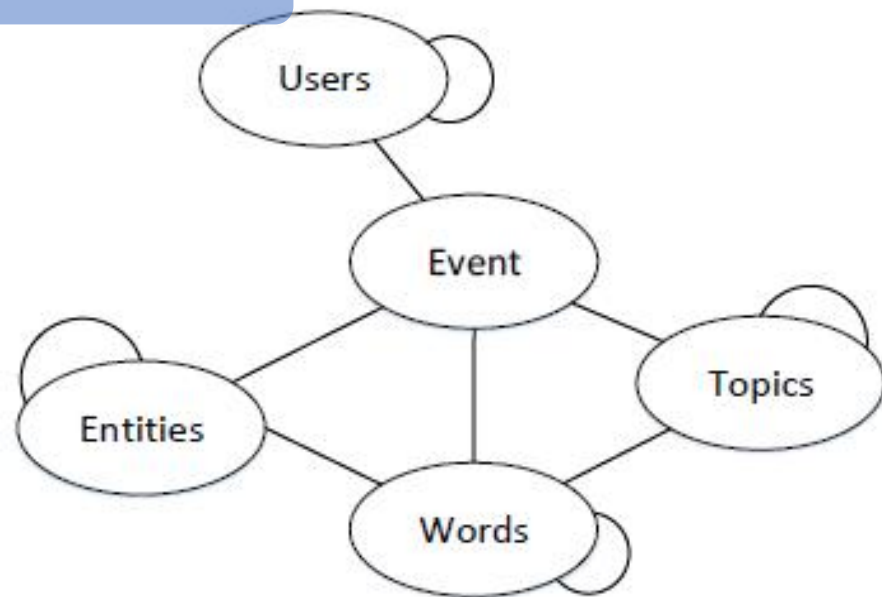
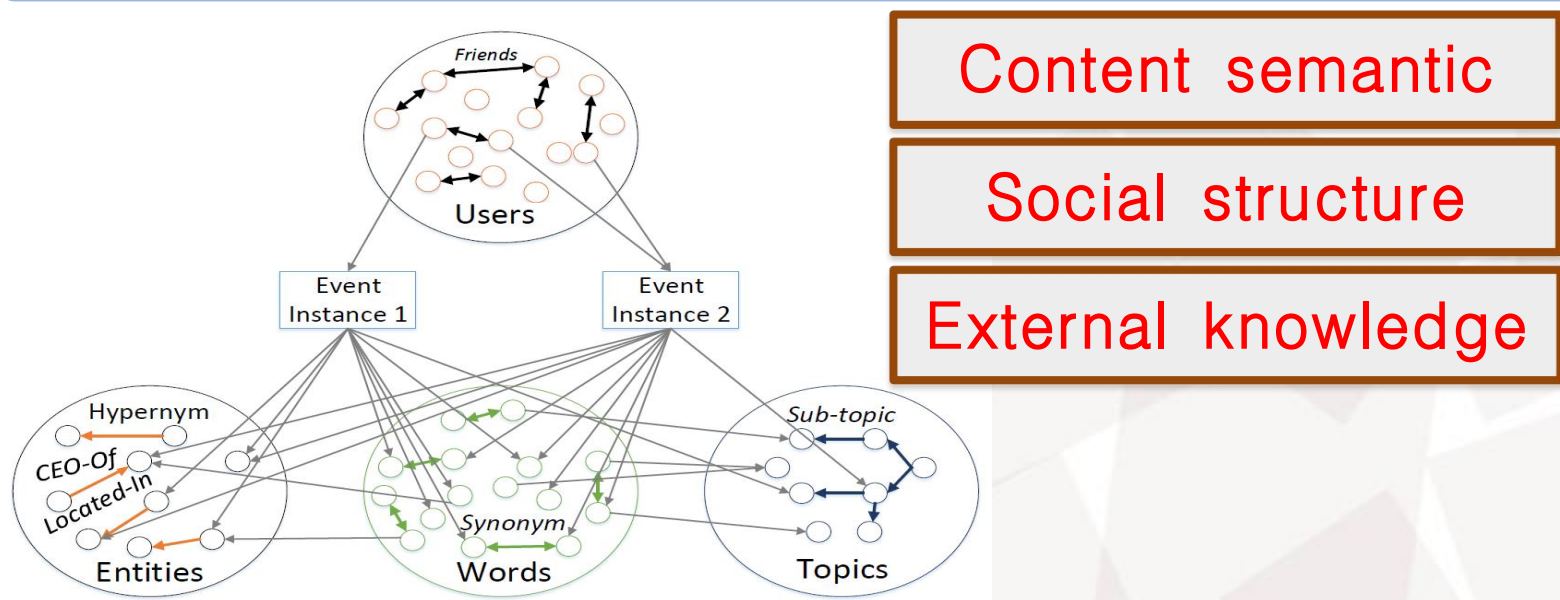
Fine-grained Social
Event Detection

Accuracy

Generalization

开放域小样本事件检测

模型1：舆情事件定义与Pairwise异构图卷积神经网络事件表征



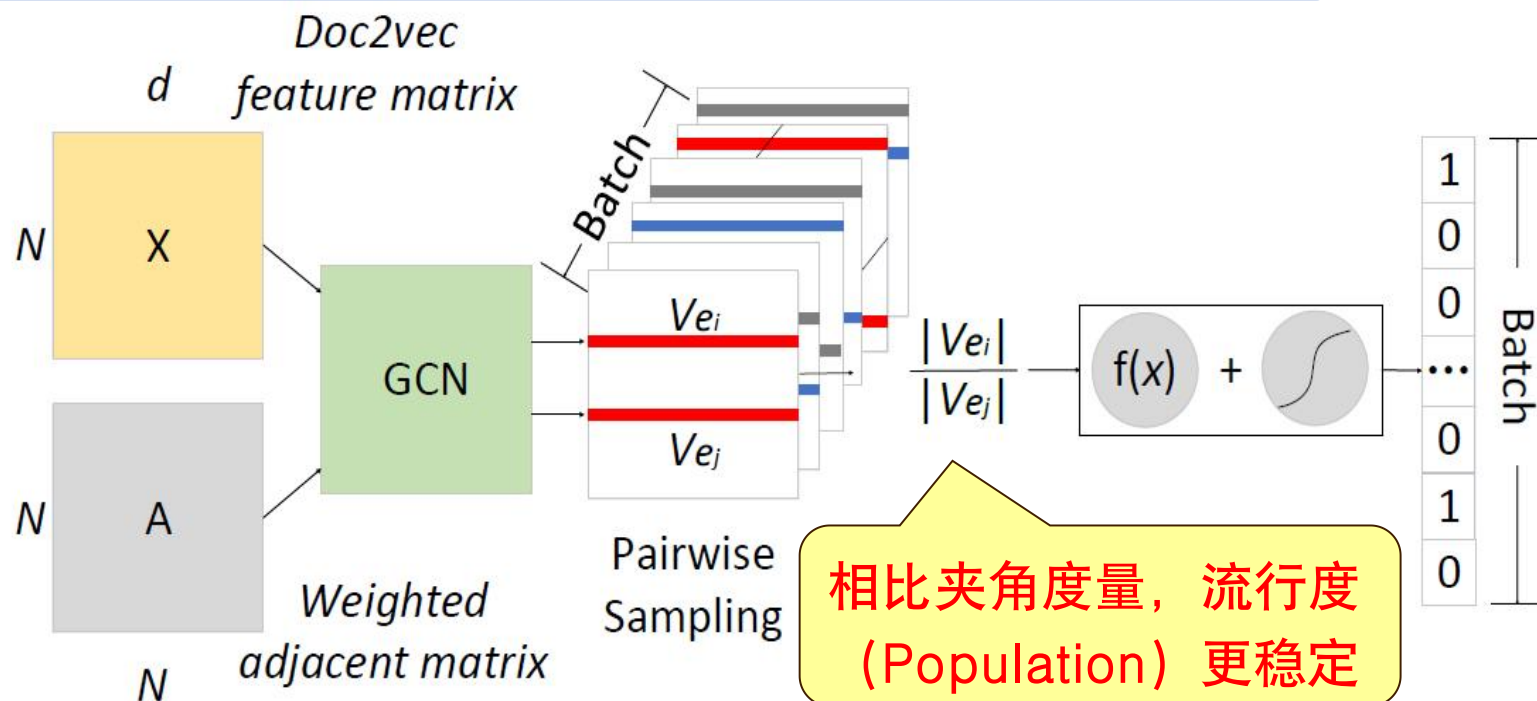
(1) Example of two event instances modeling by knowledgeable HIN

(2) Meta-schema of event-based HIN

(3) Meta-paths instances based social message similarity measure:

$$KIES(e_i, e_j) = \sum_{m=1}^{M'} \omega_m \frac{2 \times CouP_{P_m}(e_i, e_j)}{CouP_{P_m}(e_i, e_i) + CouP_{P_m}(e_j, e_j)}$$

模型1: Pairwise异构图卷积神经网络事件表征



An overview of Pairwise Popularity Graph Convolutional Network (PP-GCN)

Papadopoulos, etc., Popularity versus similarity in growing networks. Nature, 489(7417):537, 2012.

Hao Peng, etc., Fine-grained Event Categorization with Heterogeneous Graph Convolutional Networks, IJCAI2019, 谷歌学术引用 86

Datasets	Train	Validation	Test	Class
Tencent	17,438	5,813	5,812	9,941
Weibo	6,000	2,000	2,000	5,470

Table 1: Description of evaluation datasets.

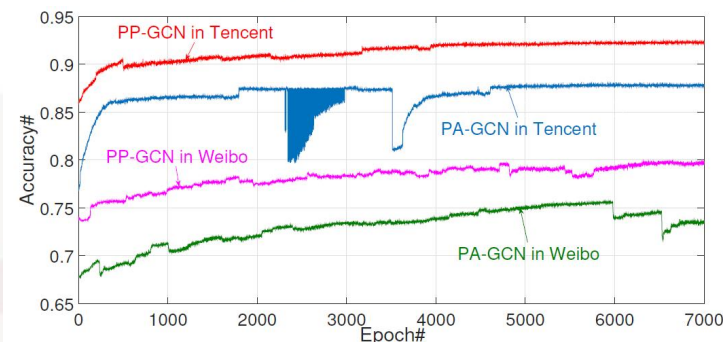


Figure 3: Illustration of the Accuracy for PP-GCN and PA-GCN.

Table 2: Accuracy and F1 results of event detection.

Algorithms	Tencent		Weibo	
	Accuracy	F1	Accuracy	F1
ARC-I	0.5384	0.4868	0.4910	0.4857
ARC-II	0.5437	0.3677	0.5277	0.5137
DUET	0.5625	0.5237	0.5397	0.5523
DSSM	0.5808	0.6468	0.5765	0.5411
C-DSSM	0.6017	0.4857	0.6170	0.5814
MV-LSTM	0.5562	0.6383	0.6252	0.6613
SVM	0.7581	0.7361	0.6511	0.6268
SE-GCN	0.7901	0.7895	0.7065	0.7015
PP-SE-GCN	0.8319	0.8383	0.7317	0.7384
PA-GCN	0.8818	0.8801	0.7567	0.7591
PP-GCN	0.9252	0.9231	0.8000	0.8134



模型2: 异构图注意力神经网络事件检测

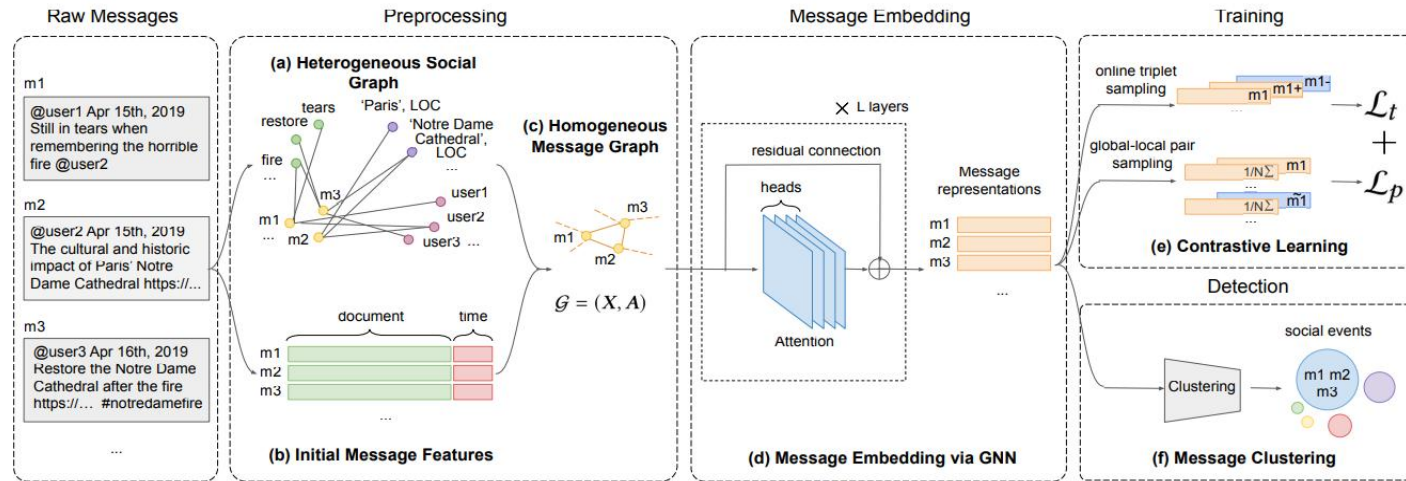
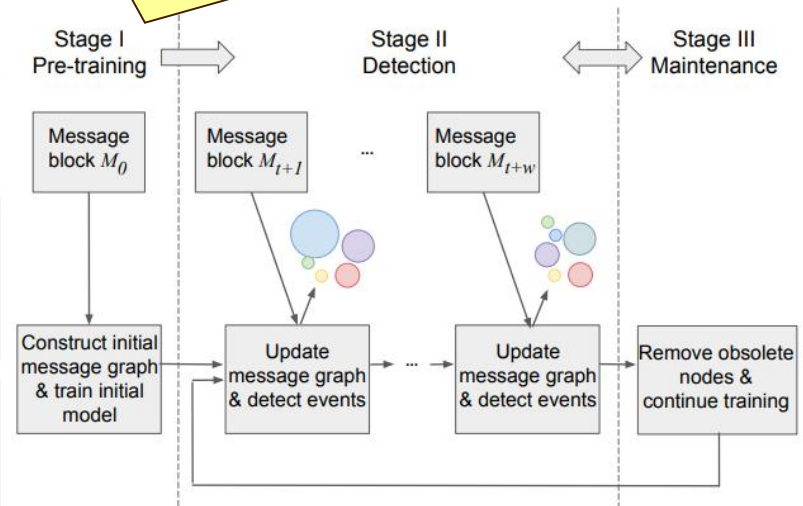


Figure 1: The architecture of the proposed KPGNN model (best viewed in color). (a) is a heterogeneous social graph that combines various types of elements contained in the raw messages. Different node colors denote different node types. (b) is the initial feature vectors of the messages. (c) is a homogeneous message graph that incorporates (a) and (b) (detailed in Section 3.2). (d) shows a GNN-based encoder that learns representations for the messages in (c). (e) calculates triplet loss $\mathcal{L}_t$ and global-local pair loss $\mathcal{L}_p$ through contrastive learning. In (e), two orange bars form a positive pair while one orange bar and one blue bar denote a negative pair. (f) clusters messages into social events.

An overview of Knowledge-Preserving Incremental Social Event Detection

Yuwei Cao, Hao Peng, etc, Knowledge-Preserving Incremental Social Event Detection via Heterogeneous GNNs. Web Conference 2021, 谷歌学术引用 54

全生命周期的神经网络增量社会事件检测架构



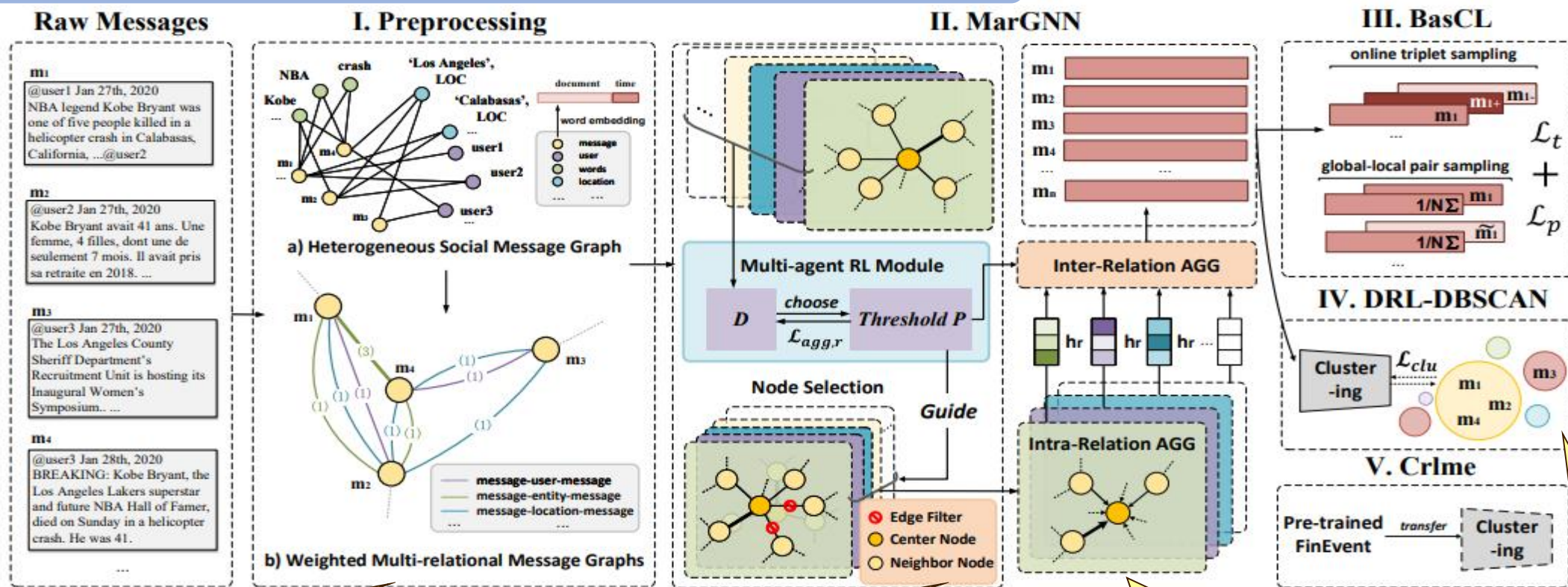
Blocks	M_1	M_2	M_3	M_4	M_5	M_6	M_7
Word2vec [27]	.19±.00	.50±.00	.39±.00	.34±.00	.41±.00	.53±.00	.25±.00
LDA [3]	.11±.00	.27±.01	.28±.00	.25±.00	.26±.00	.32±.00	.18±.01
WMD [19]	.32±.00	.71±.00	.67±.00	.50±.00	.61±.00	.61±.00	.46±.00
PP-GCN [29]	.23±.00	.57±.02	.55±.01	.46±.01	.48±.01	.57±.01	.37±.00
EventX [21]	.36±.00	.68±.00	.63±.00	.63±.00	.59±.00	.70±.00	.51±.00
KPGNN _t	.38±.01	.78±.01	.77±.00	.68±.01	.73±.01	.81±.00	.54±.01
KPGNN	.39±.00	.79±.01	.76±.00	.67±.00	.73±.01	.82±.01	.55±.01

Blocks	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}	M_{14}
Word2vec [27]	.46±.00	.35±.00	.51±.00	.37±.00	.30±.00	.37±.00	.36±.00
LDA [3]	.37±.01	.34±.00	.44±.01	.33±.01	.22±.01	.27±.00	.21±.00
WMD [19]	.67±.00	.55±.00	.61±.00	.50±.00	.00.60±.00	.54±.00	.66±.00
PP-GCN [29]	.55±.02	.51±.02	.55±.02	.50±.01	.45±.01	.47±.01	.44±.01
EventX [21]	.71±.00	.67±.00	.68±.00	.65±.00	.61±.00	.58±.00	.57±.00
KPGNN _t	.79±.01	.74±.01	.79±.01	.73±.00	.69±.01	.68±.01	.68±.01
KPGNN	.80±.00	.74±.02	.80±.01	.74±.01	.68±.01	.69±.01	.69±.00

Blocks	M_{15}	M_{16}	M_{17}	M_{18}	M_{19}	M_{20}	M_{21}
Word2vec [27]	.27±.00	.49±.00	.33±.00	.29±.00	.37±.00	.38±.00	.31±.00
LDA [3]	.21±.00	.35±.01	.19±.00	.18±.00	.29±.01	.35±.00	.19±.00
WMD [19]	.51±.00	.60±.00	.55±.00	.63±.00	.54±.00	.58±.00	.58±.00
PP-GCN [29]	.39±.01	.55±.01	.48±.00	.47±.01	.51±.02	.51±.01	.41±.02
EventX [21]	.49±.00	.62±.00	.58±.00	.59±.00	.60±.00	.67±.00	.53±.00
KPGNN _t	.57±.01	.78±.01	.69±.01	.68±.01	.73±.00	.73±.00	.59±.01
KPGNN	.58±.00	.79±.01	.70±.01	.68±.02	.73±.01	.72±.02	.60±.00



模型3: 增强的异构图事件表示与检测模型



带权多关系消息
图模型

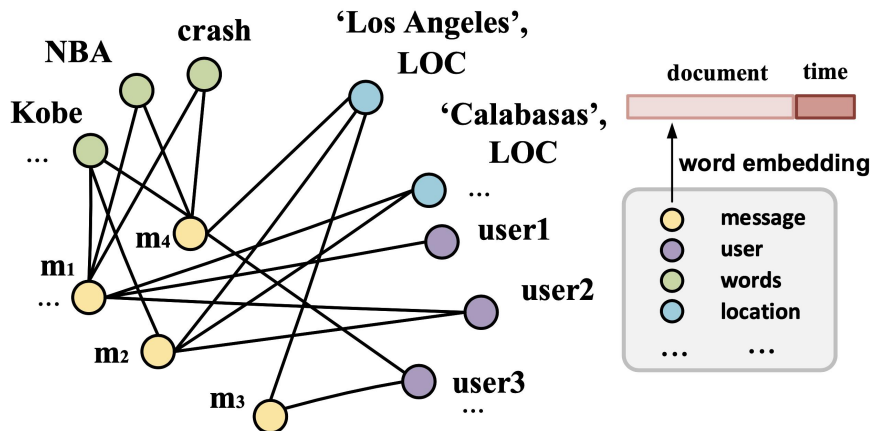
多智能体强化图神
经网络信息聚合

关系内及关系
间信息聚合

深度强化密度
聚类模型

模块 3.1：加权多关系信息图

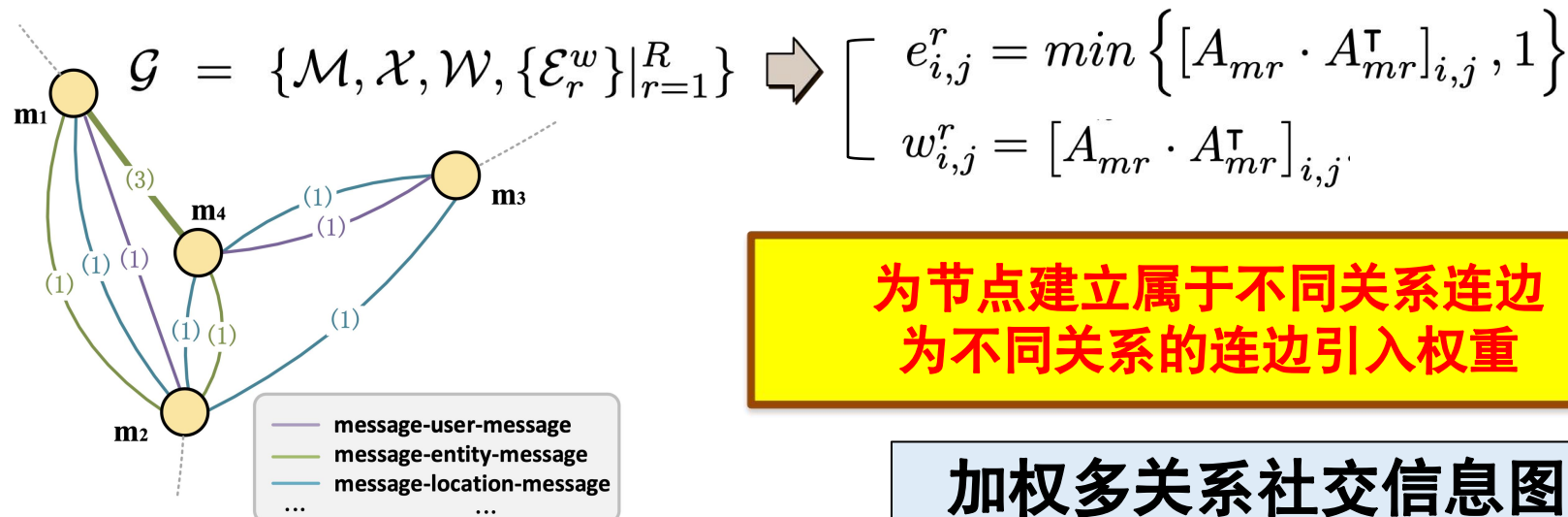
现状



异构社交信息图

瓶颈：主流的异构信息图在聚合过程中将异构化的关系同构化，保留连接信息较少

思路



为节点建立属于不同关系连边
为不同关系的连边引入权重

加权多关系社交信息图

m1

@user1 Jan 27th, 2020
NBA legend Kobe Bryant was one of five people killed in a helicopter crash in Calabasas, California, ...@user2

m2

@user2 Jan 27th, 2020
Kobe Bryant avait 41 ans. Une femme, 4 filles, dont une de seulement 7 mois. Il avait pris sa retraite en 2018. ...

m3

@user3 Jan 27th, 2020
The Los Angeles County Sheriff Department's Recruitment Unit is hosting its Inaugural Women's Symposium... ..

m4

@user3 Jan 28th, 2020
BREAKING: Kobe Bryant, the Los Angeles Lakers superstar and future NBA Hall of Famer, died on Sunday in a helicopter crash. He was 41.

模块 3.2.1：多智能体强化加权多关系图神经网络框架

强化的邻居选择

马尔可夫决策模型

Top-P 邻居采样

多智能体指导的多关系邻居选择

Actor-Critic

状态

$$s_{agg,r}^{(l)(k)} = \frac{1}{|N|} \sum_{i=1}^N \frac{\sum_{m_j \in \mathcal{N}_r^{(k)}(m_i)} w_{i,j}^{(r)} \cdot \mathcal{D}(\mathbf{h}_{m_i}^{(k)}, \mathbf{h}_{m_j}^{(k)})}{\sum_{m_j \in \mathcal{N}_r^{(k)}(m_i)} w_{i,j}^{(r)}}$$

动作

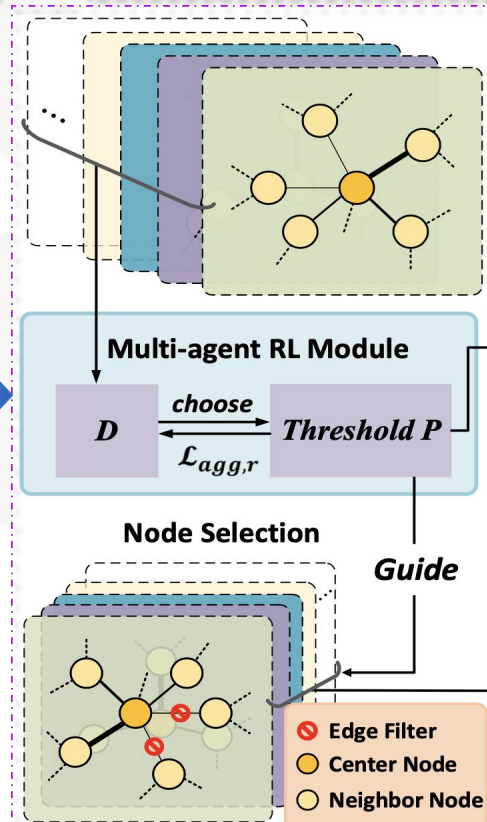
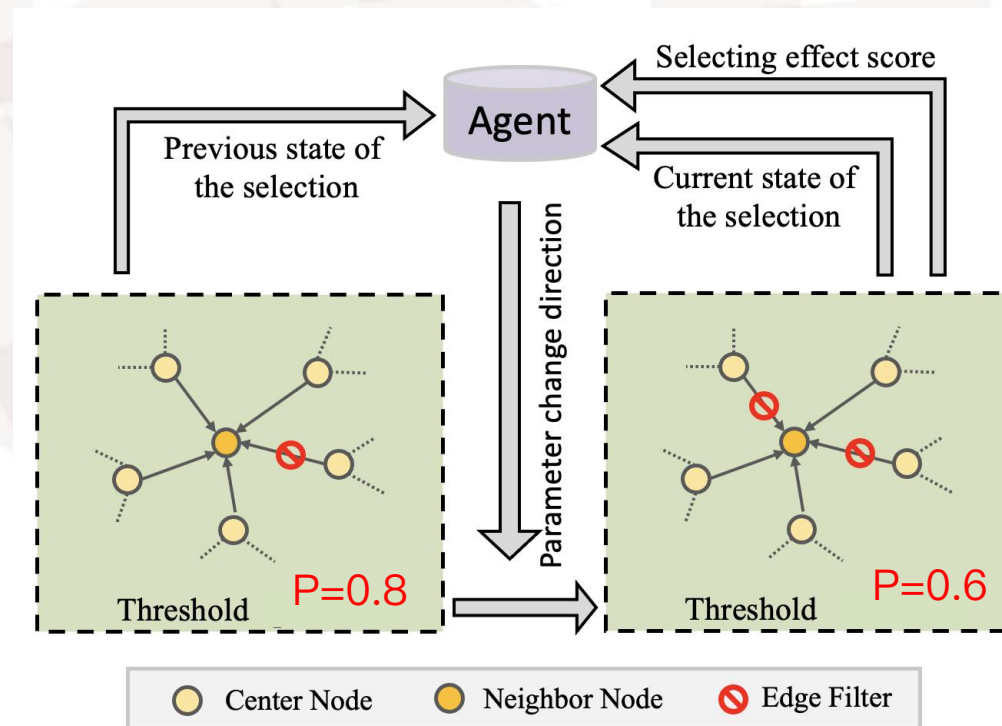
$$p_{agg,r}^{(k)} \in [0, 1]$$

奖励

$$r_{agg,r}^{(k)} = NMI_k(|E_{true}|)$$

优化

$$\mathcal{L}_{agg,r} = - \left[\left(r_{agg,r}^{(k)} + \gamma Q(s_{agg,r}^{(k+1)}, a_{agg,r}^{(k+1)}) - Q(s_{agg,r}^{(k)}, a_{agg,r}^{(k)}) \right) \cdot \log \left[\pi(a_{agg,r}^{(k)} | s_{agg,r}^{(k)}) \right] \right]$$



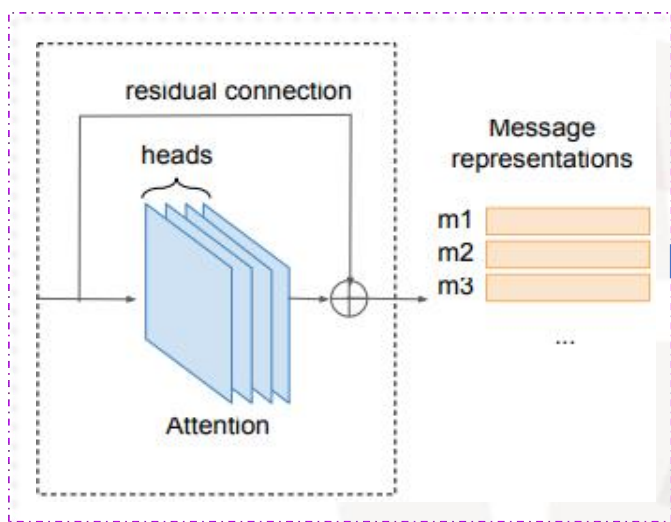
协作学习在不同关系之间找到平衡

缓解不同关系中杂质信息对聚合的影响



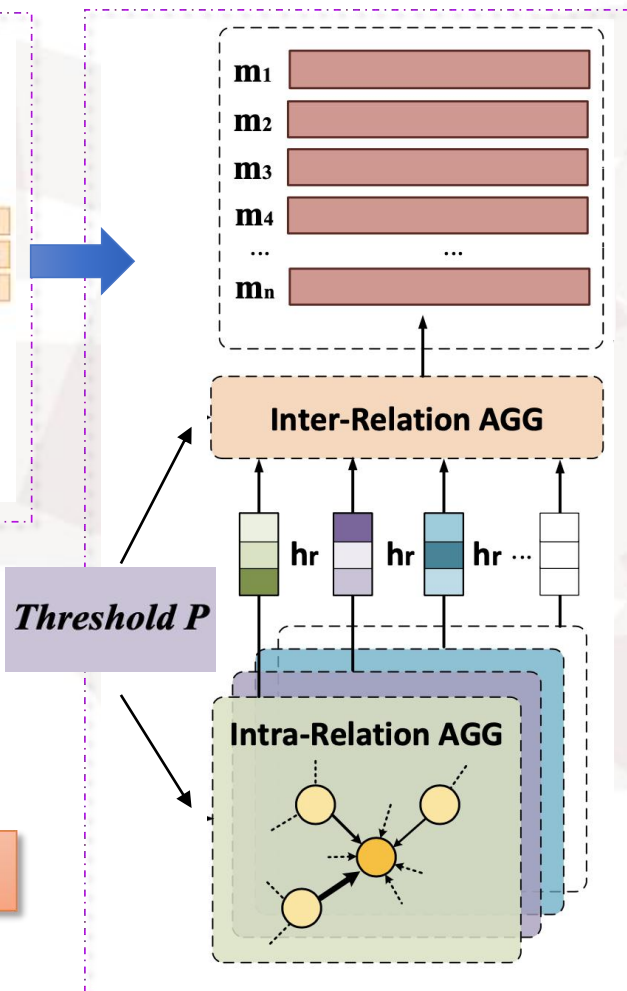
模块 3.2.2: 多智能体强化加权多关系图神经网络框架

关系感知的加权邻居聚合



An overview of
Knowledge-Preserving
Incremental Social Event
Detection via
Heterogeneous GNNs

单关系的图注意力机制



(1)

Intra-relation Aggregation

$$\mathbf{h}_{m_i, r}^{(l)} \leftarrow AGG_{intra, r}^{(l)} \left(\overset{\text{heads}}{\parallel} \left(\oplus \{ \mathbf{h}_{m_j}^{(l-1)} : m_j \in \mathcal{N}_r^l(m_i) \} \right) \right)$$

头部连接

累加聚合算子

(2)

Inter-relation Aggregation

$$\mathbf{h}_{m_i}^{(l)} \leftarrow \mathbf{h}_{m_i}^{(l-1)} \otimes AGG_{inter}^{(l)} \left(\otimes \{ p_r^{(l)} \cdot \mathbf{h}_{m_i, r}^{(l)} \}_{r=1}^R \right)$$

学习项

Threshold P

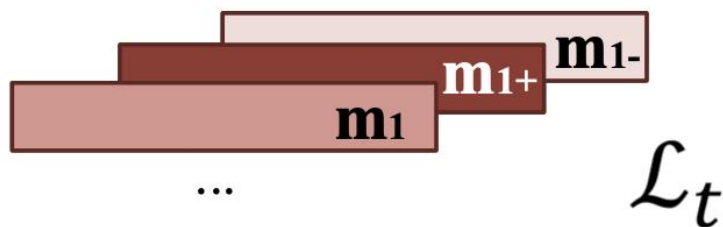


模块 3.3：基于平衡采样策略的对比学习机制

正样本对

负样本对

online triplet sampling



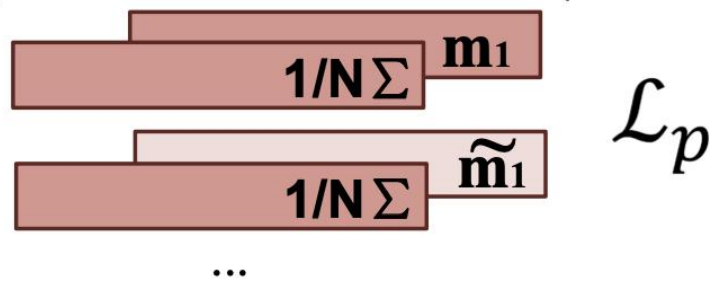
(1)

$$\mathcal{L}_t = \sum_{(m_i, m_{i+}, m_{i-}) \in T} \max\{\mathcal{D}(\mathbf{h}_{m_i}, \mathbf{h}_{m_{i+}}) - \mathcal{D}(\mathbf{h}_{m_i}, \mathbf{h}_{m_{i-}}) + a, 0\}$$

$$\mathcal{L}_{BasCL} = \mathcal{L}_t + \mathcal{L}_p$$

总体损失函数

global-local pair sampling



(2)

$$\mathcal{L}_p = \frac{1}{N} \sum_{i=1}^N \left(\log \mathcal{S}(\mathbf{h}_{m_i}, \mathbf{s}) + \log (1 - \mathcal{S}(\tilde{\mathbf{h}}_{m_i}, \mathbf{s})) \right)$$

噪声对比

不局限于固定数量的类

解决现实世界中社会事件的长尾问题

模块 3.4：深度强化学习指导的DBSCAN模型

马尔可夫决策模型

深度强化学习指导的 DBSCAN 聚类参数选择

TD3

状态

$$s_{clu}^{(\tau)} = \{\epsilon^{(\tau)}, |E|^{(\tau)}, coh^{(\tau)}, sep^{(\tau)}\}$$

动作

$$\epsilon^{(\tau)}$$

MinPts 固定

奖励

$$r_{clu}^{(\tau)} = \begin{cases} \frac{SS_B^{(\tau)}}{|E|^{(\tau)} - 1} / \frac{SS_W^{(\tau)}}{N - |E|^{(\tau)}}, & |E|^{(\tau)} \text{ in bounds.} \\ 0, & |E|^{(\tau)} \text{ out of bounds.} \end{cases}$$

轮廓系数 (无监督)

优化

$$\mathcal{L}_{clu} = -\mathbb{E} \left[\frac{1}{2} \left(r_{clu}^{(\tau)} + \gamma \min \hat{Q}(s_{clu}^{(\tau+1)}, a_{clu}^{(\tau+1)}) - Q(s_{clu}^{(\tau)}, a_{clu}^{(\tau)}) \right)^2 \right]$$

Cluster-ing

$\mathcal{L}_{clu}$

m₁
m₂
m₃
m₄

Previous state of the cluster

Agent

Clustering effect score

Current state of the cluster

Parameter change direction

Eps=0.2, MinPts=3

Eps=0.1, MinPts=3

○ Data Object ↔ Directly Reachable ○ Eps Radius

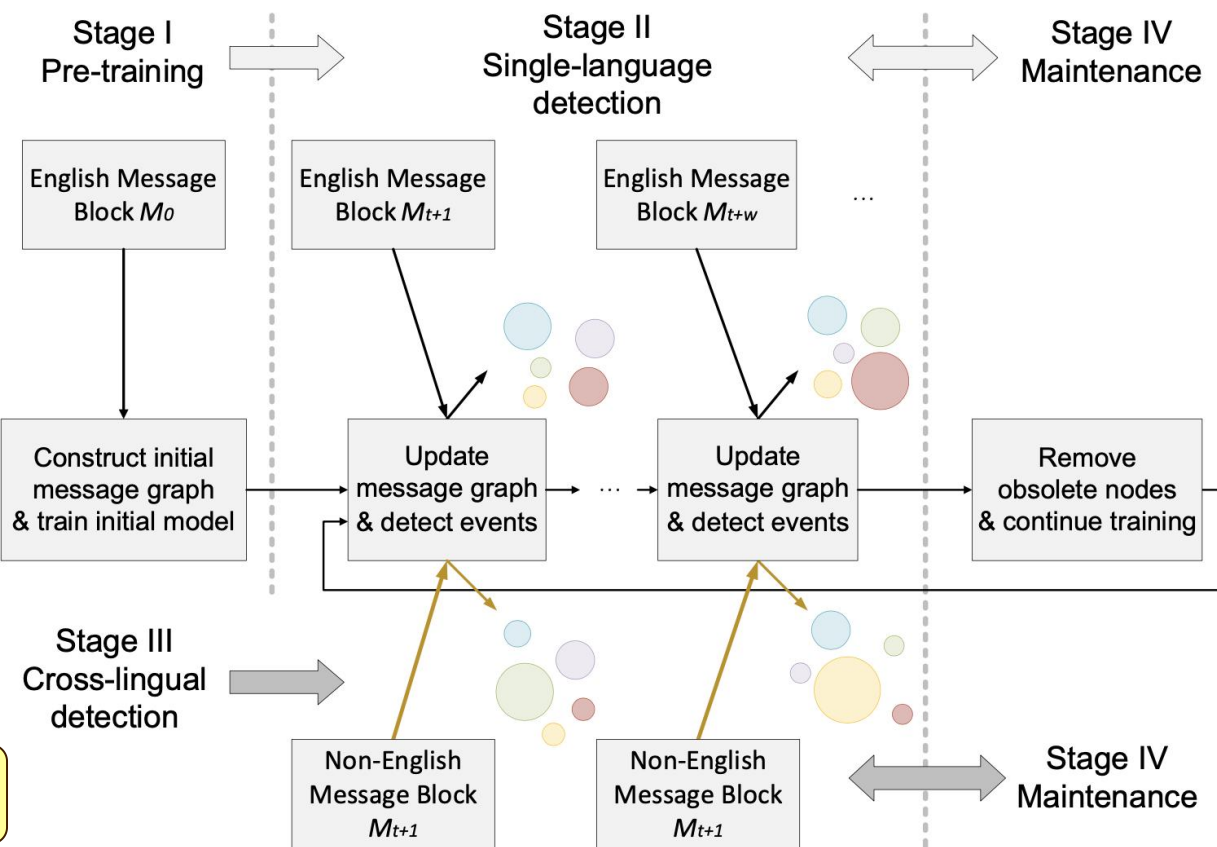
流式场景自适应的聚类参数选择

无需依赖样本的真实标签和事件类的数量



模块 3.5：跨语言的社会信息嵌入方法及生命周期机制

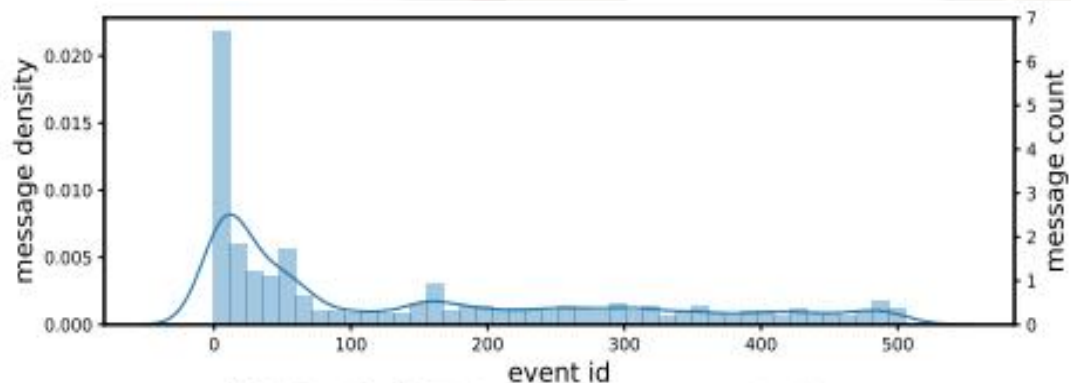
Definition 2.7. Given a social event detection algorithm $f_E(M_i; \theta) = E_i$, $M_i \in S_E$ in English social stream, **cross-lingual social event detection** algorithm learn a upgraded model $f_{NoE}(M_j; \bar{\theta}) = E_j$ from $f_E(M_i; \theta)$, where $E_j = \{e_k | 1 \leq k \leq |E_j|\}$ is also a set of events contained in no-English message block $M_j \in S_{NoE}$. Here, $\bar{\theta}$ denotes the parameter of the upgraded model f_{NoE} from f_E .



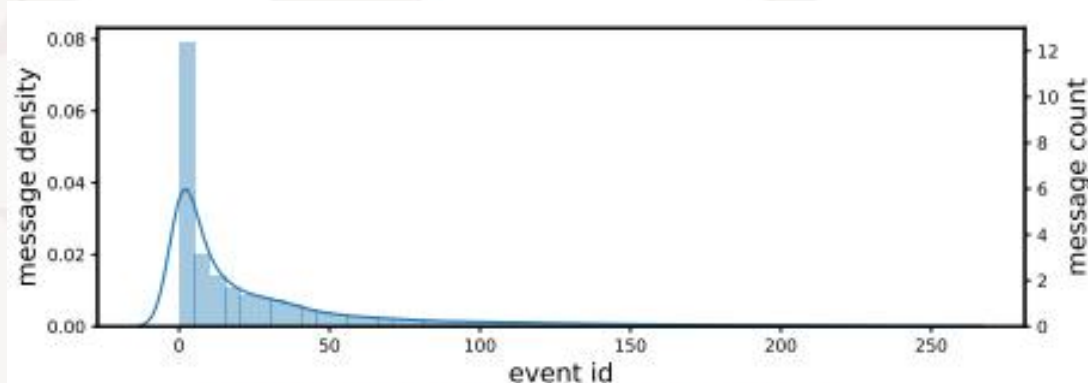
实验分析1：数据统计分析及静态事件检测实验结果

挑战1：长尾分布

挑战2：语言模态



(a) English Twitter Dataset statistics.



(b) French Twitter Dataset statistics.

TABLE 2: Offline evaluation results on the Twitter dataset. The best results are marked in **bold**.

Metrics	Word2vec	LDA	WMD	BERT	BiLSTM	PP-GCN	EventX	KPGNN _t	KPGNN	FinEvent _k	FinEvent _d
NMI	.44±.00	.29±.00	.65±.00	.64±.00	.63±.00	.68±.02	.72±.00	.69±.01	.70±.01	.79±.01	.80±.01 ↑ .08
AMI	.13±.00	.04±.00	.50±.00	.44±.00	.41±.00	.50±.02	.19±.00	.51±.00	.52±.01	.62±.01	.69±.01 ↑ .19
ARI	.02±.00	.01±.00	.06±.00	.07±.00	.17±.00	.20±.01	.05±.00	.21±.01	.22±.01	.24±.01	.48±.01 ↑ .28

实验分析 2.1：总体增量评估

NMI 提升 8%- 136%

TABLE 3: Incremental evaluation NMIs. The best results are marked in **bold** and second-best in *italic*.

Blocks	Word2vec	LDA	WMD	BERT	BiLSTM	PP-GCN	EventX	KPGNN _t	KPGNN	FinEvent _k	FinEvent _d
M ₁	.19±.00	.11±.00	.32±.00	.36±.00	.24±.00	.23±.00	.36±.00	.38±.01	.39±.00	.38±.01	.84±.01 ↑ .49
M ₂	.50±.00	.27±.01	.71±.00	.78±.00	.50±.00	.57±.02	.68±.00	.78±.01	.79±.01	.81±.00	.84±.01 ↑ .06
M ₃	.39±.00	.28±.00	.67±.00	.75±.00	.39±.00	.55±.01	.63±.00	.77±.00	.76±.00	.83±.00	.89±.01 ↑ .14
M ₄	.34±.00	.25±.00	.50±.00	.60±.00	.40±.00	.46±.01	.63±.00	.68±.01	.67±.00	.71±.01	.71±.01 ↑ .08
M ₅	.41±.00	.26±.00	.61±.00	.72±.00	.41±.00	.48±.01	.59±.00	.73±.01	.73±.01	.76±.00	.83±.01 ↑ .11
M ₆	.53±.00	.32±.00	.61±.00	.78±.00	.50±.00	.57±.01	.70±.00	.81±.00	.82±.01	.84±.00	.83±.01 ↑ .06
M ₇	.25±.00	.18±.01	.46±.00	.54±.00	.33±.00	.37±.00	.51±.00	.54±.01	.55±.01	.56±.00	.73±.01 ↑ .02
M ₈	.46±.00	.37±.01	.67±.00	.79±.00	.49±.00	.55±.02	.71±.00	.79±.01	.80±.00	.87±.01	.87±.01 ↑ .08
M ₉	.35±.00	.34±.00	.55±.00	.70±.00	.43±.00	.51±.02	.67±.00	.74±.01	.74±.02	.78±.02	.79±.01 ↑ .09
M ₁₀	.51±.00	.44±.01	.61±.00	.74±.00	.50±.00	.55±.02	.68±.00	.79±.01	.80±.01	.81±.01	.82±.01 ↑ .08
M ₁₁	.37±.00	.33±.01	.50±.00	.68±.00	.49±.00	.50±.01	.65±.00	.73±.00	.74±.01	.76±.00	.75±.01 ↑ .08
M ₁₂	.30±.00	.22±.01	.60±.00	.59±.00	.39±.00	.45±.01	.61±.00	.69±.01	.68±.01	.76±.01	.67±.01 ↑ .15
M ₁₃	.37±.00	.27±.00	.54±.00	.63±.00	.46±.00	.47±.01	.58±.00	.68±.01	.69±.01	.67±.00	.79±.01 ↑ .18
M ₁₄	.36±.00	.21±.00	.66±.00	.64±.00	.44±.00	.44±.01	.57±.00	.68±.01	.69±.00	.74±.00	.82±.01 ↑ .16
M ₁₅	.27±.00	.21±.00	.51±.00	.54±.00	.40±.00	.39±.01	.49±.00	.57±.01	.58±.00	.64±.00	.69±.01 ↑ .15
M ₁₆	.49±.00	.35±.01	.60±.00	.75±.00	.53±.00	.55±.01	.62±.00	.78±.01	.79±.01	.80±.00	.90±.01 ↑ .15
M ₁₇	.33±.00	.19±.00	.55±.00	.63±.00	.45±.00	.48±.00	.58±.00	.69±.01	.70±.01	.73±.00	.83±.01 ↑ .20
M ₁₈	.29±.00	.18±.00	.63±.00	.57±.00	.44±.00	.47±.01	.59±.00	.68±.01	.68±.02	.72±.01	.74±.01 ↑ .11
M ₁₉	.37±.00	.29±.01	.54±.00	.66±.00	.44±.00	.51±.02	.60±.00	.73±.00	.73±.01	.76±.02	.66±.01 ↑ .10
M ₂₀	.38±.00	.35±.00	.58±.00	.68±.00	.48±.00	.51±.01	.67±.00	.73±.00	.72±.02	.73±.00	.80±.01 ↑ .12
M ₂₁	.31±.00	.19±.00	.58±.00	.59±.00	.41±.00	.41±.02	.53±.00	.59±.01	.60±.00	.65±.01	.74±.01 ↑ .15

实验分析 2.2：总体增量评估

AMI 提升 11%- 147%

TABLE 4: Incremental evaluation AMIs. The best results are marked in **bold** and second-best in *italic*.

Blocks	Word2vec	LDA	WMD	BERT	BiLSTM	PP-GCN	EventX	KPGNN _t	KPGNN	FinEvent _k	FinEvent _d
M_1	.08±.00	.08±.00	.30±.00	.34±.00	.12±.00	.21±.00	.06±.00	.36±.01	.37±.00	.36±.01	.84±.00 ↑.50
M_2	.41±.00	.20±.01	.69±.00	.76±.00	.41±.00	.55±.02	.29±.02	.77±.01	.78±.01	.77±.00	.84±.00 ↑.08
M_3	.31±.00	.22±.01	.63±.00	.73±.00	.31±.00	.52±.01	.18±.01	.75±.00	.74±.00	.82±.01	.89±.00 ↑.16
M_4	.24±.00	.17±.00	.45±.00	.55±.00	.30±.00	.42±.01	.19±.01	.65±.01	.64±.01	.67±.02	.69±.00 ↑.14
M_5	.33±.00	.21±.00	.57±.00	.71±.00	.33±.00	.46±.01	.14±.00	.71±.01	.71±.01	.74±.00	.82±.00 ↑.11
M_6	.40±.00	.20±.00	.57±.00	.74±.00	.36±.00	.52±.02	.27±.00	.78±.00	.79±.01	.81±.00	.82±.00 ↑.08
M_7	.13±.00	.12±.01	.46±.00	.50±.00	.20±.00	.34±.00	.13±.00	.50±.01	.51±.01	.53±.00	.72±.00 ↑.22
M_8	.33±.00	.24±.01	.63±.00	.75±.00	.35±.00	.49±.02	.21±.00	.75±.01	.76±.01	.84±.01	.87±.00 ↑.12
M_9	.24±.00	.24±.00	.46±.00	.66±.00	.32±.00	.46±.02	.19±.00	.70±.01	.71±.02	.75±.00	.78±.00 ↑.12
M_{10}	.39±.00	.36±.01	.57±.00	.70±.00	.39±.00	.51±.02	.24±.00	.76±.01	.78±.01	.78±.00	.81±.00 ↑.11
M_{11}	.26±.00	.25±.01	.42±.00	.65±.00	.37±.00	.46±.01	.24±.00	.70±.00	.71±.01	.73±.00	.74±.00 ↑.09
M_{12}	.23±.00	.16±.01	.58±.00	.56±.00	.32±.00	.42±.01	.16±.00	.66±.01	.66±.01	.75±.01	.67±.00 ↑.17
M_{13}	.23±.00	.19±.00	.50±.00	.59±.00	.31±.00	.43±.01	.16±.00	.65±.01	.67±.01	.64±.00	.79±.00 ↑.20
M_{14}	.26±.00	.15±.00	.64±.00	.61±.00	.34±.00	.41±.01	.14±.00	.65±.01	.65±.00	.72±.00	.82±.00 ↑.18
M_{15}	.15±.00	.13±.00	.47±.00	.50±.00	.26±.00	.35±.01	.07±.00	.53±.01	.54±.00	.61±.00	.67±.00 ↑.17
M_{16}	.36±.00	.27±.01	.59±.00	.72±.00	.41±.00	.52±.01	.19±.00	.75±.01	.77±.01	.75±.01	.90±.00 ↑.20
M_{17}	.24±.00	.13±.00	.57±.00	.60±.00	.35±.00	.45±.00	.18±.00	.67±.01	.68±.01	.71±.02	.82±.00 ↑.22
M_{18}	.21±.00	.12±.00	.60±.00	.53±.00	.35±.00	.45±.01	.16±.00	.66±.01	.66±.02	.70±.00	.74±.00 ↑.14
M_{19}	.28±.00	.22±.01	.49±.00	.63±.00	.35±.00	.48±.02	.16±.00	.70±.00	.71±.01	.75±.01	.66±.00 ↑.12
M_{20}	.24±.00	.23±.00	.55±.00	.62±.00	.34±.00	.45±.02	.18±.00	.68±.00	.68±.02	.68±.00	.78±.00 ↑.16
M_{21}	.21±.00	.13±.00	.52±.00	.57±.00	.31±.00	.38±.02	.10±.00	.57±.01	.57±.00	.63±.01	.64±.00 ↑.07

实验分析 2.3：总体增量评估

ARI 提升 24%- 170%

TABLE 5: Incremental evaluation ARIs. The best results are marked in **bold** and second-best in *italic*.

Blocks	Word2vec	LDA	WMD	BERT	BiLSTM	PP-GCN	EventX	KPGNN _t	KPGNN	FinEvent _k	FinEvent _d
M ₁	.01±.00	.01±.00	.04±.00	.03±.00	.03±.00	.05±.00	.01±.00	.06±.01	.07±.01	.05±.00	.90±.00 ↑.85
M ₂	.49±.00	.08±.00	.48±.00	.64±.00	.49±.00	.67±.03	.45±.02	.76±.01	.76±.02	.67±.01	.90±.00 ↑.23
M ₃	.16±.00	.02±.01	.28±.00	.43±.00	.17±.00	.47±.01	.09±.01	.60±.02	.58±.01	.58±.00	.89±.00 ↑.42
M ₄	.07±.00	.07±.00	.11±.00	.19±.00	.11±.00	.24±.01	.07±.01	.30±.01	.29±.01	.27±.02	.27±.00 ↑.06
M ₅	.17±.00	.06±.00	.26±.00	.44±.00	.19±.00	.34±.00	.04±.00	.48±.01	.47±.03	.43±.01	.63±.00 ↑.19
M ₆	.25±.00	.07±.01	.16±.00	.44±.00	.18±.00	.55±.03	.14±.00	.67±.05	.72±.03	.65±.00	.74±.00 ↑.19
M ₇	.02±.00	.01±.00	.08±.00	.07±.00	.08±.00	.11±.02	.02±.00	.11±.01	.12±.00	.09±.01	.45±.00 ↑.34
M ₈	.17±.00	.03±.00	.22±.00	.50±.00	.08±.00	.43±.04	.09±.00	.59±.02	.60±.01	.65±.02	.72±.00 ↑.22
M ₉	.08±.00	.03±.01	.12±.00	.33±.00	.27±.00	.31±.02	.07±.00	.45±.02	.46±.02	.43±.00	.68±.00 ↑.35
M ₁₀	.23±.00	.09±.02	.20±.00	.44±.00	.22±.00	.50±.07	.13±.00	.64±.01	.70±.06	.62±.02	.74±.00 ↑.24
M ₁₁	.09±.00	.03±.01	.12±.00	.27±.00	.17±.00	.38±.02	.16±.00	.48±.01	.49±.03	.42±.01	.60±.00 ↑.22
M ₁₂	.09±.00	.02±.01	.27±.00	.31±.00	.13±.00	.34±.03	.07±.00	.50±.03	.48±.01	.44±.00	.26±.00 ↑.16
M ₁₃	.06±.00	.01±.00	.13±.00	.14±.00	.13±.00	.19±.01	.04±.00	.28±.01	.29±.03	.21±.02	.75±.00 ↑.56
M ₁₄	.10±.00	.02±.00	.33±.00	.30±.00	.16±.00	.29±.01	.10±.00	.43±.02	.42±.02	.43±.01	.81±.00 ↑.48
M ₁₅	.09±.00	.01±.00	.16±.00	.10±.00	.14±.00	.15±.00	.01±.00	.16±.02	.17±.00	.16±.00	.46±.00 ↑.31
M ₁₆	.10±.00	.11±.01	.32±.00	.41±.00	.10±.00	.51±.03	.08±.00	.62±.03	.66±.05	.56±.01	.88±.00 ↑.37
M ₁₇	.06±.00	.02±.00	.26±.00	.24±.00	.17±.00	.35±.03	.12±.00	.41±.03	.43±.05	.36±.01	.81±.00 ↑.46
M ₁₈	.21±.00	.02±.00	.35±.00	.24±.00	.19±.00	.39±.03	.08±.00	.46±.02	.47±.04	.44±.01	.52±.00 ↑.13
M ₁₉	.28±.00	.03±.00	.12±.00	.32±.00	.16±.00	.41±.02	.07±.00	.50±.01	.51±.03	.44±.00	.35±.00 ↑.10
M ₂₀	.24±.00	.02±.01	.19±.00	.33±.00	.20±.00	.41±.01	.11±.00	.51±.01	.51±.04	.43±.02	.71±.00 ↑.30
M ₂₁	.21±.00	.01±.01	.19±.00	.18±.00	.16±.00	.20±.03	.01±.00	.23±.02	.20±.01	.23±.00	.48±.00 ↑.27

实验分析 3：聚类可视化

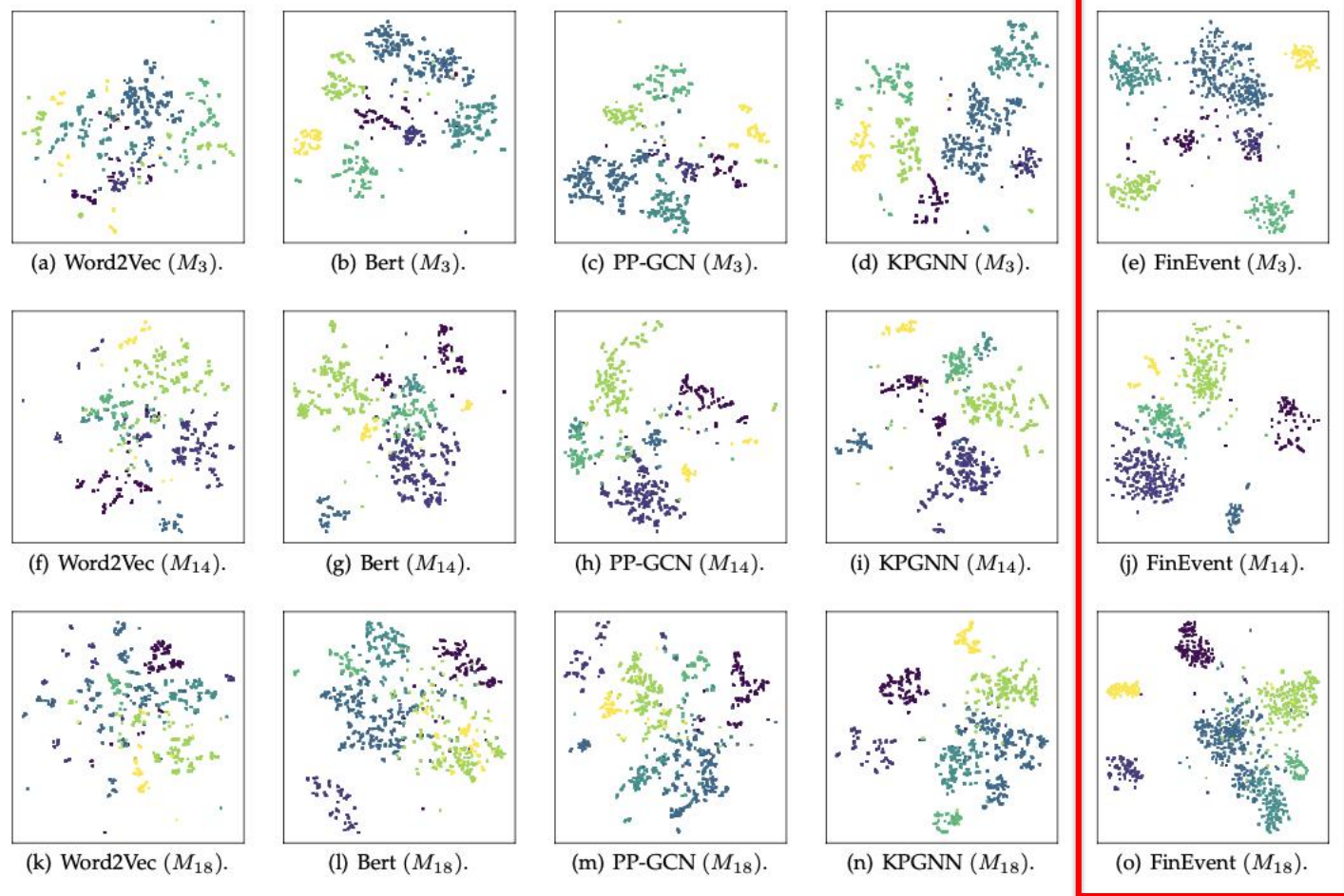


Fig. 4: Cluster visualization of message representations in the detection stage.

实验分析 4：消融对比实验

TABLE 6: Ablation study for neighbor sampler strategy, Intra-relation Aggregation AGG_{intra} and Inter-relation Aggregation AGG_{inter} . The best results are marked in **bold** and second-best in *italic*.

	Neighbor Sampling Strategy			Cluster Type	Intra-relation Aggregator		Threshold	Inter-relation Aggregator			Avg. Metrics		
	Random	Constant	Reinforced		Shared-GNN	RNN		Cat.	Sum	MLP	NMI	AMI	ARI
1	-	-	✓	DBSCAN	-	-	✓	✓	-	-	0.788	0.777	0.645
2	-	-	✓	K-Means	-	-	✓	✓	-	-	<i>0.727</i>	<i>0.702</i>	<i>0.443</i>
3	-	✓	-	K-Means	-	-	-	-	-	-	0.698	0.671	<i>0.448</i>
4	-	top 50%	-	K-Means	-	-	-	✓	-	-	0.719	0.699	0.441
5	-	bottom 50%	-	K-Means	-	-	-	✓	-	-	0.698	0.670	0.427
6	✓	-	-	K-Means	-	-	-	✓	-	-	0.718	0.694	0.437
7	-	✓	-	K-Means	-	-	-	✓	-	-	<i>0.722</i>	<i>0.696</i>	<i>0.442</i>
8	-	-	✓	K-Means	✓	-	✓	✓	-	-	0.700	0.673	0.416
9	-	-	-	K-Means	-	✓	-	-	-	-	0.449	0.324	0.168
10	-	-	✓	K-Means	-	-	-	✓	-	-	0.723	0.697	0.438
11	-	-	✓	K-Means	-	-	✓	-	✓	-	0.702	0.674	0.425
12	-	-	✓	K-Means	-	-	-	-	✓	-	0.700	0.672	0.422
13	-	-	✓	K-Means	-	-	✓	✓	-	✓	0.653	0.620	0.383
14	-	-	✓	K-Means	-	-	-	✓	-	✓	0.645	0.610	0.368



实验分析 5：多智能体强化学习进程分析——邻居采样

稳定收敛：动态纳什均衡

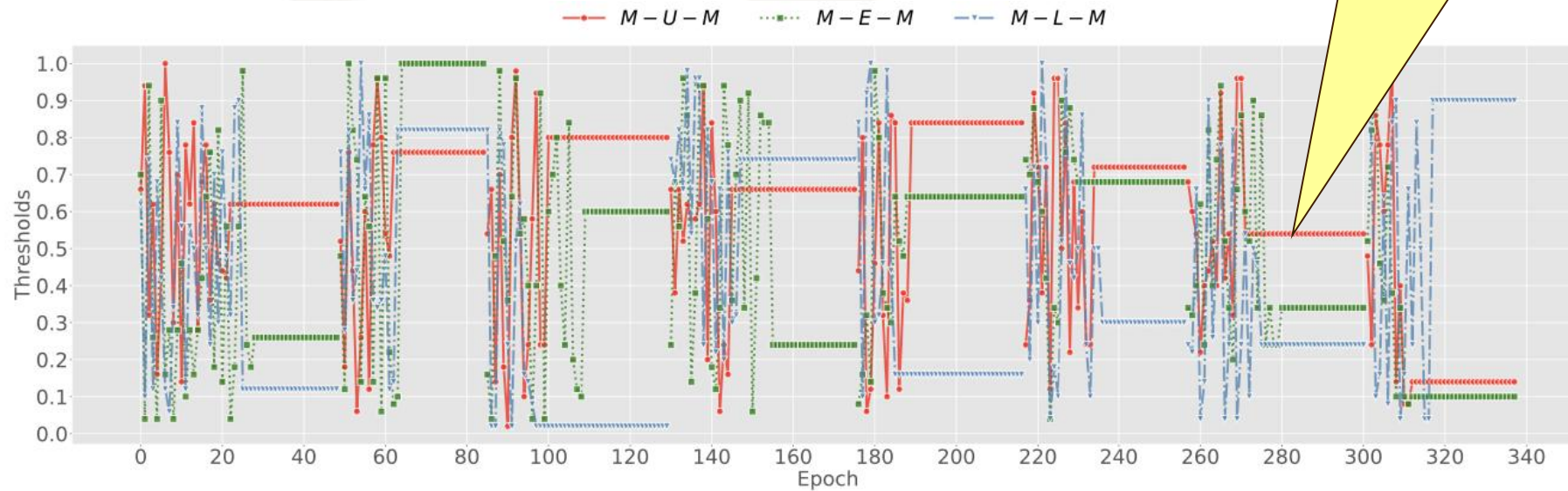
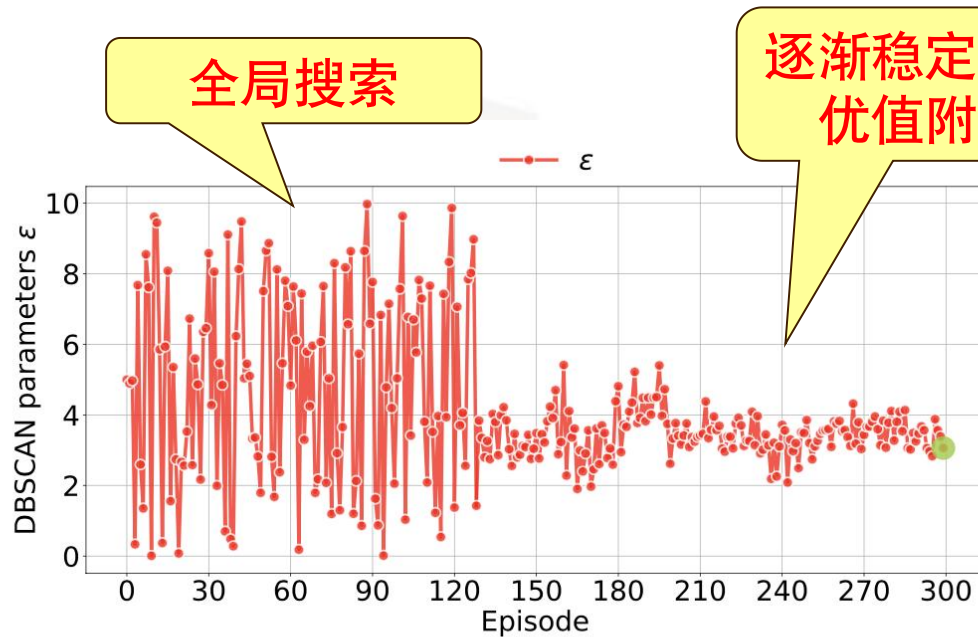


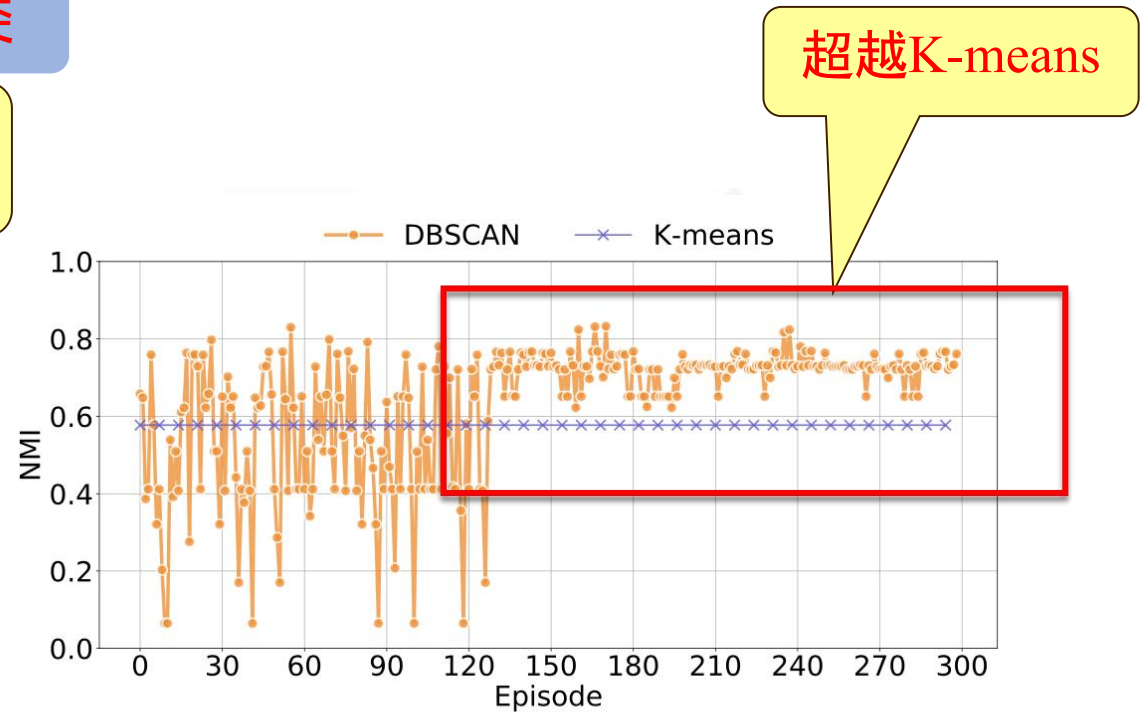
Fig. 5: Multi-agent reinforcement learning process in the online maintenance stage. We summarize the epochs of all time periods. Note that each process from fluctuation to stability is a pre-training or maintenance stage. The figure contains a total of one pre-training process and seven maintenance processes.

包含 1 个预训练阶段(从 epoch 0 到 epoch 50)和 7 个维护阶段(从 epoch 50 到 epoch 340)

实验分析 6：深度强化学习进程分析——聚类



(a) DBSCAN parameters.



(b) NMI scores.

Fig. 6: Deep reinforcement learning process in the online detection stage. We show the DRL-DBSCAN parameter adjustment and NMI change process of block M_7 as an example of DRL-DBSCAN, where the green marked points represents the final convergence parameter.

实验分析 7：增量检测中的强化学习及跨语言实验分析

TABLE 7: Preserving thresholds in the detection stage.

Blocks	M_0	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
$M-U-M$	–	.88	.24	.98	.24	.50	.40	.40	.12	.06	.96
$M-E-M$	–	.82	.56	.22	.20	.88	.90	.74	.20	.28	.50
$M-L-M$	–	.96	.08	.80	.54	.42	.78	.80	.56	.70	.86

Blocks	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}	M_{16}	M_{17}	M_{18}	M_{19}	M_{20}	M_{21}
$M-U-M$	.22	.94	.22	.14	.10	.46	.44	.24	.10	.32	.38
$M-E-M$	.66	.72	.24	.60	.76	.82	.90	.90	.20	.10	.34
$M-L-M$	.74	.64	.20	.98	.54	.18	.50	.46	.24	.16	.92

TABLE 8: DBSCAN parameters in the detection stage.

Blocks	M_0	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
ϵ	–	3.87	3.29	2.57	3.25	3.24	3.70	2.35	2.95	3.39	3.57

Blocks	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}	M_{16}	M_{17}	M_{18}	M_{19}	M_{20}	M_{21}
ϵ	3.19	3.28	3.76	3.93	2.00	3.37	3.54	3.23	4.00	2.23	3.18

TABLE 9: Cross-lingual transferring evaluation on French dataset. The best results are marked in bold.

Blocks	M_0	M_1	M_2	M_3	M_4	M_5	M_6	M_7
FinEvent_r	–	.575	.620	.577	.512	.556	.488	.584
FinEvent_{cr}	–	.576	.578	.580	.534	.633	.656	.615

Blocks	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}
FinEvent_g	–	.592	.574	.591	.474	.568	.511	.580
FinEvent_{cg}	–	.578	.575	.604	.542	.641	.659	.600

Blocks	M_8	M_9	M_{10}	M_{11}	M_{12}	M_{13}	M_{14}	M_{15}
FinEvent_r	.640	.484	.627	.529	.545	.472	.519	.586
FinEvent_{cr}	.612	.523	.597	.610	.616	.569	.622	.630
FinEvent_g	.625	.484	.623	.546	.548	.474	.530	.591
FinEvent_{cg}	.598	.512	.599	.584	.610	.563	.640	.626

语义映射+法语有标注+重训练

语义映射+法语无标注+参数利用

谷歌翻译+法语有标注+重训练

谷歌翻译+法语无标注+参数利用

模型4:递归式深度强化参数自学习的高效密度聚类 (DRL-DBSCAN)

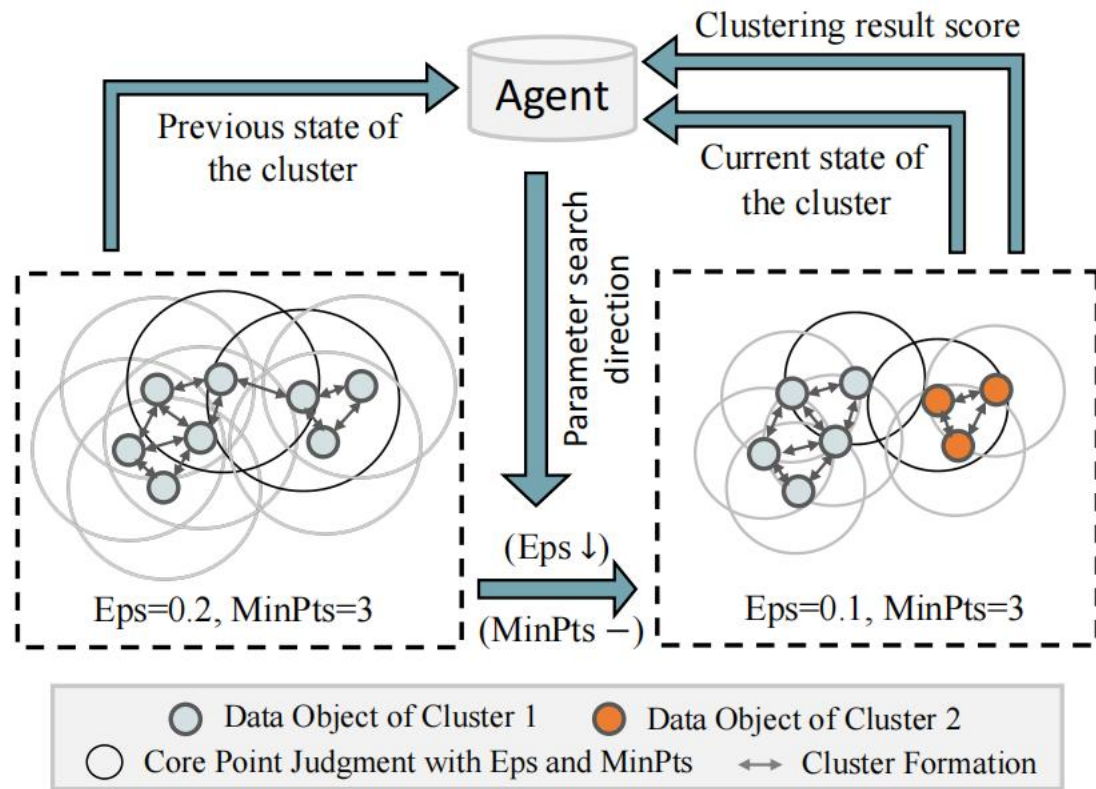
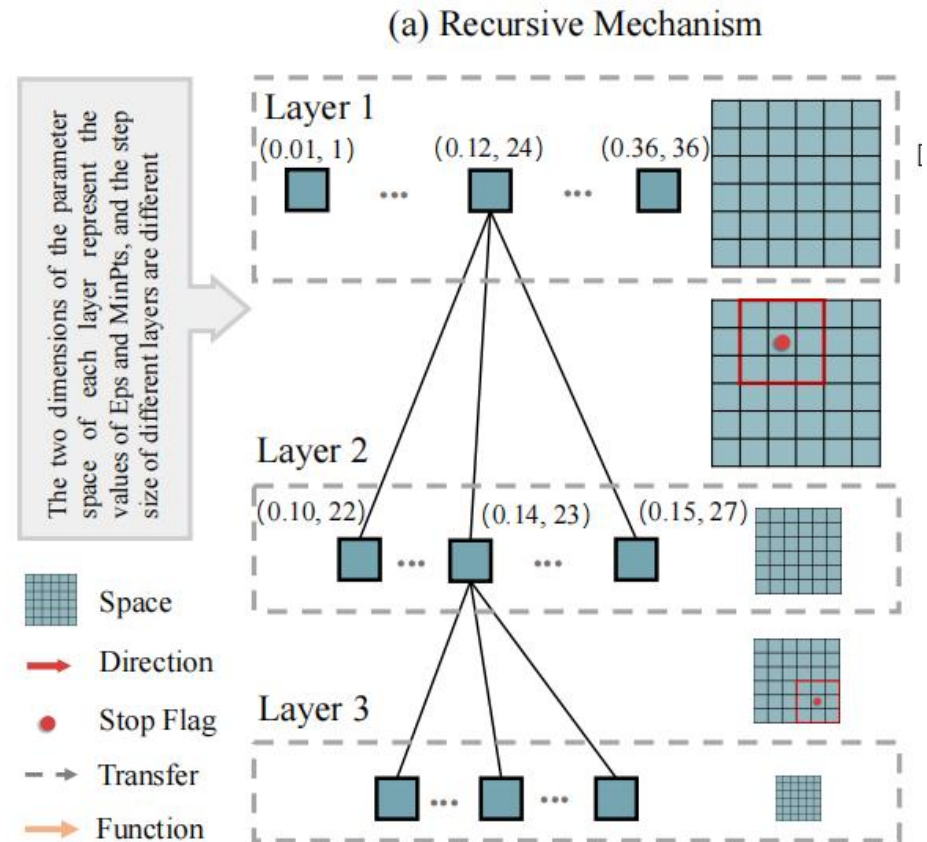


Figure 1: Markov process of parameter search. The agent uses the data as the environment, and determines the action to search by observing the clustering state and reward.



模型4:递归式深度强化参数自学习的高效密度聚类 (DRL-DBSCAN)

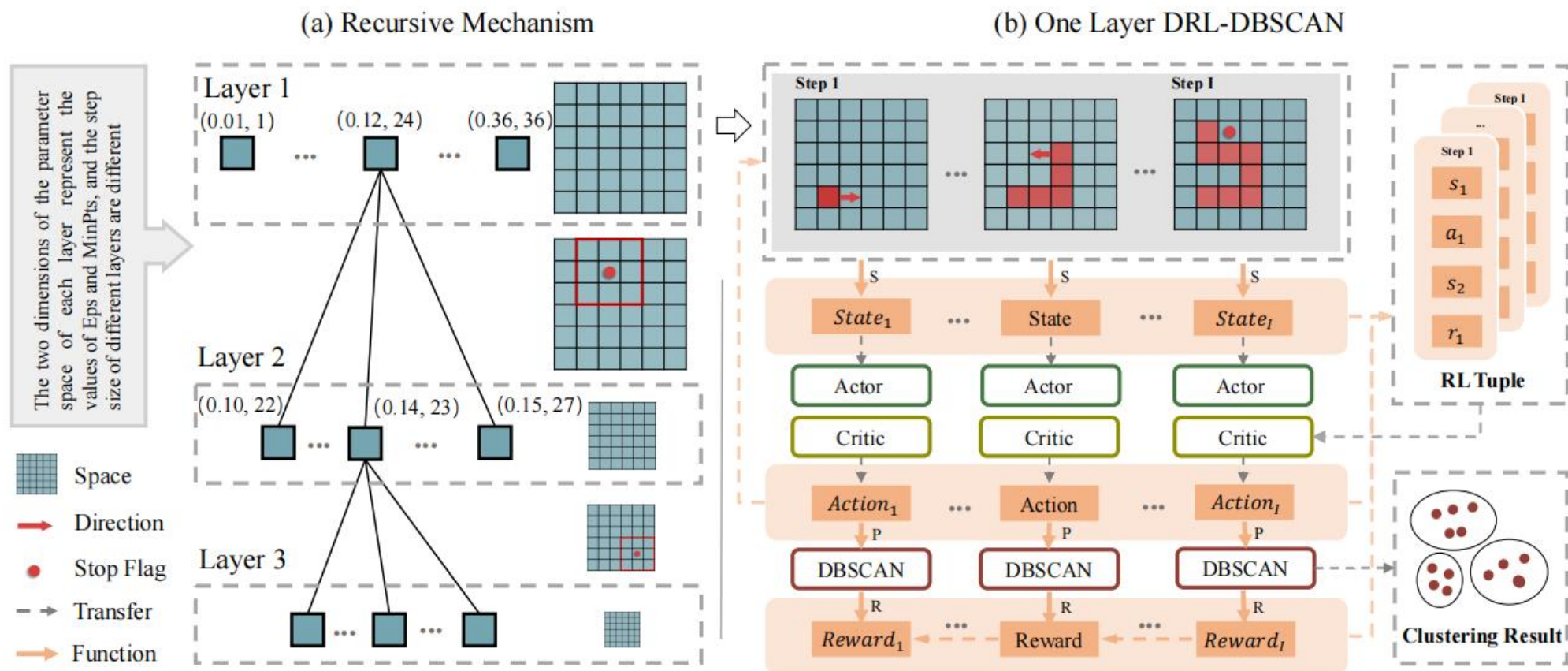


Figure 2: The core model of DRL-DBSCAN. (a) Recursive mechanism, takes 3-layer 6×6 parameter space as an example, with layerwise decreasing parameter space. (b) One layer DRL-DBSCAN, takes the search process in the 1-th layer of the recursive mechanism as an example, aims to obtain the optimal parameter combination in the parameter space of layer 1.

模型4:递归式深度强化参数自学习的高效密度聚类 (DRL-DBSCAN)

Table 2: Offline evaluation performance. The best results are bolded and second-best are underlined.

Dataset	Metrics	Traditional		Meta-heuristic				Dedicated				
		Rand	BO-TPE	Anneal	PSO	GA	DE	KDist	BDE	DRL _{re}	(Mean)	(Var.)
Pathbased	NMI	.66±.23	<u>.78±.07</u>	.65±.24	.60±.28	.68±.19	.22±.28	.40±.- -	.51±.33	.82±.03	↑.04	↓.04
	ARI	.63±.21	<u>.79±.10</u>	.66±.25	.55±.38	.67±.26	.18±.28	.38±.- -	.48±.40	.85±.04	↑.06	↓.06
Compound	NMI	<u>.75±.05</u>	.70±.24	.52±.36	.46±.34	.70±.25	.33±.35	.39±.- -	.72±.25	.78±.04	↑.03	↓.01
	ARI	<u>.73±.04</u>	.68±.24	.51±.35	.42±.36	.68±.24	.31±.34	.39±.- -	.70±.25	.76±.03	↑.03	↓.01
Aggregation	NMI	<u>.76±.11</u>	.72±.14	.75±.27	.59±.35	.75±.15	.28±.37	.60±.- -	.63±.28	.96±.02	↑.20	↓.09
	ARI	.68±.16	.63±.19	<u>.70±.27</u>	.51±.37	.68±.19	.25±.35	.52±.- -	.54±.28	.96±.03	↑.26	↓.13
D31	NMI	.31±.33	.23±.24	.17±.19	.36±.33	.23±.20	.24±.26	.07±.- -	<u>.41±.36</u>	.67±.02	↑.26	↓.17
	ARI	.14±.26	.04±.05	.03±.04	.09±.22	.04±.04	.06±.09	.00±.- -	<u>.21±.28</u>	.26±.02	↑.05	↓.02



模型4:递归式深度强化参数自学习的高效密度聚类 (DRL-DBSCAN)

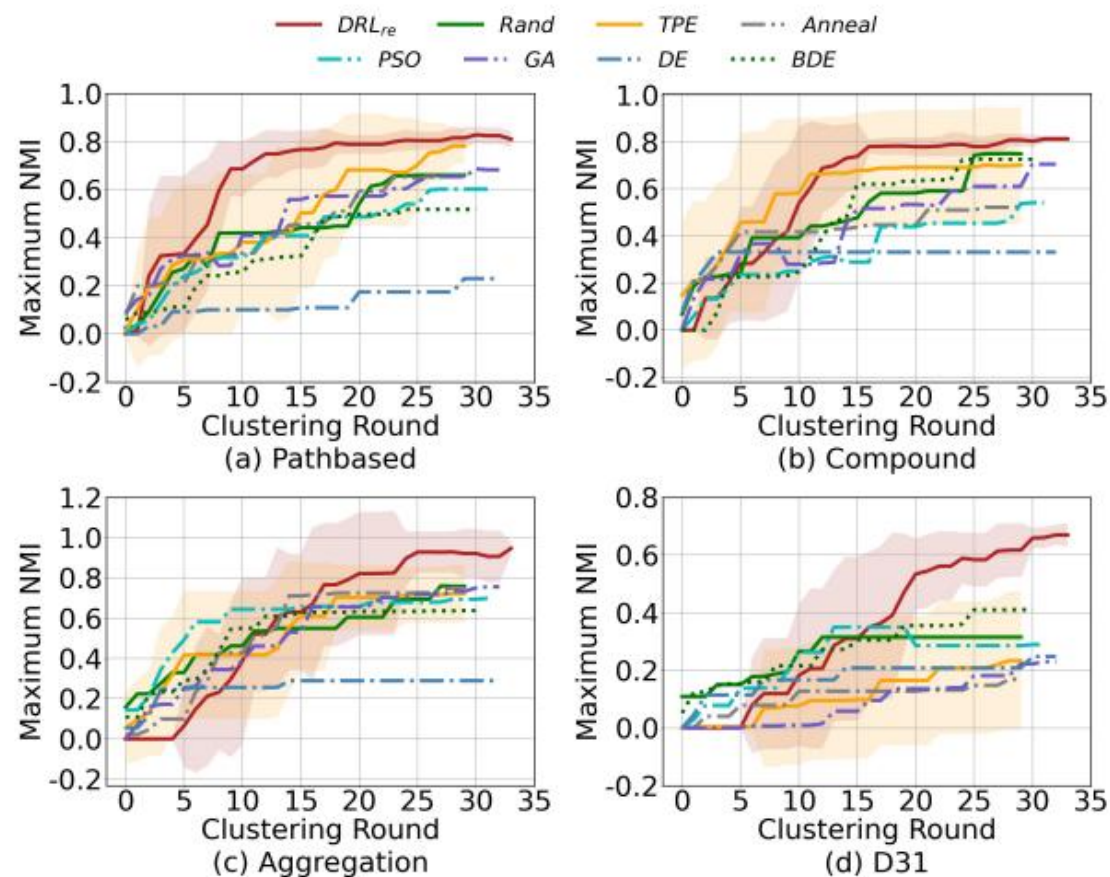


Figure 4: Offline clustering efficiency comparison.

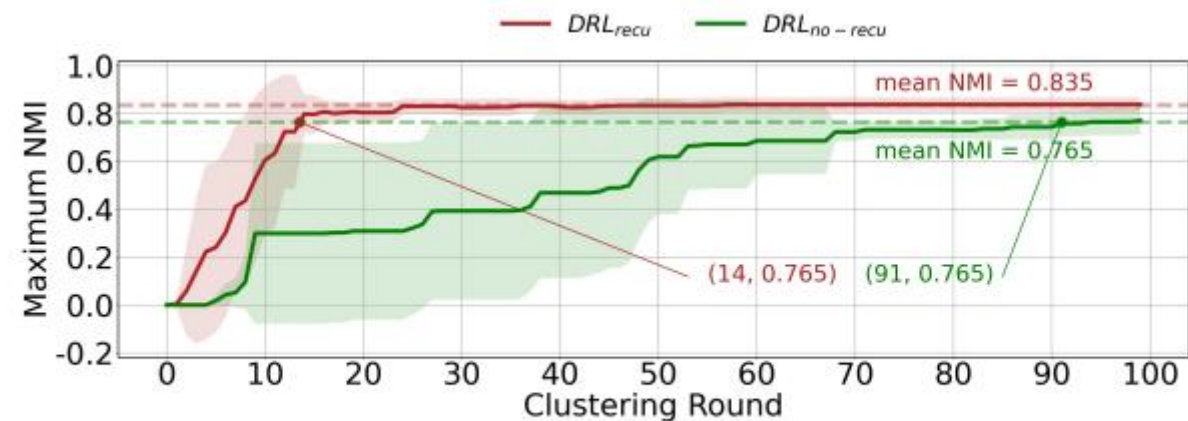


Figure 5: Efficiency comparison of recursive mechanism.

模型4:递归式深度强化参数自学习的高效密度聚类 (DRL-DBSCAN)

Table 4: Online evaluation NMI for training-based modes. The best results are bolded and second-best are underlined.

Blocks	Rand	BO-TPE	Anneal	PSO	GA	DE	KDist	BDE	DRL _{re}	DRL _{con}	(Mean)	(Var.)
$\mathcal{V}_9$	.67±.24	.83±.03	.53±.37	.74±.10	.65±.29	.19±.31	.30±.- -	.70±.21	<u>.86±.01</u>	.87±.00	↑.04	↓.03
$\mathcal{V}_{10}$	.36±.15	<u>.50±.07</u>	.45±.17	<u>.50±.20</u>	.43±.15	.15±.17	.20±.- -	.37±.20	<u>.50±.27</u>	.64±.06	↑.14	↓.01
$\mathcal{V}_{11}$	.40±.06	.43±.10	.32±.26	.55±.16	.43±.08	.09±.12	.12±.- -	.47±.16	<u>.60±.16</u>	.68±.02	↑.13	↓.04
$\mathcal{V}_{12}$	.44±.23	.62±.16	.27±.35	.66±.07	.50±.24	.19±.28	.11±.- -	.41±.31	.75±.01	<u>.72±.10</u>	↑.09	↓.06
$\mathcal{V}_{13}$	.84±.06	.87±.04	.72±.38	.68±.26	.76±.17	.38±.38	.62±.- -	.68±.23	.92±.02	.92±.02	↑.08	↓.02
$\mathcal{V}_{14}$	.74±.12	<u>.82±.04</u>	.54±.37	.63±.24	.54±.24	.25±.25	.55±.- -	.56±.25	.76±.25	.85±.00	↑.03	↓.04
$\mathcal{V}_{15}$	.68±.24	.76±.04	.66±.34	.55±.25	.62±.27	.28±.32	.36±.- -	.72±.14	.85±.07	<u>.83±.13</u>	↑.17	-
$\mathcal{V}_{16}$	.73±.13	.77±.09	.77±.10	.40±.35	.67±.22	.49±.31	.11±.- -	.67±.19	.86±.01	.86±.00	↑.09	↓.09

Table 5: Online evaluation NMI for testing-based modes. The best results are bolded and second-best are underlined.

Blocks	Rand	BO-TPE	Anneal	PSO	GA	DE	KDist	BDE	DRL _{all}	DRL _{one}	(Mean)	(Var.)
$\mathcal{V}_9$	.34±.31	.49±.33	.22±.34	.14±.29	.27±.37	.10±.26	.30±.- -	.54±.36	.68±.30	.68±.30	↑.19	-
$\mathcal{V}_{10}$	.11±.14	.28±.17	.17±.21	.24±.01	.20±.21	.12±.18	.20±.- -	.28±.24	.33±.16	.33±.15	↑.05	-
$\mathcal{V}_{11}$	.16±.15	.29±.24	.23±.18	.33±.29	.23±.23	.02±.05	.12±.- -	.21±.22	.30±.13	<u>.32±.08</u>	-	-
$\mathcal{V}_{12}$	.23±.25	.19±.24	.10±.22	<u>.38±.26</u>	.34±.27	.03±.06	.11±.- -	.29±.27	<u>.38±.17</u>	.46±.09	↑.08	-
$\mathcal{V}_{13}$	.58±.35	<u>.70±.24</u>	.47±.40	.44±.31	.36±.28	.08±.14	.62±.- -	.32±.26	.68±.34	.70±.27	-	-
$\mathcal{V}_{14}$	.36±.19	.34±.28	.47±.35	.37±.33	.27±.25	.11±.24	.55±.- -	.43±.28	<u>.60±.27</u>	.62±.16	↑.15	↓.03
$\mathcal{V}_{15}$	.45±.35	.38±.36	.37±.33	.30±.34	.36±.32	.09±.18	.36±.- -	.42±.31	<u>.64±.28</u>	.70±.03	↑.25	↓.15
$\mathcal{V}_{16}$	.22±.32	.45±.24	.32±.29	.19±.27	.36±.27	.12±.20	.11±.- -	<u>.59±.23</u>	.60±.27	.53±.20	↑.01	-

模型5：开放集社会事件检测的质量感知自提升图神经网络

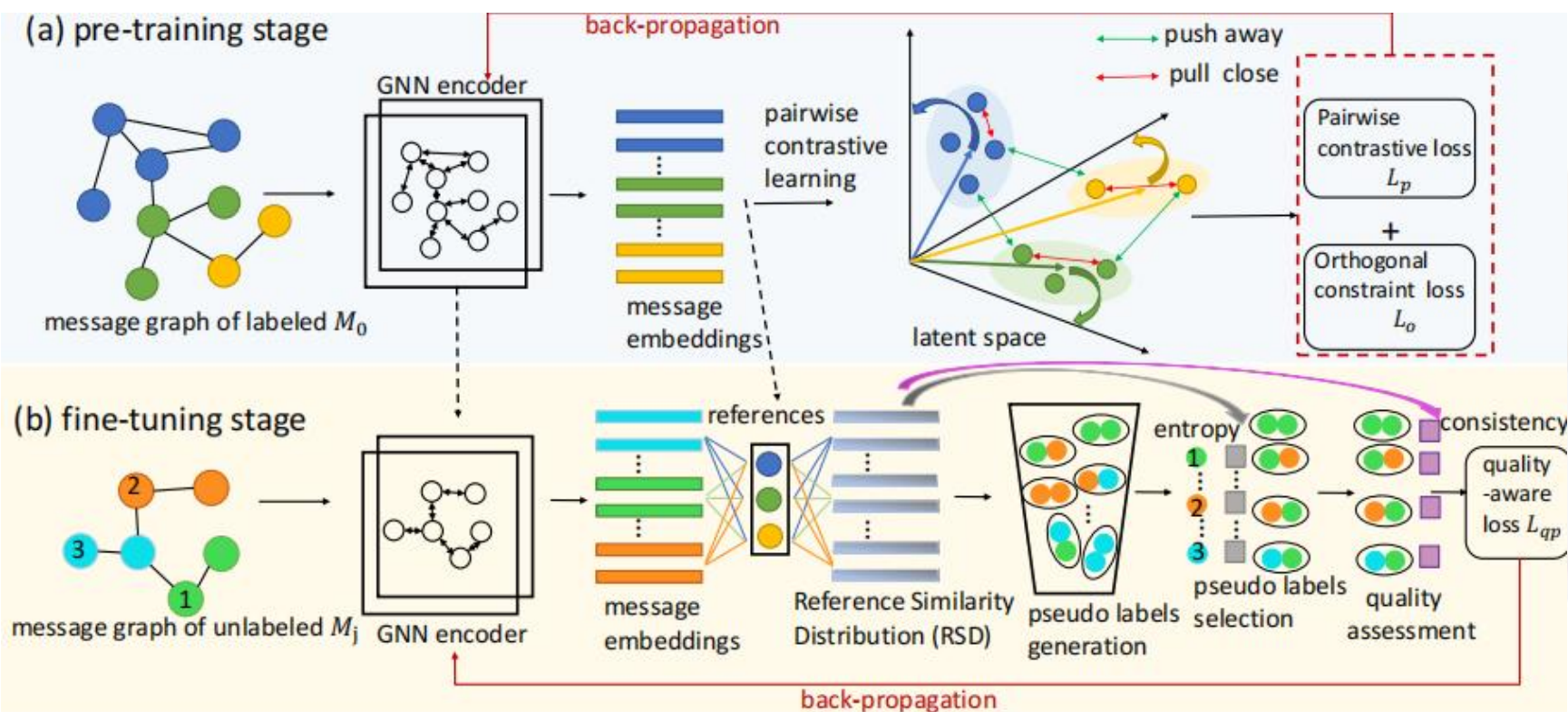
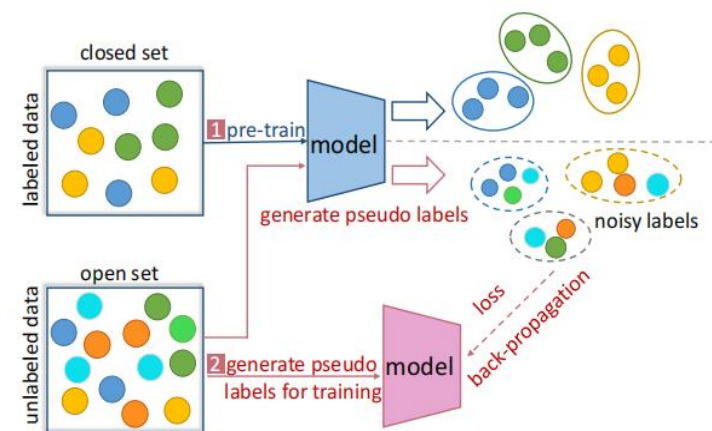


Figure 2: The architecture of QSGNN. (a) shows the pre-training stage, in which we fully explore the discriminative knowledge from the labeled data by utilizing both distance and direction information. (b) demonstrates the fine-tuning stage, in which we utilize the Reference Similarity Distribution vector for pseudo pairwise label generation, selection and quality assessment.

基于GNN编码器的自监督预训练

基于GNN编码器的自提升微调

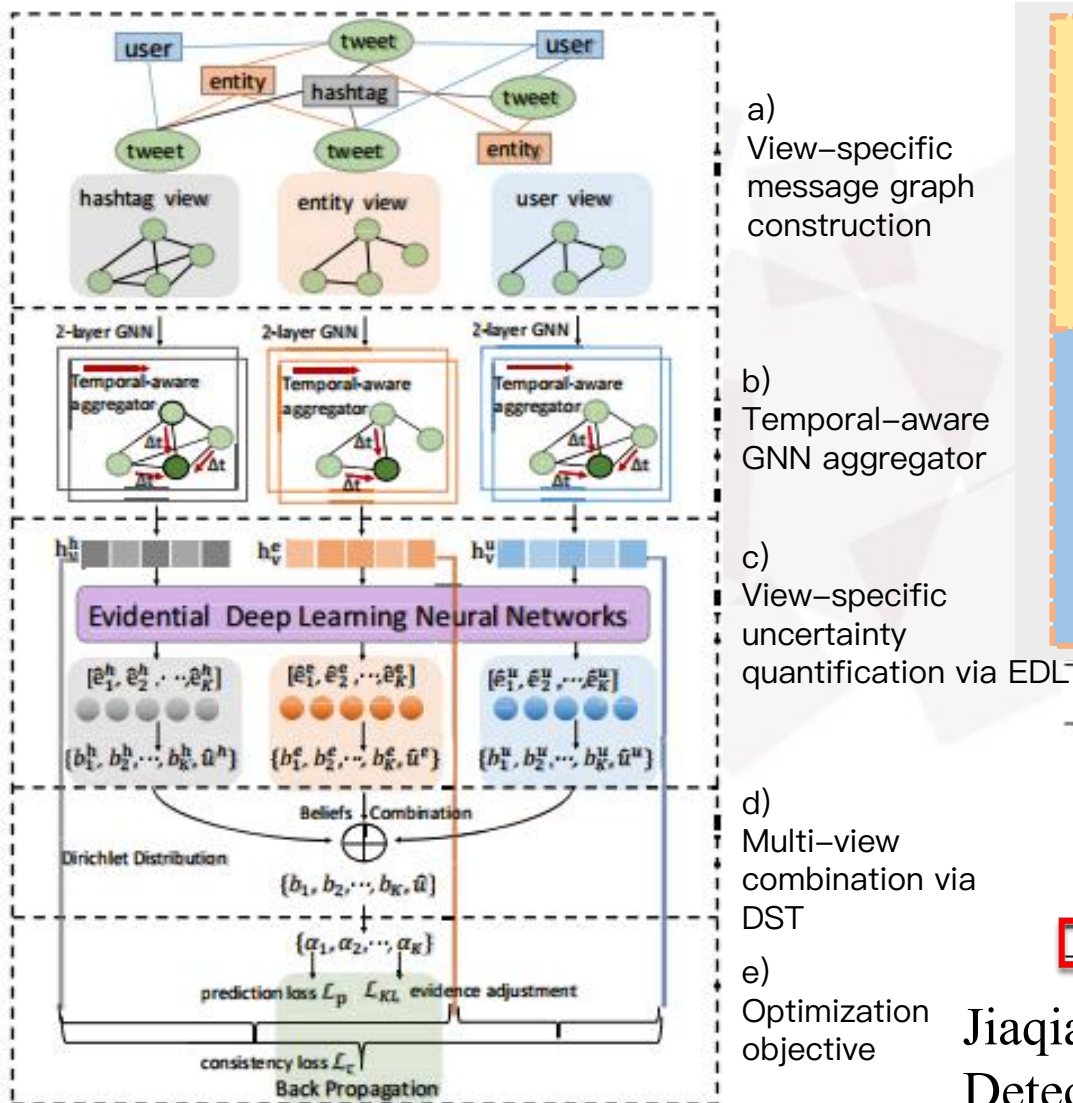


从已知到未知知识迁移两步学习策略

Methods	M_0 in Events2012		M_0 in Events2018	
	NMI	AMI	NMI	AMI
TwitterLDA [50]	.26±.00	.17±.00	.22±.00	.16±.00
Word2Vec [27]	.47±.00	.21±.00	.24±.00	.20±.00
BERT [8]	.63±.01	.44±.00	.42±.00	.34±.00
EventX [22]	.68±.00	.29±.00	.57±.00	.56±.00
PP-GCN [31]	.70±.02	.56±.01	.60±.01	.49±.02
KPGNN [6]	.76±.02	.64±.02	.66±.03	.60±.02
QSGNN w/o L_o	.77±.00	.65±.00	.68±.02	.61±.01
QSGNN	.79±.01	.68±.01	.71±.02	.64±.02
promotion	↑ 3%	↑ 4%	↑ 5%	↑ 4%



模型6：基于时间感知图的社会事件检测



瓶颈

不可靠观点引发的噪声信息

作为事件重要指标的时间信息

措施

时间感知聚合器

EDL+DST

Methods	Events2012		Kawaritz		CrisisLexT26	
	ACC	F1	ACC	F1	ACC	F1
TwitterLDA	0.3796	0.1376	0.5281	0.4484	0.5275	0.5212
JETS	0.5943	0.5614	0.6579	0.6444	0.7154	0.7008
Word2Vec	0.4484	0.4300	0.3980	0.3565	0.6484	0.6464
BERT	0.7768	0.6420	0.9140	0.9083	0.9242	0.9240
PP-GCN	0.6120	0.6171	0.8776	0.8671	0.8113	0.8057
KPGNN	0.7876	0.6899	0.9360	0.9309	0.9643	0.9646
MVGAN	0.7800	0.7358	0.9432	0.9410	0.9720	0.9680
ETGNN	0.8480	0.7565	0.9636	0.9610	0.9792	0.9794

Jiaqian Ren, etc., Evidential Temporal-aware Graph-based Social Event Detection, IEEE ICWS 2022

模型6：基于时间感知图的社会事件检测

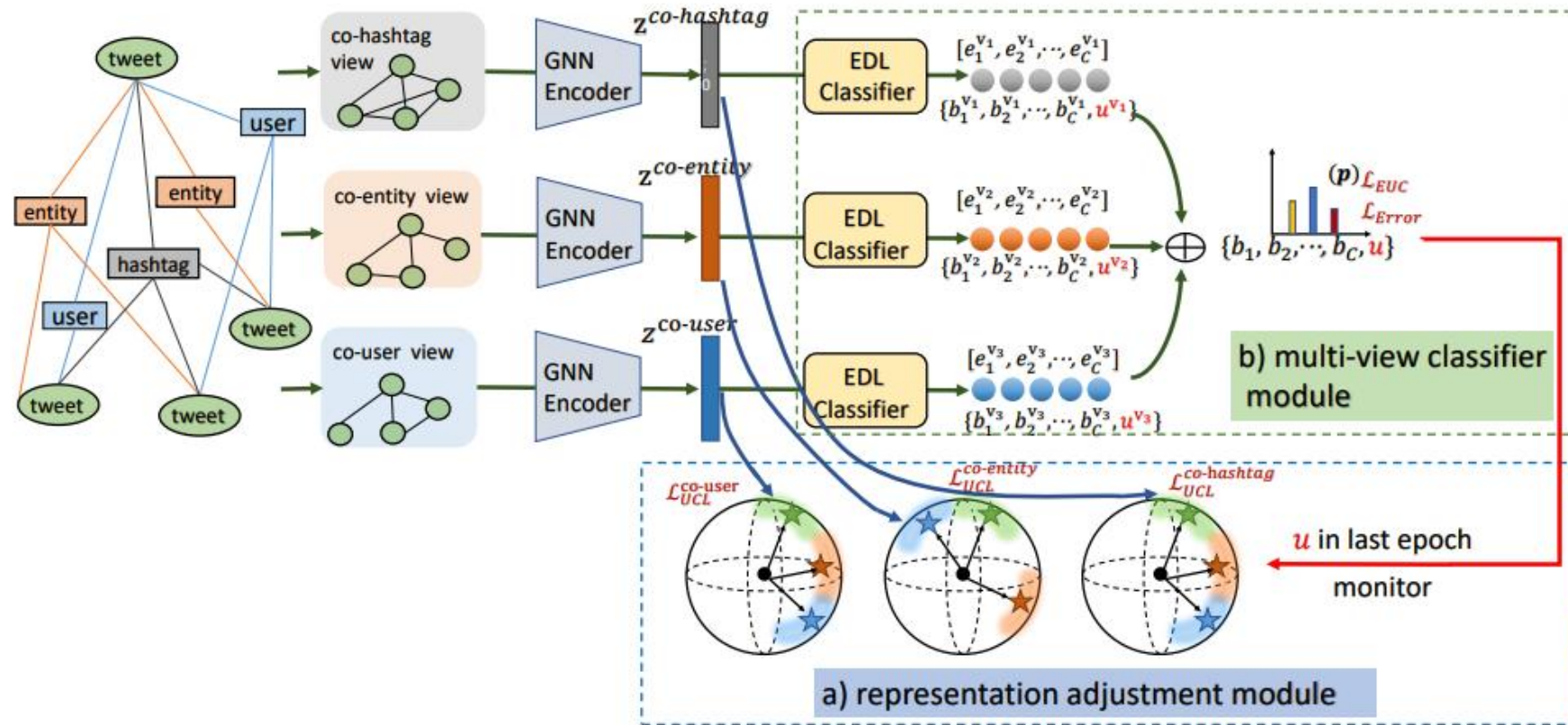


Fig. 2: The architecture of the proposed uncertainty-guided class imbalance learning framework (UCL_{SED} and $UCL-EC_{SED}$). The whole framework contains two modules: a) representation adjustment module, in which larger margins are assigned to more uncertain classes to ensure class separability; b) multi-view classifier module, in which multi-view results are combined via Dempster-Shafer theory with an additional calibration method to ensure robustness and accuracy.

模型6：基于时间感知图的社会事件检测

TABLE 2: Comparison with social event detection methods.

Methods	Events2012_100		Events2018_100		CrisisLexT_7	
	ACC	F1	ACC	F1	ACC	F1
TwitterLDA [25]	0.0937	0.0827	0.0690	0.0483	0.3190	0.2259
Word2Vec [42]	0.7467	0.7489	0.3517	0.3389	0.4429	0.3780
BERT [43]	0.7650	0.7681	0.5197	0.5213	0.5571	0.4597
PP-GCN [37]	0.6333	0.5462	0.7000	0.5099	0.7333	0.6871
KPGNN [38]	0.7333	0.5908	0.7667	0.6190	0.7567	0.7111
MVGNN [24]	0.8183	0.8214	0.6993	0.6817	0.7190	0.7055
ETGNN [21]	0.8645	0.8656	0.6043	0.6012	0.7667	0.7430
UCL _{SED}	0.9007	0.9022	0.7753	0.7854	0.8064	0.8053
UCL-EC _{SED}	0.9143	0.9160	0.7830	0.7903	0.8381	0.8364

TABLE 3: Comparison with long-tail recognition methods.

Methods	Events2012_100		Events2018_100		CrisisLexT_7	
	ACC	F1	ACC	F1	ACC	F1
CE	0.8577	0.8588	0.7373	0.7400	0.7476	0.7231
CB+Focal [7]	0.8767	0.8684	0.7557	0.7523	0.7524	0.7495
LDAM [56]	0.8983	0.8995	0.7730	0.7814	0.7684	0.7661
Hybrid-PSC [17]	0.8857	0.8865	0.7687	0.7660	0.7810	0.7620
BLC [19]	0.9083	0.9098	0.7733	0.7816	0.8190	0.8132
DRO-LT [20]	0.8943	0.8942	0.7707	0.7752	0.7886	0.7830
TSC [18]	0.9033	0.9080	0.7740	0.7833	0.8143	0.8058
UCL _{SED}	0.9007	0.9022	0.7753	0.7854	0.8064	0.8053
UCL-EC _{SED}	0.9143	0.9160	0.7830	0.7903	0.8381	0.8364

模型6：基于时间感知图的社会事件检测

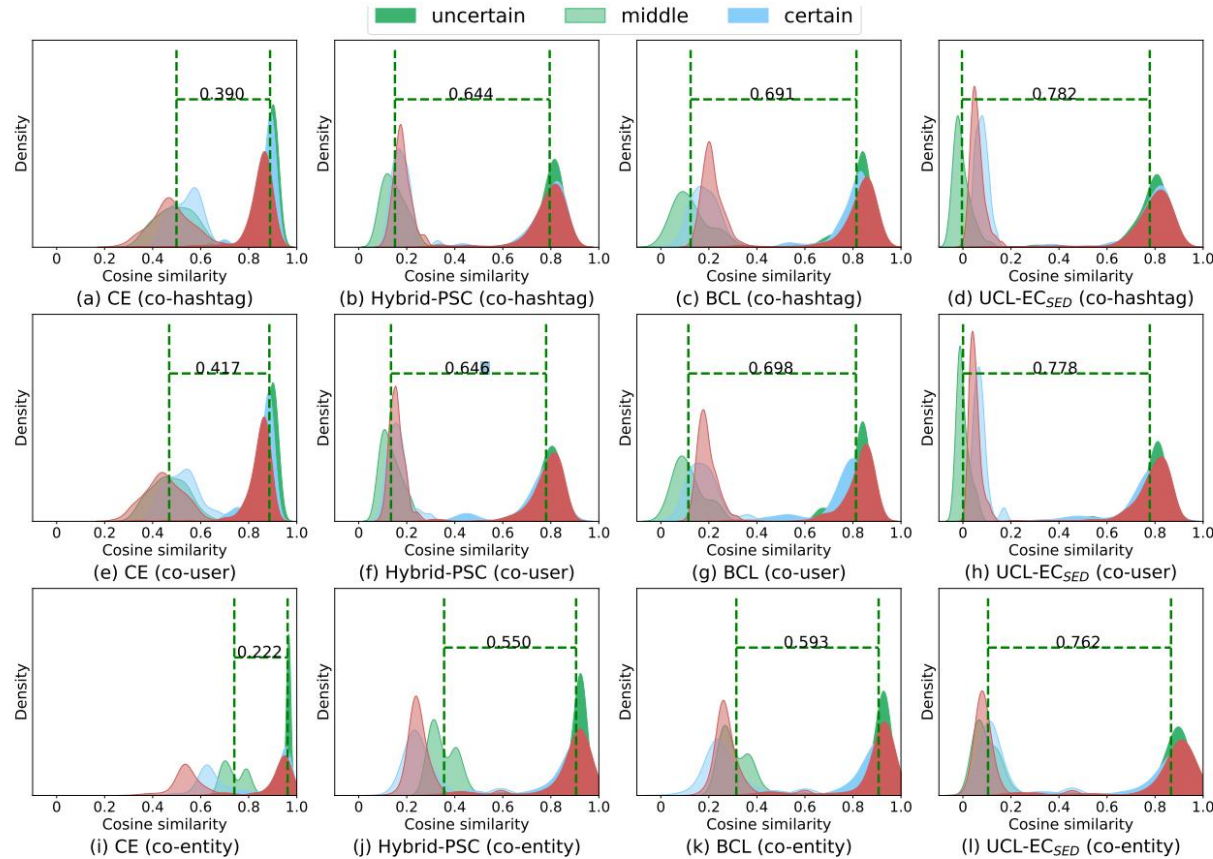
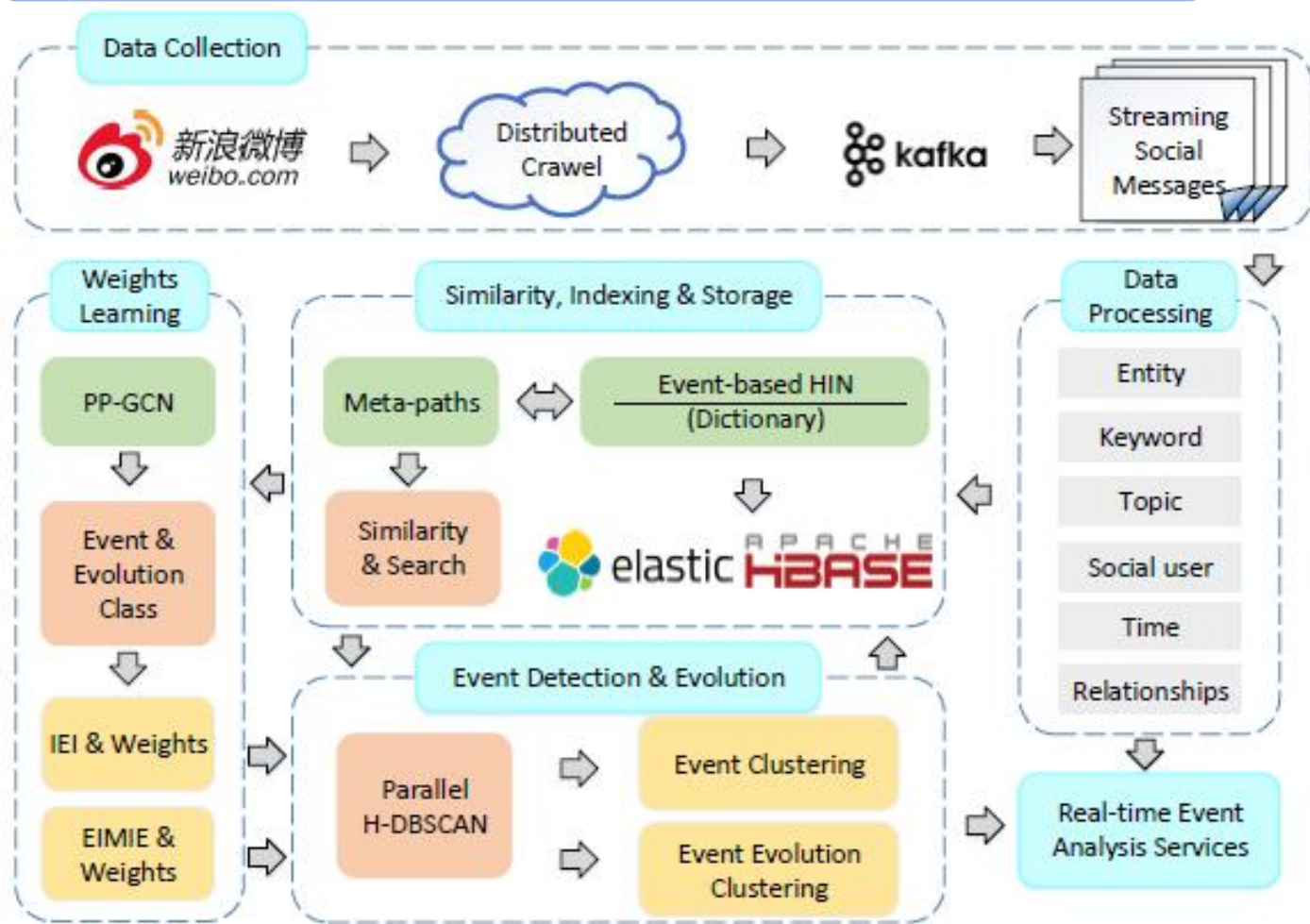
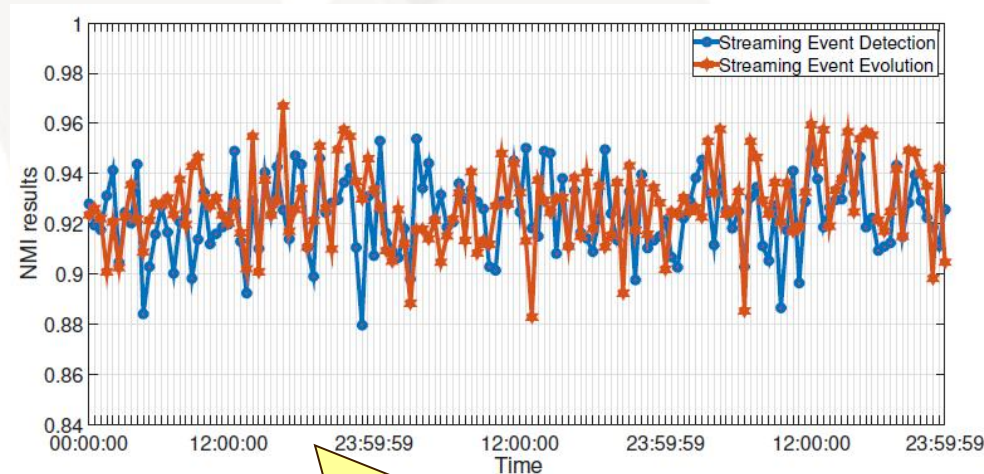
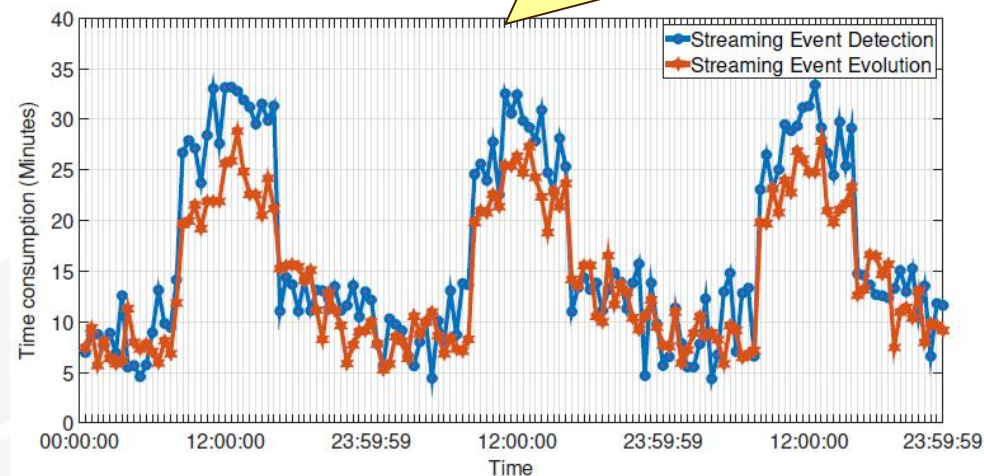


Fig. 5: Visualization of mean intra-class cosine similarity and mean inter-class cosine similarity distribution on Events2012_100. The first, second and third rows show the similarity distribution of *co-hashtag*, *co-entity* and *co-user* views, respectively. Dark color area indicates intra-class similarities while light color area indicates inter-class ones. Uncertain, middle, and certain classes are plotted separately. Average inter-class cosine similarity and average intra-class similarity of the uncertain group are marked with dashed green vertical lines.

模型7：异构图神经网络事件检测系统



流式在线处理耗时



流式事件检测准确率

Hao Peng, etc., Streaming Social Event Detection and Evolution Discovery in Heterogeneous Information Networks, ACM TKDD 2021, 谷歌学术引用 58

系统应用：全球突发舆情事件预警系统

全球突发威胁预警系统

中文

欢迎 screenadmin

信息筛选

信息标注

信息分发

数据采集展示

突发事件态势

国家安全报告

主题配置

采集配置

用户管理

起始: 2020-05-08 结束: 2020-05-09 是否推送: 全部 请输入关键字 查询

总30条, 每页20条, 总2页

1

2

利比亚军事行动中有至少10名民兵死亡、数十名民兵受伤

采集: 2020-05-09 01:41:52 站点: twitter.com 相关度: 62 类别: 军事

事发: 2020-05-08 19:40:00 地点: 的黎波里市 预警: 1-黄色

描述: 当地时间5月8日, 利比亚政府在声明中表示, 当天在梵蒂冈军事空军基地等地举行的空军军事行动中有至少10名民兵死亡、数十名民兵受伤。

利比亚的黎波里市Mitiga国际机场遭到轰炸

采集: 2020-05-09 01:10:54 站点: twitter.com 相关度: 100 类别: 袭击

事发: 2020-05-08 19:10:00 地点: 的黎波里市Mitiga国际机场 预警: 1-黄色

描述: 当地时间5月8日, 利比亚的黎波里市Mitiga国际机场遭到陆军元帅Khalifa Belqasim Haftar的部队轰炸, 请注意安全。

吉尔吉斯斯坦和塔吉克斯坦边界地区发生冲突事件

采集: 2020-05-08 23:55:36 站点: twitter.com 相关度: 42 类别: 暴力冲突

事发: 2020-05-08 21:00:00 地点: 巴特肯区 预警: 1-黄色

描述: 当地时间5月8日, 吉尔吉斯斯坦和塔吉克斯坦边界地区发生冲突事件, 造成3人受伤, 请注意安全。

利比亚Qasr al Hajj地区的Khafter车队遭遇枪击

采集: 2020-05-08 22:58:12 站点: twitter.com 相关度: 57 类别: 枪击

事发: 2020-05-08 16:40:00 地点: 西山区Qasr al Hajj地区 预警: 1-黄色

描述: 当地时间5月8日, 利比亚西山区Qasr al Hajj地区的Khafter车队遭遇枪击。

欧盟对利比亚发生的袭击事件表示强烈谴责

采集: 2020-05-08 22:25:08 站点: twitter.com 相关度: 100 类别: 政治

事发: 2020-05-08 16:20:00 地点: 的黎波里市 预警: 0-绿色

描述: 当地时间5月8日, 欧盟外交事务发言人Peter Stano在新闻发布会上对利比亚的黎波里市多地发生的火箭弹袭击事件表示强烈谴责。



系统应用：全球突发舆情事件预警系统

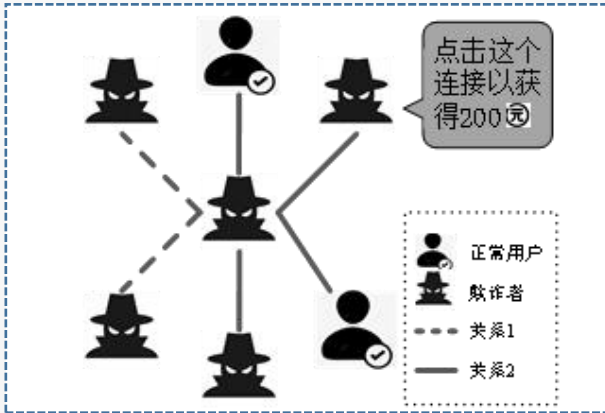




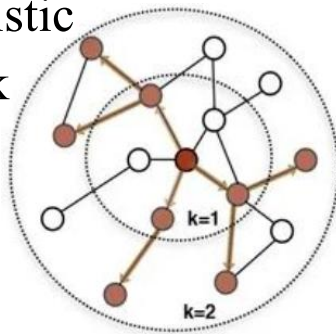
研究工作：网络用户行为异常检测

现状

Camouflaged Social Behaviors

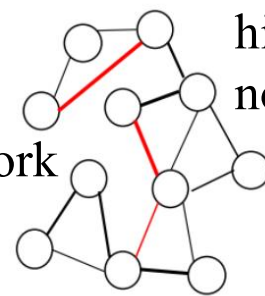


heuristic
black
box



Traditional Methods

static
network



Q: Reliability

sample annotation missing



Q: Platforms

高噪声行为异常检测难题

思路

特征融合学习

Neighborhood Selection

Structural Reconstruction

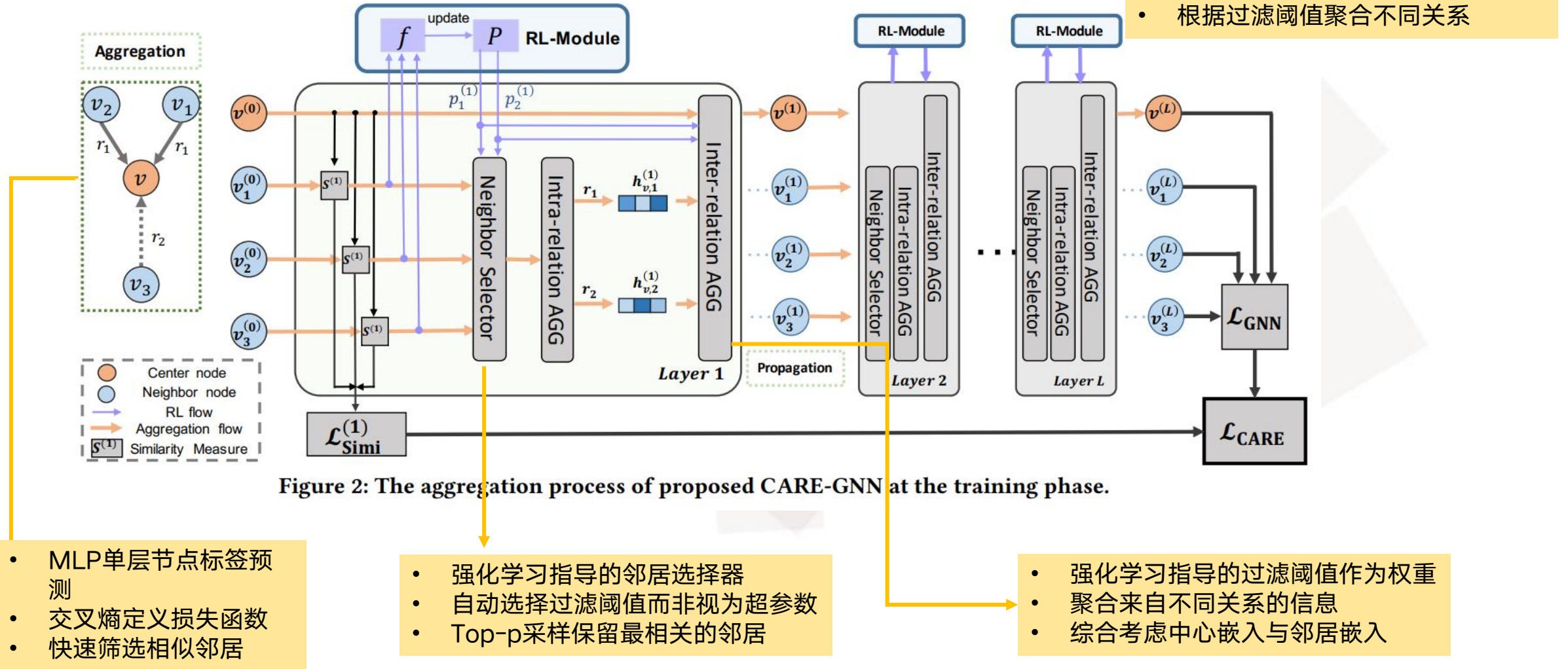
Federated Cross-Platform Detection

Abnormal Behavior
Detection

Accuracy

Generalization

模型8：增强图学习异常用户行为检测



模型8：增强的图神经网络异常用户检测

Table 2: Dataset and graph statistics.

	#Nodes (Fraud%)	Relation	#Edges	Avg. Feature Similarity	Avg. Label Similarity
Yelp	45,954 (14.5%)	<i>R-U-R</i>	49,315	0.83	0.90
		<i>R-T-R</i>	573,616	0.79	0.05
		<i>R-S-R</i>	3,402,743	0.77	0.05
		<i>ALL</i>	3,846,979	0.77	0.07
Amazon	11,944 (9.5%)	<i>U-P-U</i>	175,608	0.61	0.19
		<i>U-S-U</i>	3,566,479	0.64	0.04
		<i>U-V-U</i>	1,036,737	0.71	0.03
		<i>ALL</i>	4,398,392	0.65	0.05

数据集特点

- 巨量真实节点
- 复杂异构关系
- 多种实体行为
- 较低欺诈比例

包含对酒店和餐馆的良性实体评论与垃圾评论

包含对乐器的良性实体评论与垃圾评论

Table 3: Fraud detection performance (%) on two datasets under different percentage of training data.

	Metric	Train%	GCN	GAT	RGCN	Graph-SAGE	Genie-Path	Player-2Vec	Semi-GNN	Graph-Consis	CARE-Att	CARE-Weight	CARE-Mean	CARE-GNN
Yelp	AUC	5%	54.98	56.23	50.21	53.82	56.33	51.03	53.73	61.58	66.08	71.10	69.83	71.26
		10%	50.94	55.45	55.12	54.20	56.29	50.15	51.68	62.07	70.21	71.02	71.85	73.31
		20%	53.15	57.69	55.05	56.12	57.32	51.56	51.55	62.31	73.26	74.32	73.32	74.45
		40%	52.47	56.24	53.38	54.00	55.91	53.65	51.58	62.07	74.98	74.42	74.77	75.70
	Recall	5%	53.12	54.68	50.38	54.25	52.33	50.00	52.28	62.60	63.52	66.64	68.09	67.53
		10%	51.10	52.34	51.75	52.23	54.35	50.00	52.57	62.08	67.38	68.35	68.92	67.77
		20%	53.87	53.20	50.92	52.69	54.84	50.00	52.16	62.35	68.34	69.07	69.48	68.60
		40%	50.81	54.52	50.43	52.86	50.94	50.00	50.59	62.08	71.13	70.22	69.25	71.92
Amazon	AUC	5%	74.44	73.89	75.12	70.71	71.56	76.86	70.25	85.46	89.49	89.36	89.35	89.54
		10%	75.25	74.55	74.13	73.97	72.23	75.73	76.21	85.29	89.58	89.37	89.43	89.44
		20%	75.13	72.10	75.58	73.97	71.89	74.55	73.98	85.50	89.58	89.68	89.34	89.45
		40%	74.34	75.16	74.68	75.27	72.65	56.94	70.35	85.50	89.70	89.69	89.52	89.73
	Recall	5%	65.54	63.22	64.23	69.09	65.56	50.00	63.29	85.49	88.22	88.31	88.02	88.34
		10%	67.81	65.84	67.22	69.36	66.63	50.00	63.32	85.38	87.87	88.36	88.12	88.29
		20%	66.15	67.13	65.08	70.30	65.08	50.00	61.28	85.59	88.40	88.60	88.00	88.27
		40%	67.45	65.51	67.68	70.16	65.41	50.00	62.89	85.53	88.41	88.45	88.22	88.48

半监督学习快速训练模型

动态调整阈值强泛化能力

模型8：增强的图神经网络异常用户检测

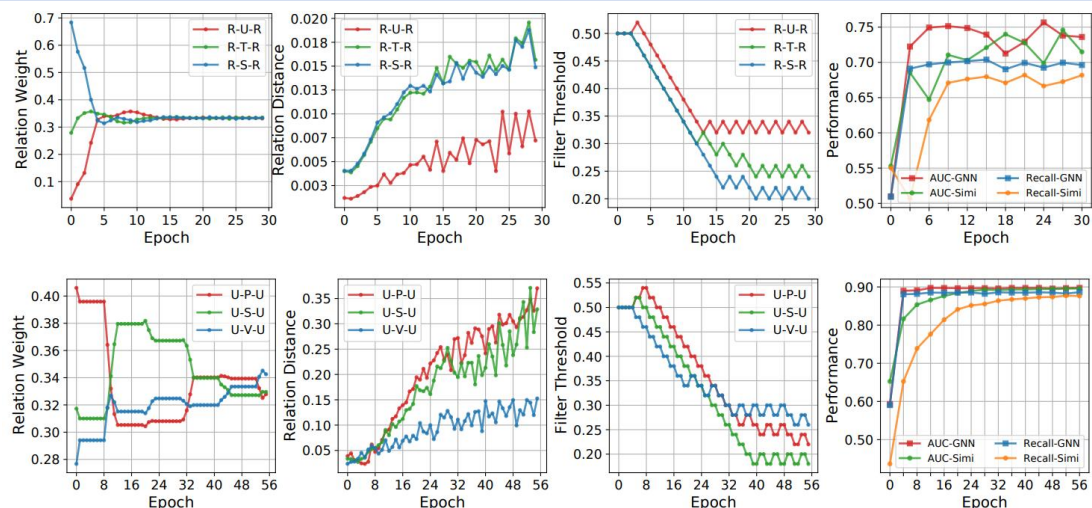


Figure 3: The training process and testing performance of CARE-Weight on Yelp (upper) and Amazon (lower) dataset.

邻居选择的必要性

不同关系的区分度

滤波阈值的稳定性

模型综合表现对比

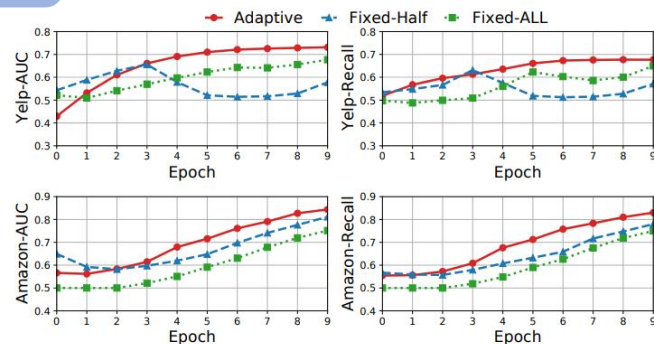


Figure 4: The testing AUC and Recall for CARE-GNN with different neighbor filtering methods during training.

自适应滤波优越性

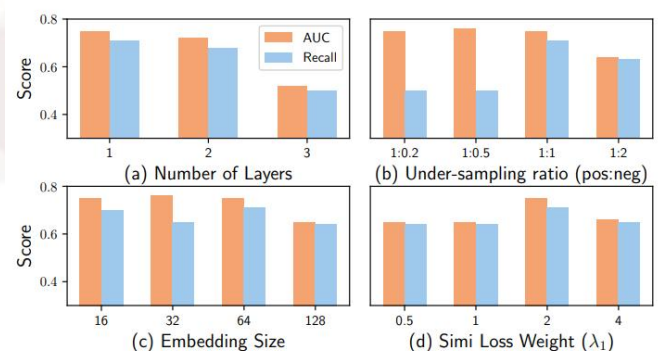


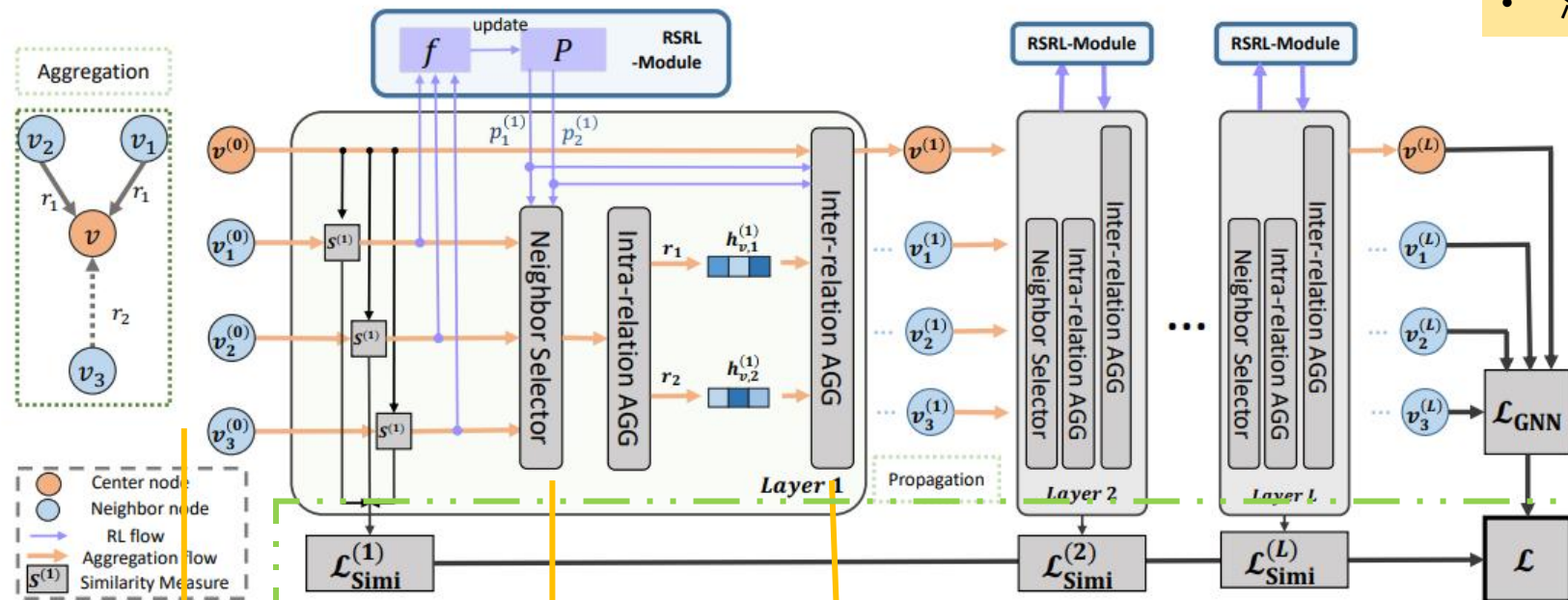
Figure 5: Parameter Sensitivity. For each parameter configuration, only the best results among 30 epochs are recorded.

参数灵敏度分析

- 训练层数
- 欠采样比例
- 嵌入规模
- 相似损失权重



模型8：递归增强的图学习异常用户检测



改进

- 泛化能力更强的RSRL框架

- 标签感知相似性度量
- 交叉熵定义损失函数
- GNN半监督过滤邻居

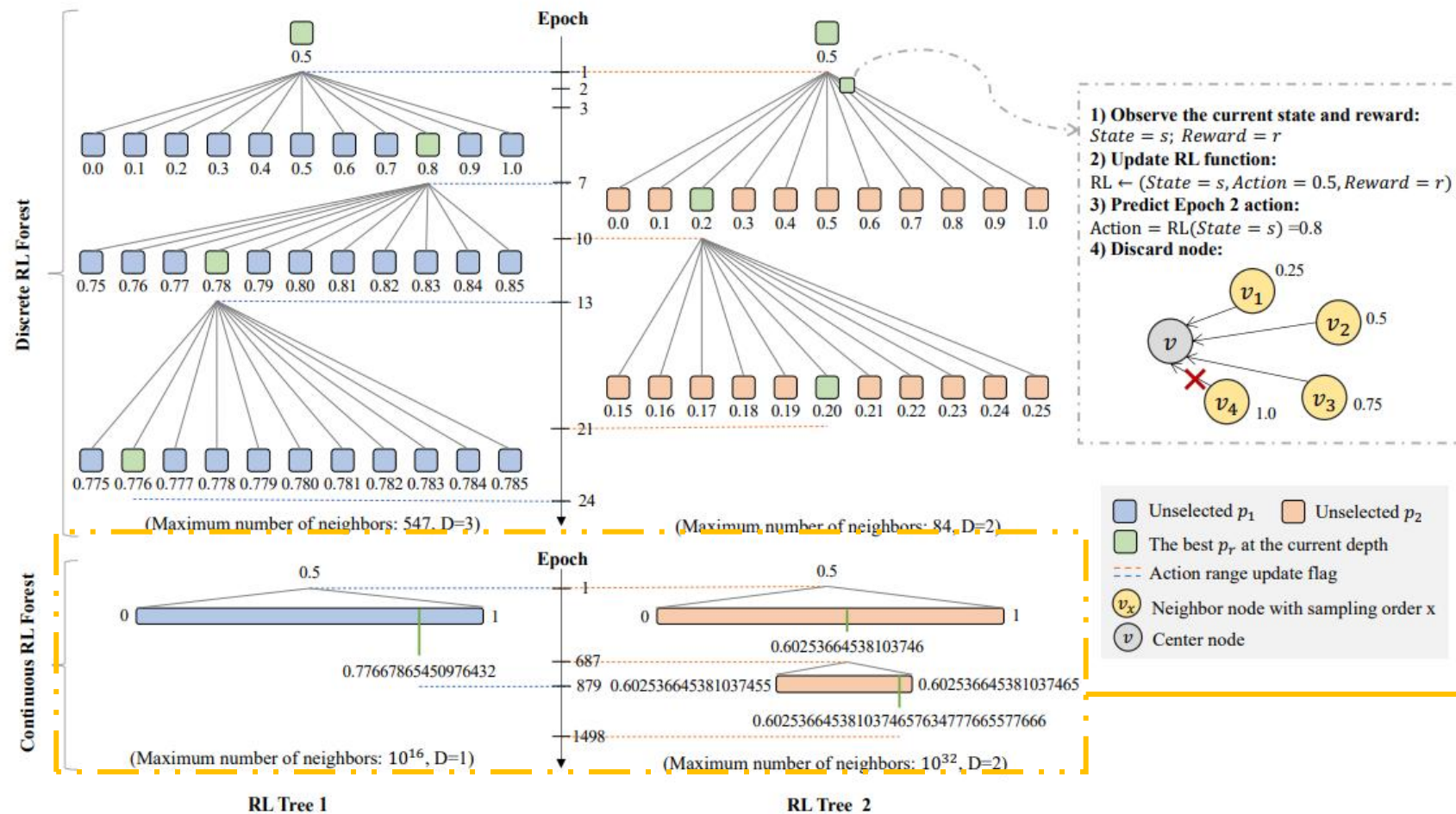
- RSRL框架指引邻居选择器
- 自动选择过滤阈值而非视为超参数
- Top-p采样保留最相关的邻居

- 强化学习指导的过滤阈值作为权重
- 聚合来自不同关系的信息
- 综合考虑中心嵌入与邻居嵌入

- GNN节点分类损失函数
- 相似性度量损失函数



模型8：递归增强的图学习异常用户检测



RSRL框架要点

- 递归选择更高精度阈值
- 连续动作空间离散化
- 标签感知距离度量作为状态
- 节点相似度视为奖励
- 连续三次相同动作终止迭代

离散&连续空间下强化学习滤波阈值自适应选择

Fig. 4. One layer Reinforcement Learning Forest.



模型8：递归增强的图学习异常用户检测

准确性分析

Models	Yelp								Amazon							
	AUC				Recall				AUC				Recall			
	5%	10%	20%	40%	5%	10%	20%	40%	5%	10%	20%	40%	5%	10%	20%	40%
GCN	54.98	50.94	53.15	52.47	53.12	51.10	53.87	50.81	74.44	75.25	75.13	74.34	65.54	67.81	66.15	67.45
GAT	56.23	55.45	57.69	56.24	54.68	52.34	53.20	54.52	73.89	74.55	72.10	72.16	63.22	65.84	67.13	65.51
GraphSAGE	53.82	54.20	56.12	54.00	54.25	52.23	52.69	52.86	70.71	73.97	73.97	75.27	69.09	69.36	70.30	70.16
RGCN	50.21	55.12	55.05	53.38	50.38	51.75	50.92	50.43	75.12	74.13	75.58	74.68	64.23	67.22	65.08	67.68
GeniePath	56.33	56.29	57.32	55.91	52.33	54.35	54.84	50.94	71.56	72.23	71.89	72.65	65.56	66.63	65.08	65.41
Player2Vec	51.03	50.15	51.56	53.65	50.00	50.00	50.00	50.00	76.86	75.73	74.55	56.94	50.00	50.00	50.00	50.00
SemiGNN	53.73	51.68	51.55	51.58	52.28	52.57	52.16	50.59	70.25	76.21	73.98	70.35	63.29	63.32	61.28	62.89
GraphConsis	61.58	62.07	62.31	62.07	62.60	62.08	62.35	62.08	85.46	85.29	85.50	85.50	85.49	85.38	85.59	85.53
GAS	54.43	52.58	52.51	52.60	53.40	53.26	53.37	51.61	71.40	77.49	74.51	71.03	64.31	64.57	62.08	63.74
FdGars	61.77	62.15	62.81	62.66	62.83	62.16	62.73	62.40	85.58	85.41	85.88	85.81	85.83	85.73	85.84	85.93
GraphNAS^H	52.93	54.69	56.73	54.46	52.40	54.15	55.69	56.16	71.01	72.48	73.52	76.05	69.17	69.48	70.35	70.16
GraphNAS	53.26	55.31	57.15	55.59	53.69	55.47	56.04	57.00	72.41	73.04	73.58	76.25	70.36	70.53	71.73	71.88
Policy-GNN^H	54.04	55.73	59.30	60.60	53.08	55.35	58.75	59.99	72.20	73.30	74.11	77.20	70.10	71.20	73.08	74.44
Policy-GNN	55.75	56.29	60.01	61.52	54.15	56.16	58.95	60.33	73.69	74.06	75.29	78.85	71.34	72.46	74.55	76.70
CARE-GNN	71.26	73.31	74.45	75.70	67.53	67.77	68.60	71.92	89.54	89.44	89.45	89.73	88.34	88.29	88.27	88.48
RioGNN	81.97	83.72	82.31	83.54	75.33	75.78	75.51	76.19	95.44	95.41	95.63	96.19	90.17	89.48	89.51	89.82

Table 4. Fraud Detection results (%) compared to the baselines.



模型8：递归增强的图学习异常用户检测

- 多层GNN效果更优
- 多深度结构的快速准确收敛
- 过滤阈值作为关系间权重的优越性

Table 5. Fraud Detection classification results (%) compared to RioGNN variants.

Models	Yelp		Amazon	
	AUC	Recall	AUC	Recall
RioGNN _{2l}	76.01	63.15	91.28	72.46
BIO-GNN	78.67	71.21	95.47	88.35
ROO-GNN	83.59	75.56	95.58	89.22
RIO-Att	78.65	71.69	93.97	83.78
RIO-Weight	80.40	72.83	96.25	89.61
RIO-Mean	77.84	71.43	94.57	89.47
RioGNN	83.54	75.55	96.19	88.66

RSRL框架要点

- 递归选择更高精度阈值
- 连续动作空间离散化
- 标签感知距离度量作为状态
- 节点相似度视为奖励
- 连续三次相同动作终止迭代

Table 6. Fraud detection clustering results (%) compared to RioGNN variants.

Dataset	Metric	RioGNN _{2l}	BIO-GNN	ROO-GNN	RIO-Att	RIO-Weight	RIO-Mean	RioGNN
Yelp	NMI	3.18	9.36	12.39	9.80	12.05	8.39	12.22
	ARI	6.12	11.84	16.61	11.88	15.88	8.80	16.45
Amazon	NMI	58.87	59.83	57.81	55.76	58.76	58.72	61.26
	ARI	76.53	77.38	76.09	76.54	76.73	76.51	78.40

- RIOGNN在密集数据集上的显著优势

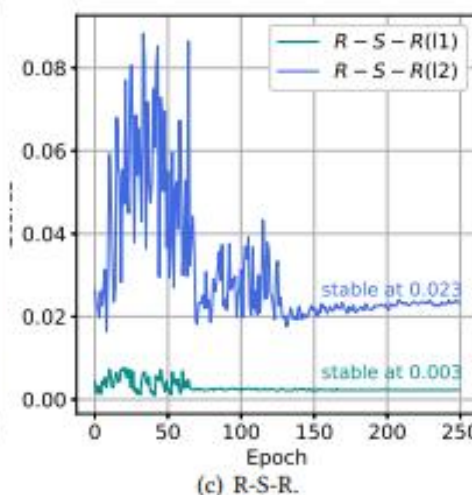
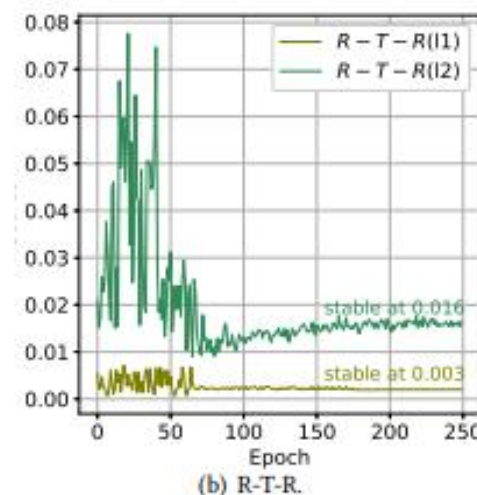
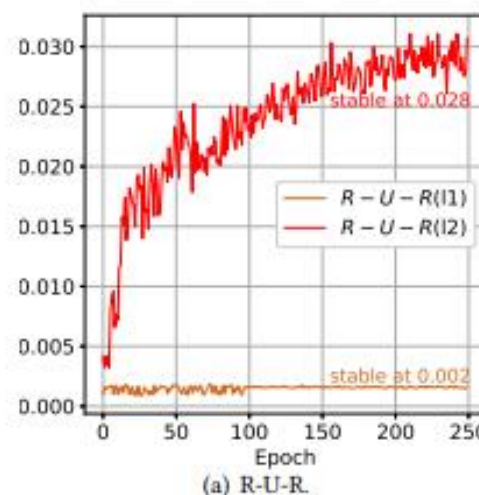


Fig. 7. Scores of Multi-Layer RioGNN on Yelp.



模型8：递归增强的图学习异常用户检测

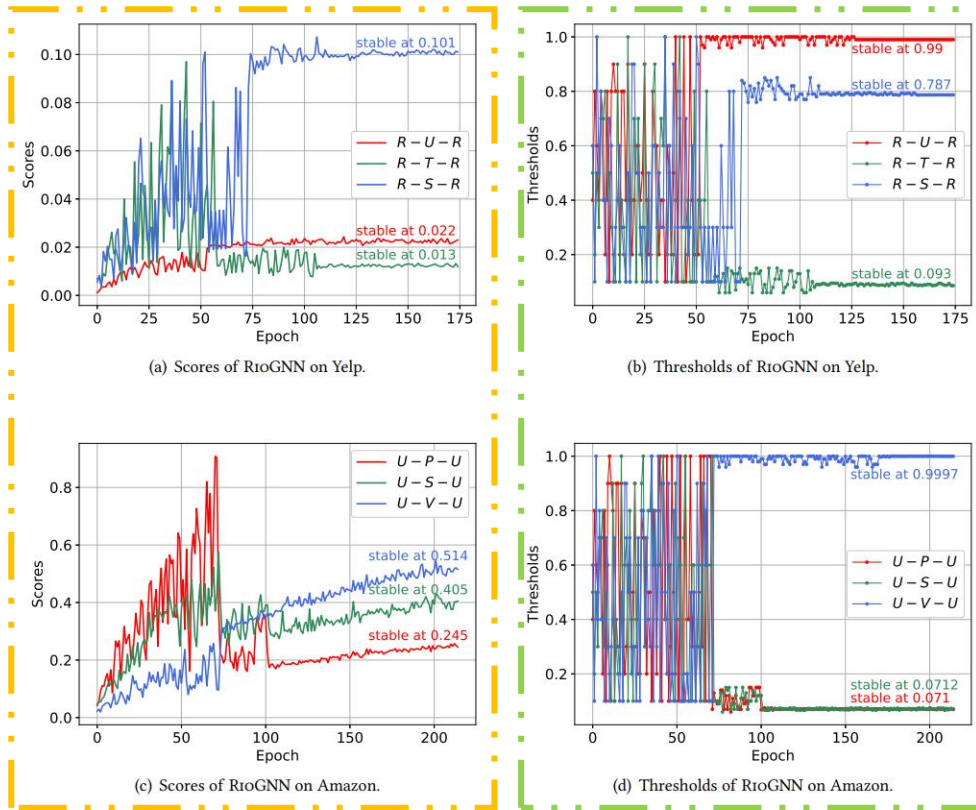


Fig. 5. The training scores and thresholds of RioGNN on Yelp and Amazon.

奖励对照

不同关系的重要性对照

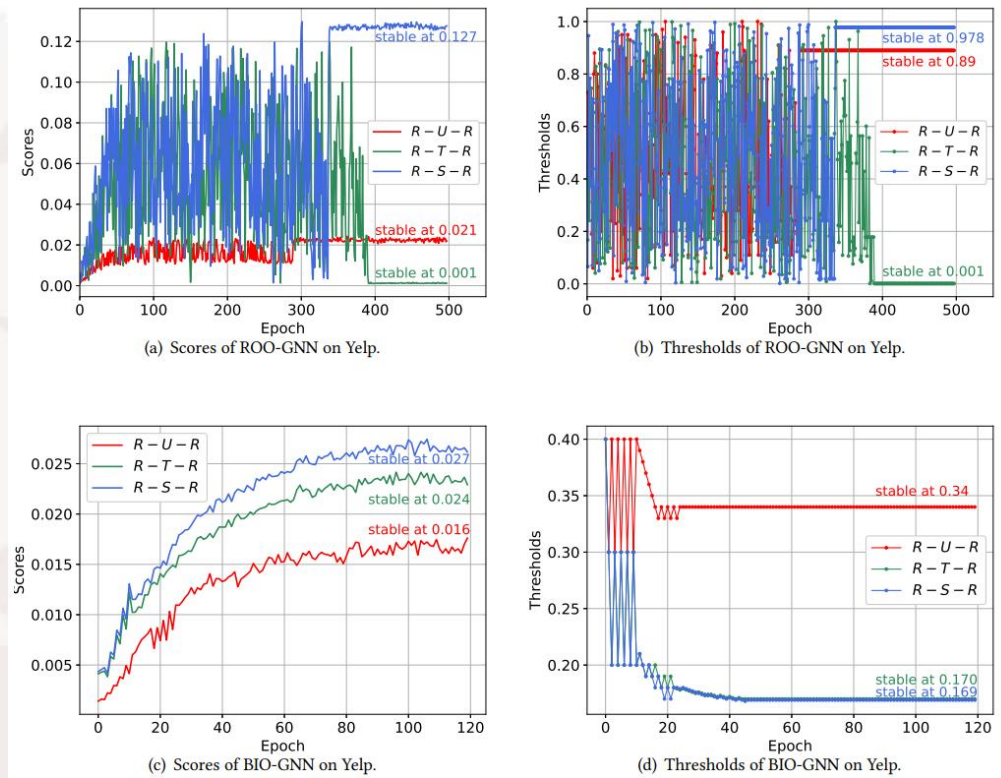


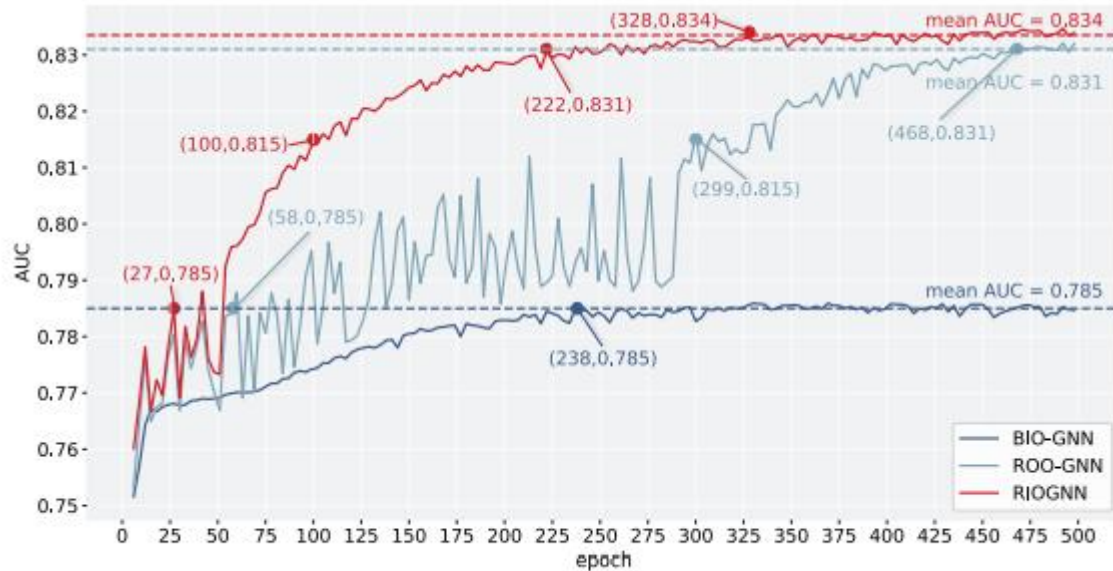
Fig. 6. The training scores and thresholds of RioGNN variants on Yelp.

RIOGNN变体对照（消融实验）

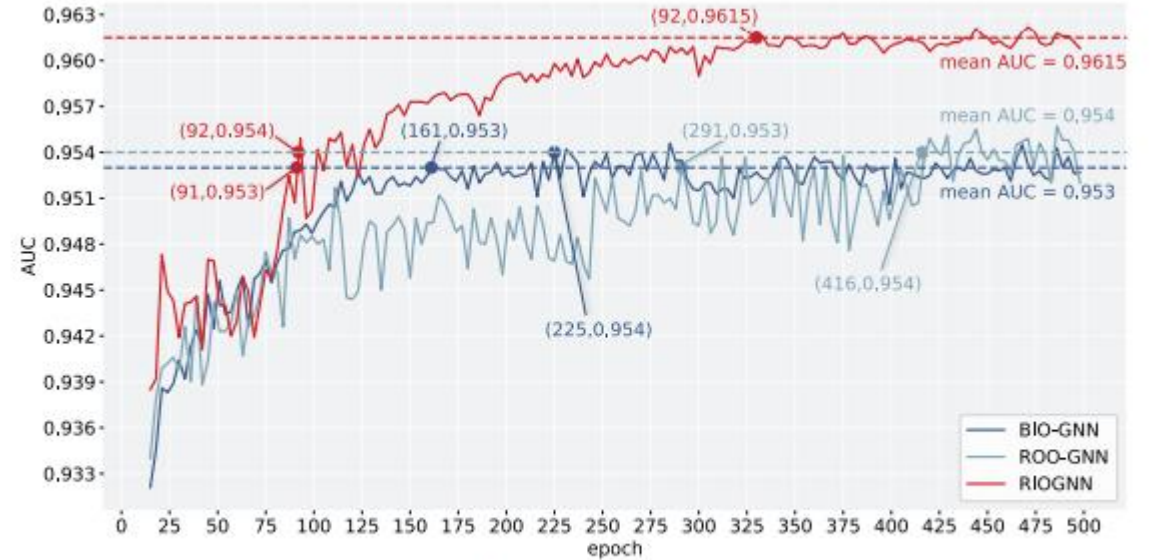


模型8：递归增强的图学习异常用户检测

递归框架影响分析



(a) AUC of Rio-GNN, BIO-GNN and ROO-GNN on Yelp.



(b) AUC of RioGNN and RioGNN No Recursion on Amazon.



模型8：递归增强的图学习异常用户检测

RS
RL
框架
通用
性
分
析

Table 10. Results (%) compared to different RL algorithms and strengthening strategies.

Methods		Yelp	Amazon	MIMIC-III
Discrete	AC [50]	83.54	96.19	81.36
	DQN [69]	84.08	95.13	80.96
	PPO [86]	80.52	94.99	80.98
Continuous	AC [50]	81.31	94.72	80.98
	DDPG [55]	83.80	95.39	81.17
	SAC [31]	80.42	94.76	80.87
	TD3 [23]	84.18	95.11	81.51

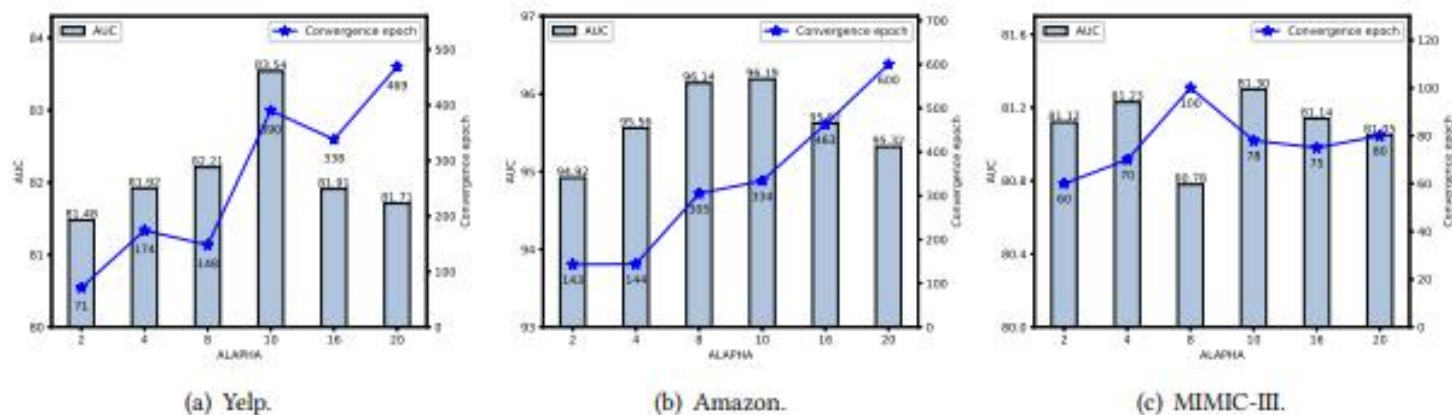
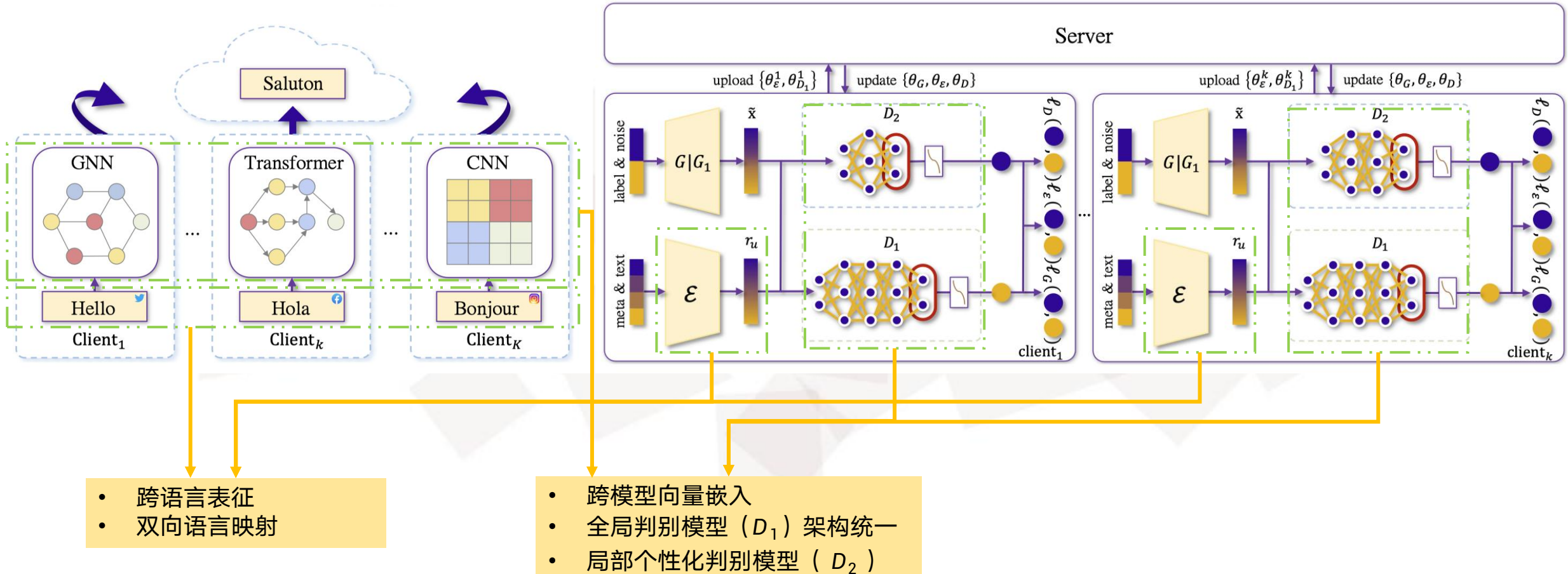


Fig. 11. Depth and Width for Different Task Scenarios.

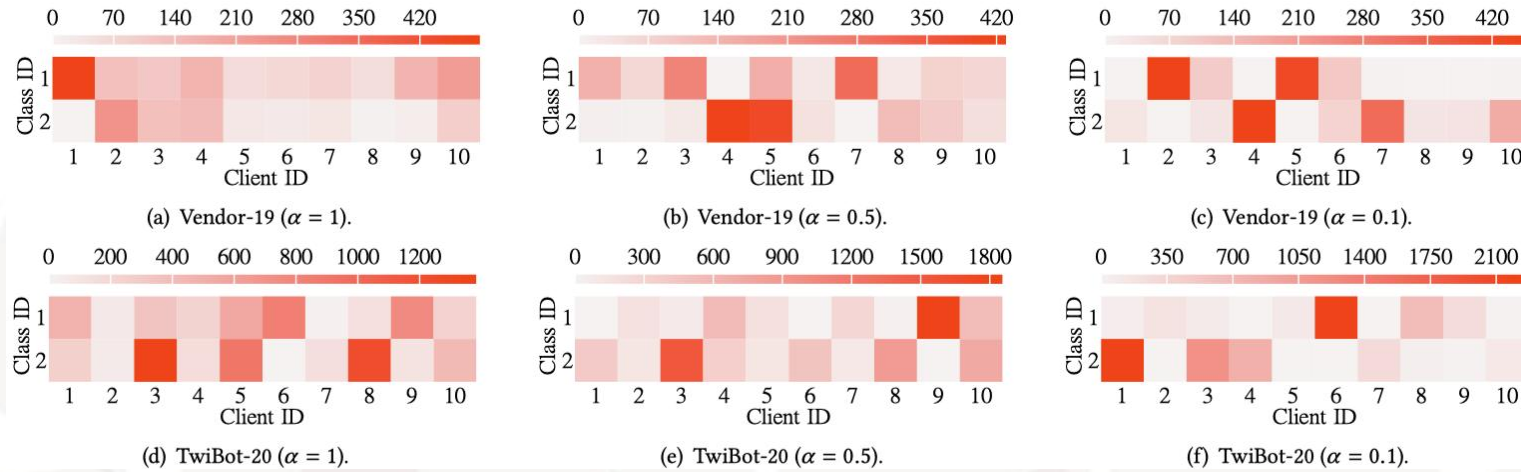
模型9：联邦的跨语言、跨平台、跨模型水军检测框架





模型9：联邦的跨语言、跨平台、跨模型水军检测框架

准确性分析



- 跨平台
- 数据分布异构
- 标签比例占比悬殊

Table 2: Comparison of the average maximum accuracy of different methods for social bot detection (%).

Dataset	Vendor-19				TwiBot-20			
Setting	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 0.1$	$\alpha = 0.05$
FedAvg	71.30±0.60	61.06±1.52	60.88±2.85	59.81±2.48	54.04±0.50	55.41±1.35	51.37±0.77	52.46±0.02
FedProx	84.37±0.43	78.25±1.02	51.86±0.04	63.27±2.32	74.34±0.06	73.32±0.25	51.86±0.04	52.30±0.63
FedDF	86.37±1.23	80.17±2.21	63.16±1.37	67.01±1.78	72.12±1.96	71.25±1.03	55.23±1.32	53.35±1.41
FedEnsemble	81.12±2.22	76.70±1.21	64.51±2.56	68.05±1.15	55.98±2.55	54.15±0.04	54.21±0.04	54.15±0.04
FedDistill	79.68±0.58	68.77±1.13	52.88±0.06	70.25±0.39	64.11±0.29	63.34±0.56	50.00±0.00	54.30±0.05
FedGen	90.05±0.33	84.83±0.96	65.12±0.60	70.79±2.39	74.14±0.47	73.12±2.09	59.19±2.70	55.78±1.79
FedFTG	88.31±1.41	82.17±1.52	66.01±1.25	68.39±1.94	74.27±1.21	74.13±0.53	60.14±1.74	56.17±1.27
FEDACK-A	91.31±0.52	84.79±1.05	66.10±2.90	68.21±1.95	77.16±1.09	74.70±1.64	63.52±1.09	55.39±1.24
FEDACK	88.58±1.91	87.05±2.03	76.04±3.40	75.27±2.50	77.08±1.83	78.26±2.60	67.81±2.20	60.14±1.32
Gain	↑ 1.26~20.01	↑ 2.22~25.99	↑ 10.03~24.18	↑ 4.48~15.46	↑ 2.82~23.12	↑ 4.13~24.11	↑ 7.67~16.44	↑ 3.97~7.84



模型9：联邦的跨语言、跨平台、跨模型水军检测框架

FedACK 框架效率分析

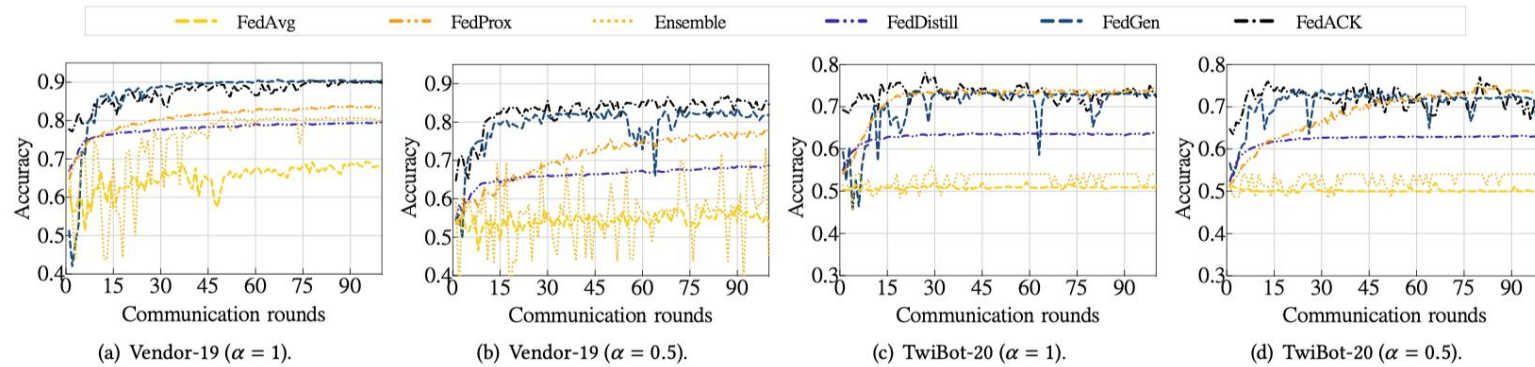
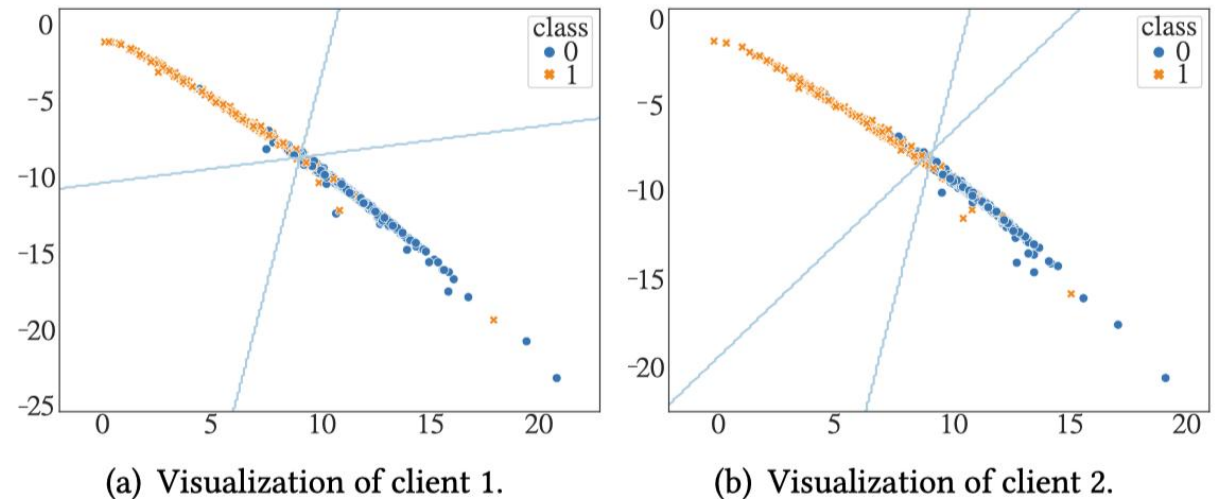


Figure 4: Learning Curve of (a-b) Vendor-19 and (c-d) TwiBot-20 in 100 communication rounds in different α settings.

Table 3: The round number to reach the target accuracy on Vendor-19 (80%, 70%) and TwiBot-20 (70%, 65%).

Dataset	Vendor-19		TwiBot-20	
Setting	$\alpha = 1$ (80)	$\alpha = 0.5$ (70)	$\alpha = 1$ (70)	$\alpha = 0.5$ (65)
FedAvg	unreached	unreached	unreached	unreached
FedProx	25.3±3.1	32.6±2.3	13.3±2.8	24.0±7.3
FedDF	22.3±2.4	38.4±3.1	50.3±5.2	60.2±6.4
Ensemble	9.0±1.1	6.0±1.4	unreached	unreached
FedDistill	60.0±1.0	unreached	unreached	unreached
FedGen	7.3±0.4	5.0±0.8	10.6±0.9	4.6±1.2
FedFTG	43.5±37.5	15.6±16.5	12.6±0.5	9.4±2.3
FEDACK	4.6±3.8	2.3±0.9	2.33±1.25	1.67±0.94

不同平台上样本
的表征空间与表
征效果





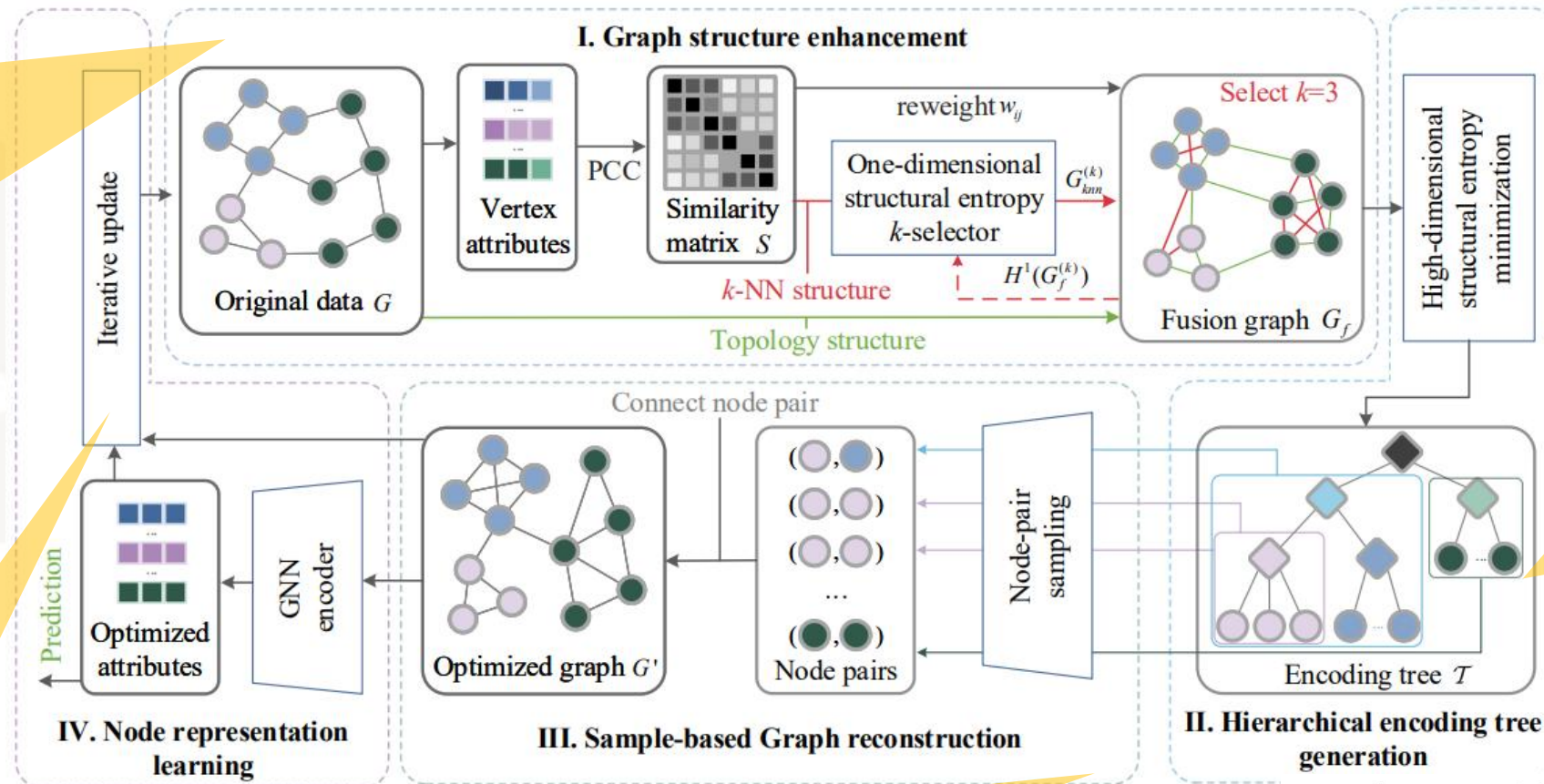
模型10：结构熵优化增强图学习模型

图增强模块

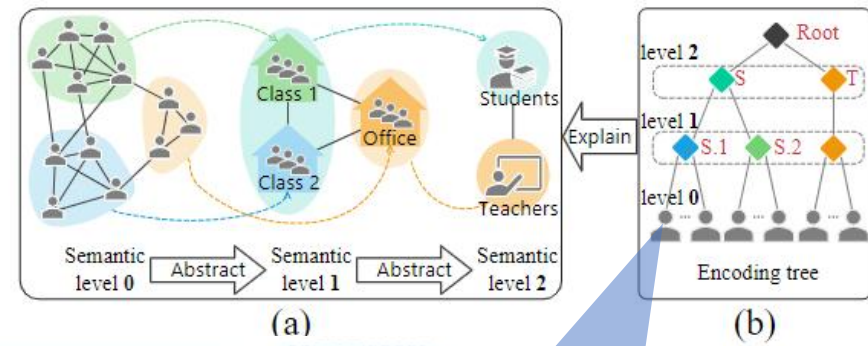
- 真实图结构常有噪且不完整。
- 图增强模块使图结构更加清晰，增大信噪比。
- 一维结构熵 k 选择器可以通过熵分析选择最合适的结构化参数

节点表示学习和迭代更新

- 图实质结构与节点表示（节点相似度）有关，需要迭代更新，同步优化二者



- 基于编码树节点采样的图重构：将节点结构熵转化成采样概率，通过在编码树上采样节点对来恢复有效的图结构。

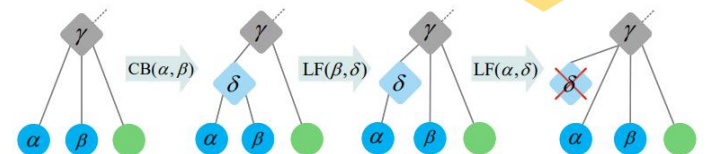


编码树与层次化抽象

- 复杂网络有固有多层社区结构
- 编码树旨在描述这种结构并利用信息熵确定嵌入层次结构的信息量

编码树构造

- 设计CB和LF算子。
- 通过贪心执行算子，获得使得图结构熵最小的编码树。





模型10：结构熵优化增强图学习模型

节点分类准确率对比分析

Table 1: Classification Accuracy (%) comparison, with improvement range of SE-GSL against the baselines. The best results are bolded and the second-best are underlined. Green denotes the outperformance percentage, while yellow denotes underperformance.

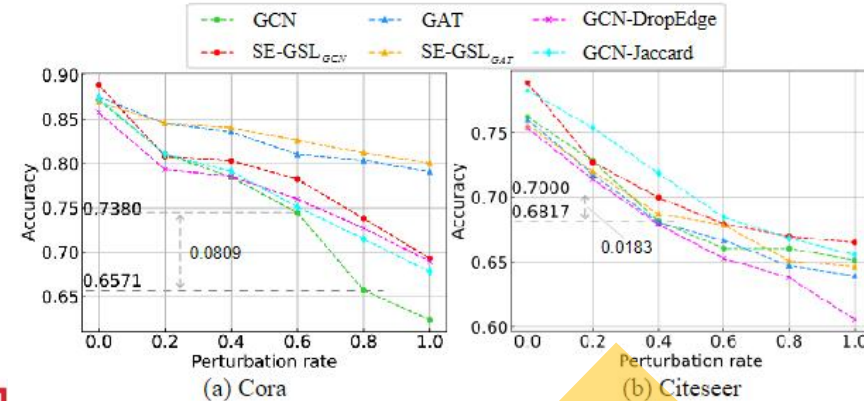
Method	Cora	Citeseer	Pubmed	PT	TW	Actor	Cornell	Texas	Wisconsin
GCN	87.26 $\pm$ 0.63	76.22 $\pm$ 0.71	87.46 $\pm$ 0.12	67.62 $\pm$ 0.21	62.46 $\pm$ 1.94	27.65 $\pm$ 0.55	49.19 $\pm$ 1.80	57.30 $\pm$ 2.86	48.57 $\pm$ 4.08
GAT	87.52 $\pm$ 0.69	76.04 $\pm$ 0.78	86.61 $\pm$ 0.15	68.76 $\pm$ 1.01	61.68 $\pm$ 1.20	27.77 $\pm$ 0.59	57.09 $\pm$ 6.32	58.10 $\pm$ 4.14	51.34 $\pm$ 4.78
GCNII	87.57 $\pm$ 0.87	75.47 $\pm$ 1.01	<u>88.64$\pm$0.23</u>	68.93 $\pm$ 0.93	65.17 $\pm$ 0.47	30.66 $\pm$ 0.66	58.76 $\pm$ 7.11	55.36 $\pm$ 6.45	51.96 $\pm$ 4.36
Grand	87.93 $\pm$ 0.71	77.59 $\pm$ 0.85	86.14 $\pm$ 0.98	69.80 $\pm$ 0.75	<u>66.79$\pm$0.22</u>	29.80 $\pm$ 0.60	57.21 $\pm$ 2.48	56.56 $\pm$ 1.53	52.94 $\pm$ 3.36
Mixhop	85.71 $\pm$ 0.85	75.94 $\pm$ 1.00	87.31 $\pm$ 0.44	66.48 $\pm$ 0.30	62.34 $\pm$ 0.22	33.72 $\pm$ 0.76	64.47 $\pm$ 4.78	63.16 $\pm$ 6.28	72.12 $\pm$ 3.34
Dropege	86.32 $\pm$ 1.09	76.12 $\pm$ 1.32	87.58 $\pm$ 0.34	68.49 $\pm$ 0.91	65.24 $\pm$ 1.45	30.10 $\pm$ 0.71	58.94 $\pm$ 5.95	59.20 $\pm$ 5.43	60.45 $\pm$ 4.48
Geom-GCN-P	84.93	75.14	88.09	-	-	31.63	60.81	67.57	64.12
Geom-GCN-S	85.27	74.71	84.75	-	-	30.30	55.68	59.73	56.67
GDC	87.17 $\pm$ 0.36	76.13 $\pm$ 0.53	88.08 $\pm$ 0.16	66.14 $\pm$ 0.54	64.14 $\pm$ 1.40	28.74 $\pm$ 0.76	59.46 $\pm$ 4.35	56.42 $\pm$ 3.99	48.30 $\pm$ 4.29
GEN	87.84 $\pm$ 0.69	<u>78.77$\pm$0.88</u>	86.13 $\pm$ 0.41	<u>71.62$\pm$0.78</u>	65.16 $\pm$ 0.77	36.69 $\pm$ 1.02	65.57 $\pm$ 6.74	73.38 $\pm$ 6.65	54.90 $\pm$ 4.73
H ₂ GCN-2	87.81 $\pm$ 1.35	76.88 $\pm$ 1.77	89.59$\pm$0.33	68.15 $\pm$ 0.30	63.33 $\pm$ 0.77	35.62 $\pm$ 1.30	82.16$\pm$6.00	<u>82.16$\pm$5.28</u>	<u>85.88$\pm$4.22</u>
SE-GSL	87.93 $\pm$ 1.24	<u>77.63$\pm$1.65</u>	88.16 $\pm$ 0.76	71.91$\pm$0.66	66.99$\pm$0.93	<u>36.34$\pm$2.07</u>	75.21 $\pm$ 3.69	82.49$\pm$3.34	86.27$\pm$2.23
Improvement	0.00~3.00	-1.14~2.92	-1.43~3.41	0.29~5.77	0.20~5.31	-0.35~8.69	-6.95~26.02	0.33~27.13	0.39~37.97

- 相比Baseline方法和SOTA方法有普遍提升

Table 2: Classification accuracy(%) of SE-GSL and corresponding backbones. Wisc. is short for Wisconsin.

Method	Actor	TW	Texas	Wisc.	Improvement
SE-GSL _{GCN}	35.03	66.88	75.68	79.61	↑ 5.20~31.04
SE-GSL _{SAGE}	36.20	66.92	82.49	86.27	↑ 0.25~6.79
SE-GSL _{GAT}	32.46	63.57	74.59	78.82	↑ 4.69~27.48
SE-GSL _{APPNP}	36.34	66.99	81.28	83.14	↑ 2.01~12.16

- 针对不同Backbone图神经网络，提升显著

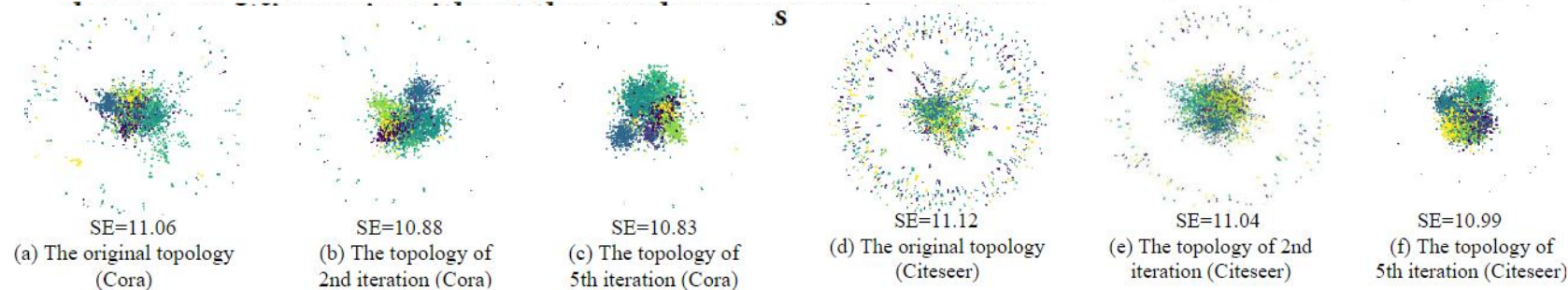
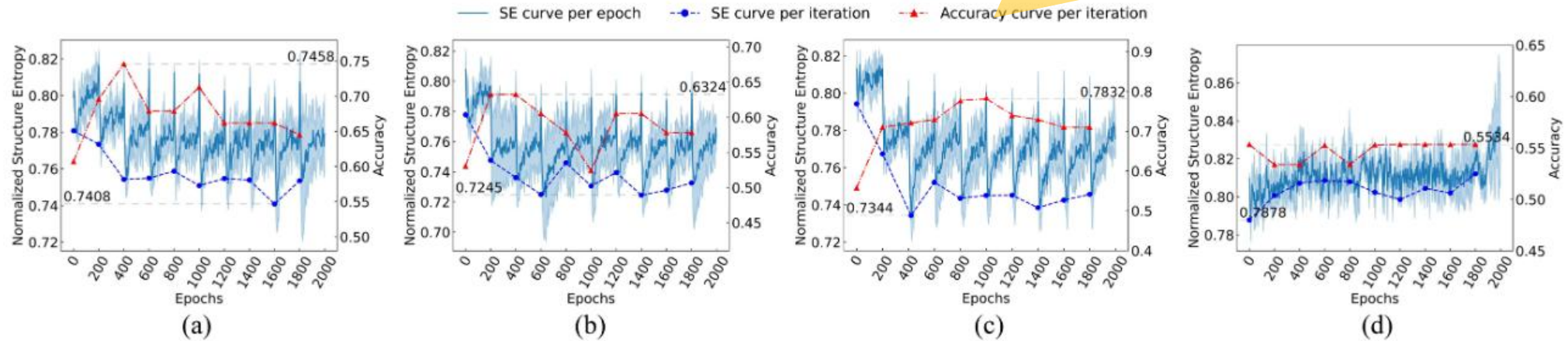


- 在人工噪声数据集上，有效减缓性能下降，提升模型鲁棒性

模型10：结构熵优化增强图学习模型

可视化分析

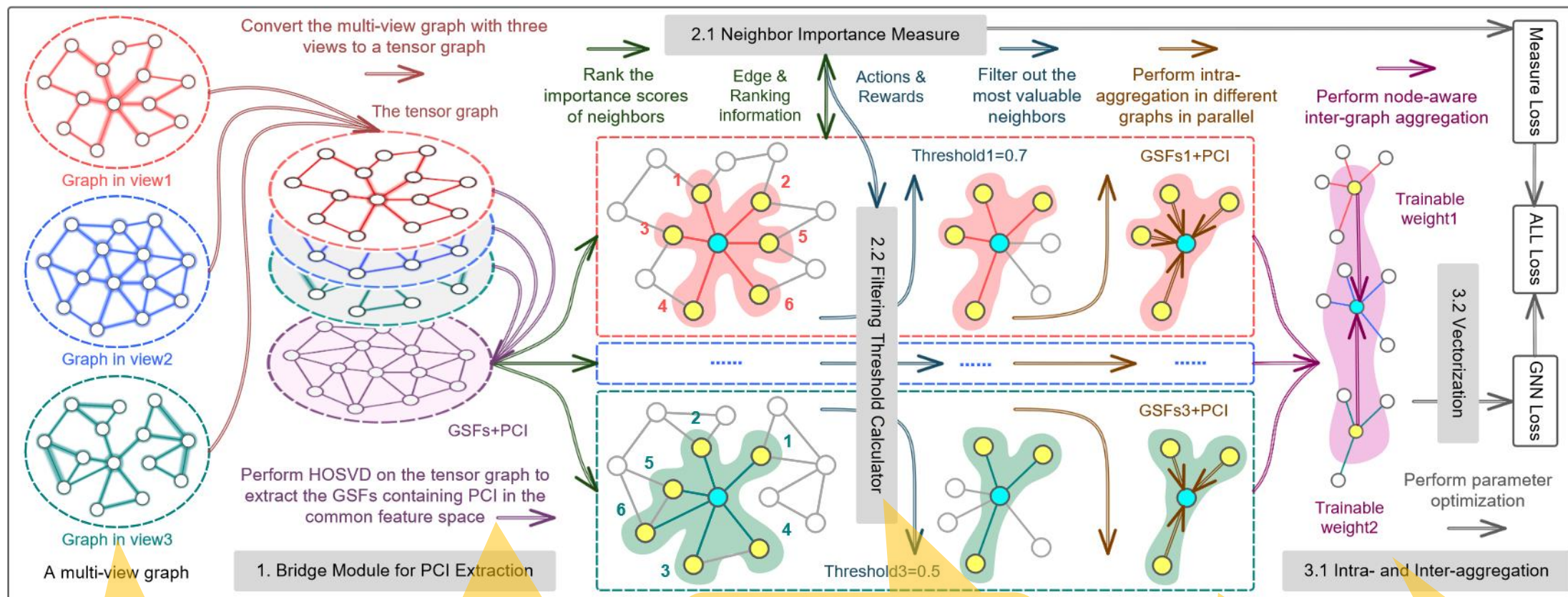
- 正常情况下，图神经网络的优化更新，会使图结构熵增大
- 我们的方法有效抑制了结构熵增加，减小了图结构的不确定性



- 结构学习过程中，社区结构更加清晰，离群点减少，最终获得更适合图神经网络学习的结构



模型11：增强的多视角张量图神经网络



多视图图表示

- 转化为张量图

PCI抽取桥接模块

- HOSVD算法
- 抽取潜在关联信息和图结构信息

强化学习引导的邻居过滤器

- 邻居重要性评价
- 过滤阈值计算

图内特征聚合

图间特征聚合



模型11：增强的多视角张量图神经网络

邻居重要性度量

FNN节点标签
预测器

图中的边
权信息

二维范式节
点距离

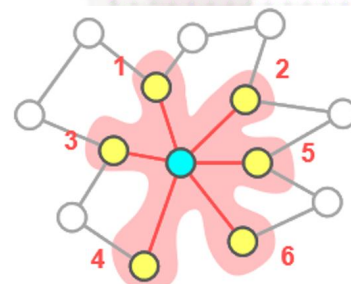
$$IMP_j^{(l)}(k, k') = |A_{i,j}(k, k')| \otimes (1 - DIST_j^{(l)}(k, k')),$$

$$DIST_j^{(l)}(k, k') = NORM \left(\left\| \sigma \left(FNN_j^{(l)} \left(\mathbf{F}_{i,j}^{(l)}(k) \right) \right) - \sigma \left(FNN_j^{(l)} \left(\mathbf{F}_{i,j}^{(l)}(k') \right) \right) \right\|_2 \right),$$

$$\mathcal{L}_{edge}^{(l)} = - \sum_{\mathcal{G}_{train}} \sum_{j=1}^V \sum_{k=1}^{|N|} \mathbf{X}(k) \log \left(\sigma \left(FNN_j^{(l)} \left(\mathbf{F}_{i,j}^{(l)}(k) \right) \right) \right).$$

FNN与图神经网络
协同优化

计算所有邻居节点
的重要性并排序



强化学习

动作

动作描述如何计算阈值更新， ϵ -greedy 算法优化策略

奖励

相邻迭代轮船间，邻居的平均重要性差异决定奖励

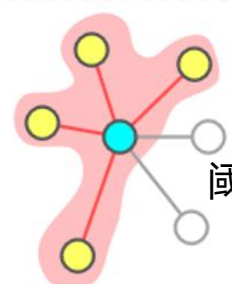
$$AVG_j^{(l)[p]} = \sum_{\mathcal{G}_{train}} \frac{\sum_{A_{i,j}^{(l^*)[p]}(k, k') > 0} IMP_j^{(l)}(k, k')}{\sum_{n_k \in N} T_j^{(l)[p]} |N(k)|},$$

$$REW_j^{(l)[p]} = \begin{cases} -1, & AVG_j^{(l)[p]} \leq AVG_j^{(l)[p-1]} \\ +1, & AVG_j^{(l)[p]} > AVG_j^{(l)[p-1]} \end{cases}.$$

终止
条件

$$\left| \sum_{p-10}^p REW_j^{(l)[p]} \right| \leq 1.$$

过滤器阈值
计算器

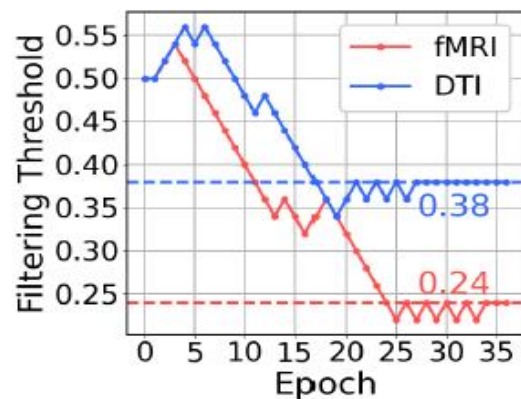


阈值=0.7

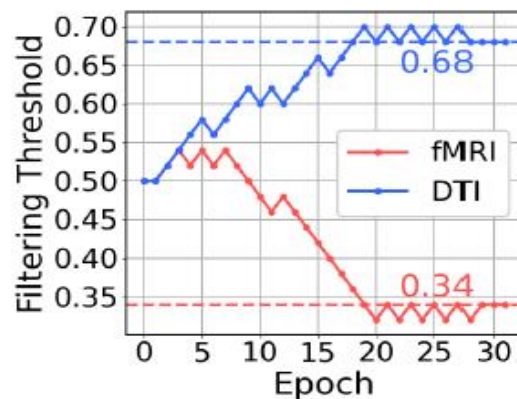


模型11：增强的多视角张量图神经网络

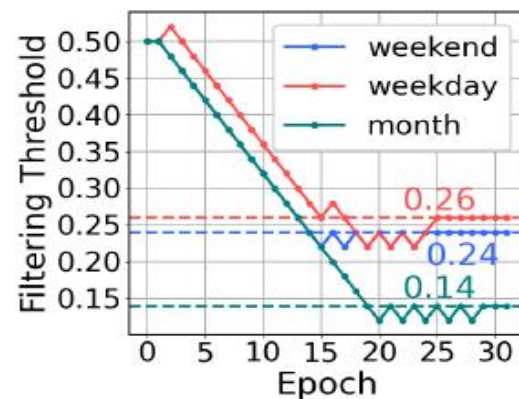
邻居过滤器参数变化



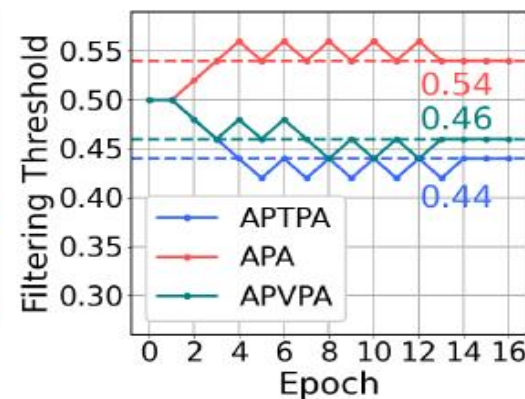
(a) HIV



(b) BP

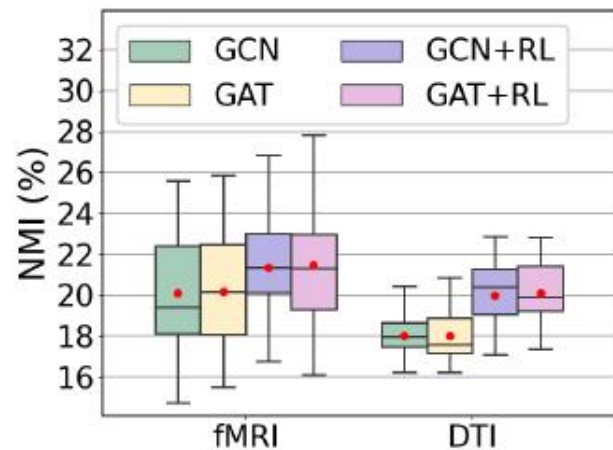


(c) BikeDC

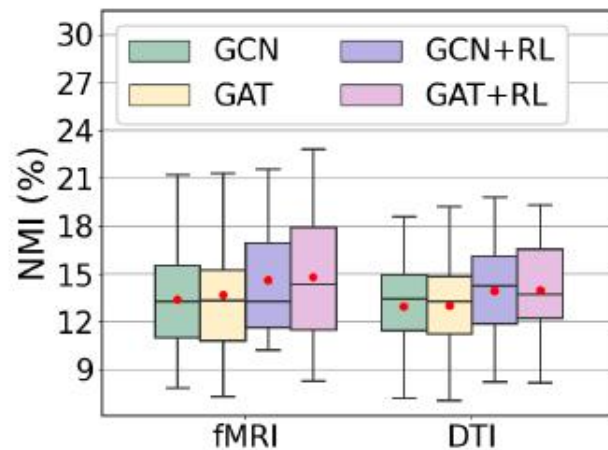


(d) DBLP

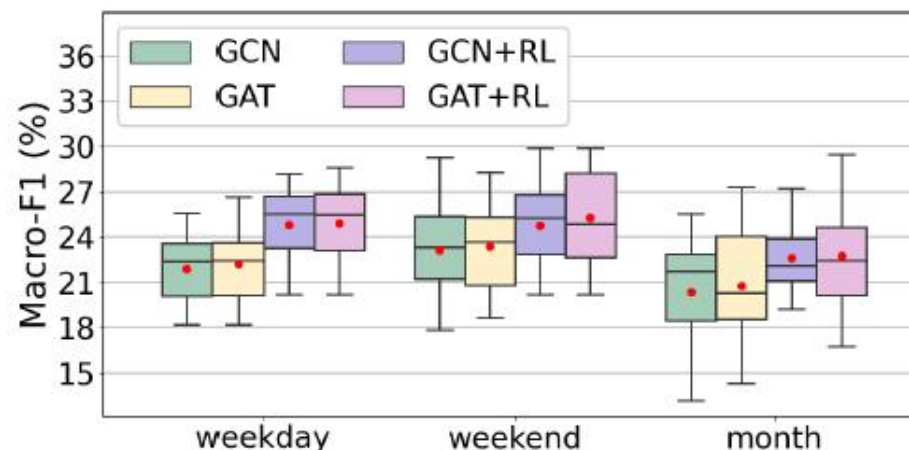
强化学习引导GNN与原生GNN



(a) HIV



(b) BP



(c) BikeDC



模型11：增强的多视角张量图神经网络

分类任务表现

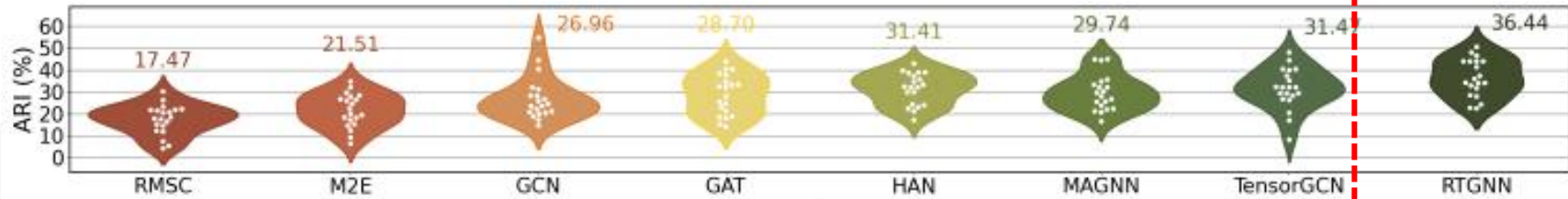
Methods		Train%	HIV		BP		BikeDC		DBLP	
			Ma-F1	Mi-F1	Ma-F1	Mi-F1	Ma-F1	Mi-F1	Ma-F1	Mi-F1
Graph-based Methods	CoRegSC	20%	49.34 ± 03.17	50.51 ± 04.72	46.31 ± 08.02	54.00 ± 02.63	20.47 ± 04.67	35.29 ± 03.72	80.63 ± 00.28	81.61 ± 00.93
		60%	59.19 ± 07.16	60.14 ± 09.56	54.13 ± 07.24	57.82 ± 04.36	29.05 ± 03.75	43.52 ± 05.99	82.90 ± 00.15	83.88 ± 00.33
	MultiNMF	20%	48.85 ± 06.94	49.19 ± 03.81	53.84 ± 09.24	58.22 ± 05.60	22.24 ± 02.75	37.35 ± 02.29	78.43 ± 00.43	79.20 ± 00.70
		60%	53.36 ± 08.89	54.71 ± 04.19	55.10 ± 07.58	58.69 ± 06.22	26.21 ± 02.77	42.35 ± 04.40	82.02 ± 00.20	82.57 ± 00.19
	RMSC	20%	51.24 ± 06.93	51.69 ± 06.47	43.95 ± 08.19	53.33 ± 03.29	20.38 ± 03.58	34.11 ± 02.99	83.46 ± 00.51	83.82 ± 00.18
		60%	62.50 ± 09.54	63.12 ± 06.07	53.96 ± 06.98	57.39 ± 03.79	30.57 ± 06.14	45.88 ± 08.64	84.89 ± 00.27	85.03 ± 00.32
	AMGL	20%	48.26 ± 08.68	49.48 ± 03.06	33.82 ± 10.68	51.11 ± 04.24	19.21 ± 05.97	35.58 ± 05.17	80.52 ± 00.21	81.38 ± 00.91
		60%	56.60 ± 09.54	58.76 ± 08.47	34.28 ± 09.31	52.17 ± 05.49	25.84 ± 04.85	43.52 ± 06.55	84.50 ± 00.37	85.11 ± 00.13
	M2E	20%	53.98 ± 03.03	56.25 ± 05.13	51.47 ± 07.45	55.55 ± 03.71	22.85 ± 05.22	38.23 ± 04.36	83.45 ± 00.12	83.82 ± 00.18
		60%	63.75 ± 03.75	65.00 ± 05.00	54.70 ± 06.98	56.95 ± 05.97	30.21 ± 03.75	45.88 ± 06.85	85.74 ± 00.05	86.41 ± 00.09
GNN-based Methods	GCN	20%	55.60 ± 03.03	56.56 ± 03.93	50.54 ± 08.94	55.77 ± 03.50	22.75 ± 03.28	39.11 ± 03.49	89.12 ± 00.17	89.94 ± 00.16
		60%	65.73 ± 04.85	67.50 ± 04.67	57.24 ± 07.12	59.56 ± 05.84	30.73 ± 03.57	48.23 ± 04.40	90.17 ± 00.05	90.40 ± 00.12
	GAT	20%	55.98 ± 04.73	56.25 ± 02.82	50.58 ± 06.98	55.55 ± 02.81	23.83 ± 04.33	39.70 ± 03.54	91.06 ± 00.03	91.19 ± 00.07
		60%	65.69 ± 04.19	68.75 ± 04.14	57.57 ± 05.88	60.00 ± 05.07	31.98 ± 04.73	48.23 ± 06.33	91.94 ± 00.18	92.35 ± 00.09
	HAN	20%	58.90 ± 05.04	60.00 ± 03.64	56.18 ± 06.31	59.33 ± 04.97	25.40 ± 02.26	40.00 ± 03.52	91.92 ± 00.08	92.42 ± 00.07
		60%	69.55 ± 04.23	70.62 ± 04.88	60.29 ± 02.50	62.60 ± 03.47	33.07 ± 05.16	50.00 ± 03.94	92.10 ± 00.04	93.56 ± 00.12
	MAGNN	20%	56.66 ± 04.75	58.43 ± 03.14	54.01 ± 05.91	57.11 ± 02.63	24.04 ± 03.53	37.05 ± 04.20	93.30 ± 00.27	93.82 ± 00.32
		60%	69.84 ± 04.59	71.25 ± 05.00	59.74 ± 06.62	62.17 ± 05.16	32.10 ± 06.47	49.41 ± 06.55	94.35 ± 00.06	94.58 ± 00.41
	TensorGCN	20%	59.02 ± 04.38	60.31 ± 02.81	53.07 ± 07.54	56.88 ± 04.68	25.30 ± 02.68	40.29 ± 03.73	92.40 ± 00.24	93.08 ± 00.11
		60%	70.98 ± 04.36	72.50 ± 05.00	60.29 ± 08.42	61.73 ± 08.43	33.99 ± 08.56	50.58 ± 09.18	93.30 ± 00.13	93.99 ± 00.09
Our Methods	RTGNN -mean	20%	64.48 ± 05.97	66.87 ± 05.44	57.13 ± 09.55	59.77 ± 05.74	31.84 ± 06.01	44.41 ± 03.82	94.13 ± 00.07	94.58 ± 00.06
		60%	72.23 ± 08.59	73.75 ± 08.29	63.86 ± 05.51	64.34 ± 05.43	45.35 ± 09.68	54.11 ± 10.12	95.19 ± 00.41	95.48 ± 00.36
	RTGNN -att	20%	65.26 ± 07.05	67.18 ± 05.45	58.09 ± 03.15	60.44 ± 02.39	33.48 ± 03.98	45.58 ± 01.97	94.30 ± 00.12	94.76 ± 00.30
		60%	73.07 ± 06.88	74.37 ± 05.89	65.67 ± 05.50	66.08 ± 05.43	46.27 ± 08.05	55.88 ± 05.42	95.21 ± 00.16	95.56 ± 00.05
	RTGNN	20%	65.74 ± 05.75	67.50 ± 06.43	60.70 ± 04.62	61.11 ± 04.44	35.22 ± 07.25	45.88 ± 03.76	94.50 ± 00.13	94.91 ± 00.15
		60%	75.14 ± 08.18	76.25 ± 08.29	66.12 ± 06.00	66.52 ± 06.16	48.25 ± 05.10	57.05 ± 04.59	95.42 ± 00.26	95.62 ± 00.44
	Gain	20%	6.72 ↑	7.19 ↑	4.52 ↑	1.78 ↑	9.82 ↑	5.59 ↑	1.20 ↑	1.09 ↑
		60%	4.16 ↑	3.75 ↑	5.83 ↑	3.92 ↑	14.26 ↑	6.47 ↑	1.07 ↑	1.04 ↑



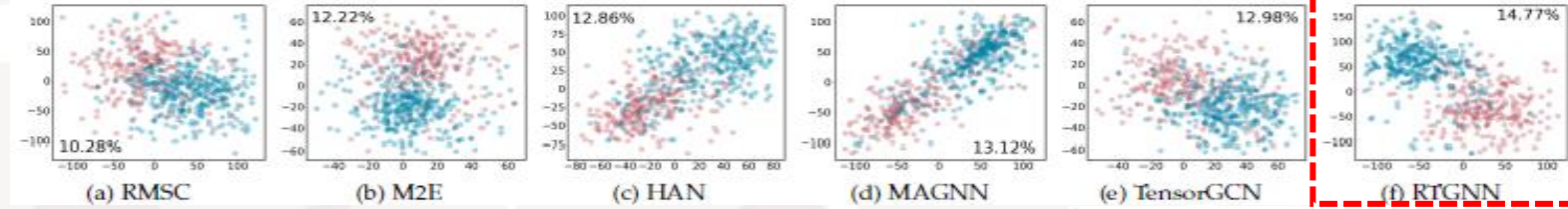
模型11：增强的多视角张量图神经网络

聚类任务表现

BIKEBC数据ARI



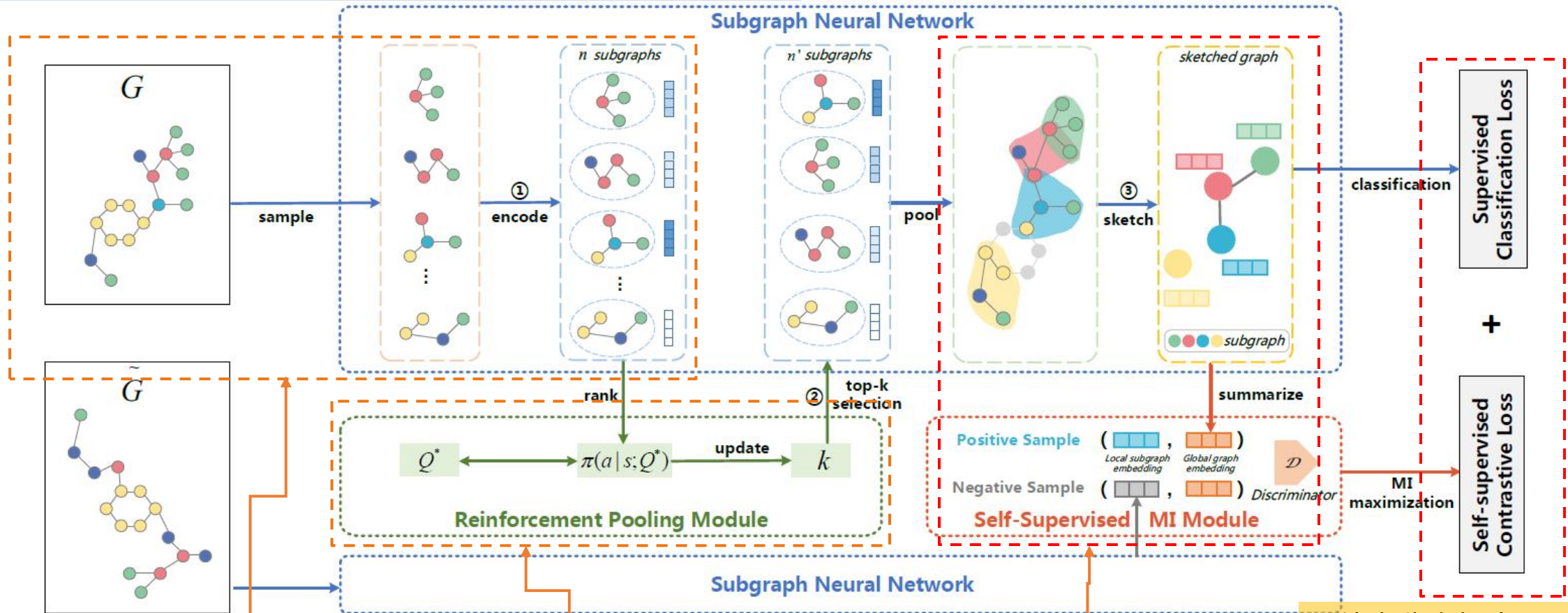
PROTEINS数据集聚类结果可视化



Methods	PROTEINS			
	Ma-F1		Mi-F1	
	20%	60%	20%	60%
GCN	71.06 ± 0.24	74.60 ± 0.69	72.16 ± 0.21	76.25 ± 0.67
GAT	71.32 ± 0.59	75.16 ± 0.85	72.19 ± 0.49	75.58 ± 0.94
HAN	71.98 ± 0.59	76.04 ± 0.51	72.58 ± 0.62	76.83 ± 0.50
MAGNN	72.19 ± 0.59	75.97 ± 1.25	72.66 ± 0.65	76.58 ± 1.24
TensorGCN	72.19 ± 0.60	75.93 ± 0.69	72.75 ± 0.58	76.50 ± 0.67
RTGNN	73.28 ± 0.48	77.08 ± 1.18	73.81 ± 0.54	77.79 ± 1.19
Gain	1.51 ↑	1.37 ↑	1.46 ↑	1.25 ↑

Methods		Metrics	HIV	BP
Graph-based Methods	CoRegSC	NMI	18.51 $\pm$ 06.53	12.22 $\pm$ 06.62
		ARI	11.01 $\pm$ 06.32	07.53 $\pm$ 08.68
	MultiNMF	NMI	14.21 $\pm$ 07.68	14.17 $\pm$ 06.26
		ARI	08.33 $\pm$ 06.49	11.58 $\pm$ 07.11
	RMSC	NMI	15.85 $\pm$ 08.54	13.65 $\pm$ 08.57
		ARI	07.73 $\pm$ 04.06	10.50 $\pm$ 05.02
GNN-based Methods	AMGL	NMI	09.62 $\pm$ 07.72	09.10 $\pm$ 07.46
		ARI	06.90 $\pm$ 07.05	05.27 $\pm$ 09.58
	M2E	NMI	19.52 $\pm$ 08.46	13.33 $\pm$ 10.84
		ARI	07.86 $\pm$ 05.12	12.55 $\pm$ 10.12
	GCN	NMI	20.67 $\pm$ 09.70	13.49 $\pm$ 05.18
		ARI	11.61 $\pm$ 07.42	11.88 $\pm$ 04.47
Our Method	GAT	NMI	20.16 $\pm$ 07.17	13.65 $\pm$ 08.57
		ARI	11.83 $\pm$ 05.35	12.59 $\pm$ 09.05
	HAN	NMI	23.73 $\pm$ 07.17	18.51 $\pm$ 08.86
		ARI	17.59 $\pm$ 14.75	13.34 $\pm$ 08.19
	MAGNN	NMI	19.30 $\pm$ 10.96	16.20 $\pm$ 06.08
		ARI	20.33 $\pm$ 11.98	13.68 $\pm$ 07.63
TensorGCN	NMI	24.38 $\pm$ 13.50	17.27 $\pm$ 06.93	
	ARI	19.44 $\pm$ 09.14	13.98 $\pm$ 07.93	
Our Method	RTGNN	NMI	35.02 $\pm$ 11.45	19.70 $\pm$ 06.58
		ARI	30.86 $\pm$ 10.81	17.26 $\pm$ 09.80
Gain		NMI	10.64 $\uparrow$	1.19 $\uparrow$
		ARI	10.53 $\uparrow$	3.28 $\uparrow$

模型12：增强的显著性子图挖掘框架



子图抽样与编码：
• BFS提取子图
• 基于GNN的子图编码器

• Top-k采样：选择有显著模式的子图
• 强化学习模块优化自适应池化率k

• 利用子图重建图特征速写
• 子图间注意力机制
• 由自监督的互信息模块引导

• 结合分类损失和自监督损失综合优化

$$\mathcal{L} = \mathcal{L}_{\text{Classify}} + \beta \sum_{G \in \mathcal{G}} \mathcal{L}_{\text{MI}}^G + \lambda \|\Theta\|^2$$

强化汇聚模块



Qingyun sun, etc., SUGAR Subgraph Neural Network with Reinforcement Pooling and Self-Supervised Mutual Information Mechanism. WWW 2021; 谷歌学术引用 79

模型12：增强的显著性子图挖掘框架

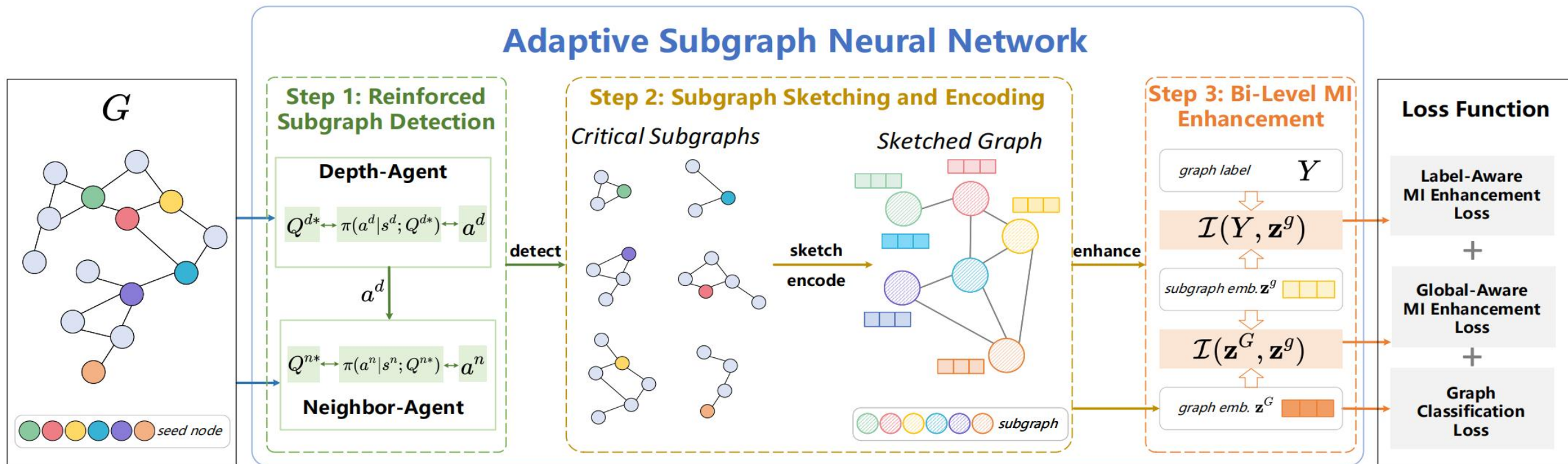


Fig. 2: An illustration of the AdaSNN architecture. Step-1: the Reinforced Subgraph Detection Module aims to detect critical subgraphs $\{g_i\}$ for an input graph G . Step-2: the critical subgraphs $\{g_i\}$ are used to contract a sketched graph G^{ske} and encoded into subgraph representations $\{z_i^g\}$. Step-3: the Bi-Level MI Enhancement Mechanism further enhances the subgraph representations and uses them for classification.

模型12：增强的显著性子图挖掘框架

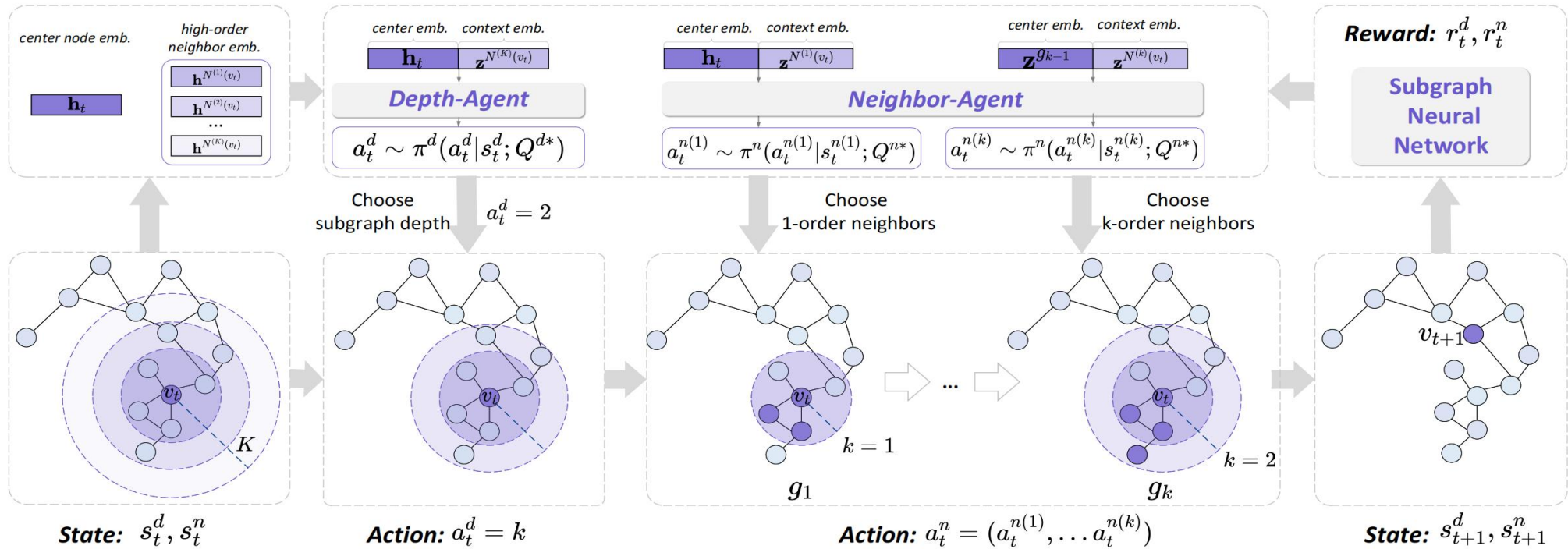


Fig. 3: An illustration of the reinforced subgraph detection module. For a given center node, a critical subgraph is detected in two phases: (1) For a center node v_t , Depth-Agent generates an action a_t^d by policy π^d to specify the number of hops for the current subgraph; (2) With the specified depth $k = a_t^d$, Neighbor-Agent generates actions $(a_t^{n(1)}, a_t^{n(2)}, \dots, a_t^{n(k)})$ by policy π^n to sample the member nodes of subgraph within k -hop neighbors of node v_t hop by hop.

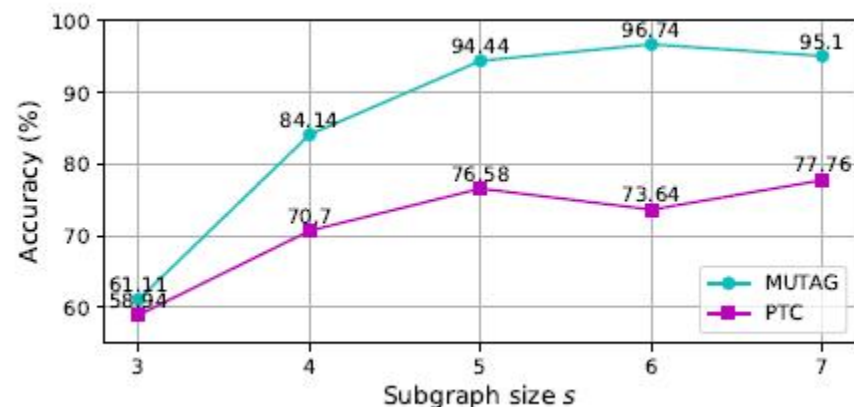
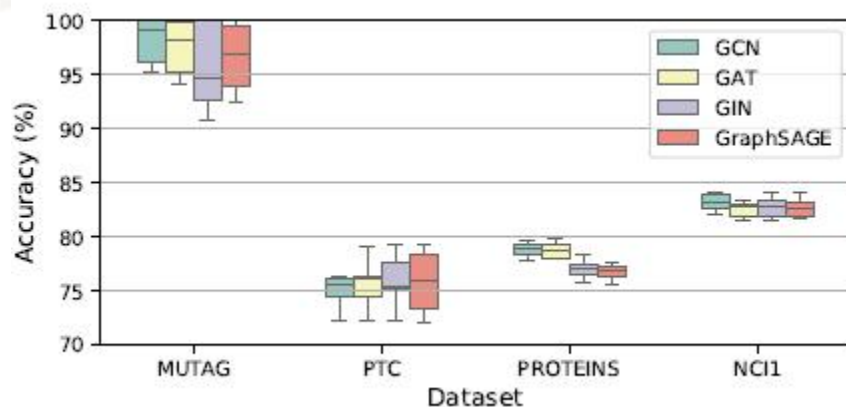


模型12：增强的显著性子图挖掘框架

分类任务表现

Method	Dataset						Avg. Rank
	MUTAG	PTC	PROTEINS	D&D	NCI1	NCI109	
WL [40]	82.05±0.36 (13)	-	-	79.78±0.36 (5)	82.19±0.18 (6)	82.46±0.24 (3)	6.75
GK [41]	83.50±0.60 (12)	59.65±0.31 (9)	-	74.62±0.12 (13)	-	-	11.33
DGK [57]	87.44±2.72 (9)	60.08±2.55 (8)	75.68±0.54 (12)	-	80.31±0.46 (9)	80.32±0.33 (7)	9.00
PATCHY-SAN [32]	92.63±4.21 (3)	62.29±5.68 (7)	75.89±2.76 (11)	77.12±2.41 (10)	78.59±1.89 (10)	-	8.20
ECC [42]	89.44 (6)	-	-	73.65 (14)	83.80 (2)	81.87 (4)	6.50
GIN [55]	89.40±5.60 (7)	64.60±7.00 (5)	76.20±2.80 (10)	-	82.70±1.70 (5)	-	6.75
GCAPS-CNN [50]	-	66.01±5.91 (4)	76.40±4.17 (7)	77.62±4.99 (9)	82.72±2.38 (4)	81.12±1.28 (6)	6.00
CapsGNN [54]	86.67±6.88 (10)	-	76.28±3.63 (8)	75.38±4.17 (12)	78.35±1.55 (11)	-	10.25
AWE [20]	87.87±9.76 (8)	-	-	71.51±4.02 (15)	-	-	11.50
S2S-N2N-PP [45]	89.86±1.10 (5)	64.54±1.10 (6)	76.61±0.50 (4)	-	83.72±0.40 (3)	83.64±0.30 (2)	4.00
NEST [58]	91.85±1.57 (4)	67.42±1.83 (3)	76.54±0.26 (6)	78.11±0.36 (8)	81.59±0.46 (8)	81.72±0.41 (5)	5.67
MA-GCNN [35]	93.89±5.24 (2)	71.76±6.33 (2)	79.35±1.74 (2)	81.48±1.03 (3)	81.77±2.36 (7)	-	3.20
SortPool [63]	85.83±1.66 (11)	58.59±2.47 (10)	75.54±0.94 (13)	79.37±0.94 (6)	74.44±0.47 (13)	-	10.60
DiffPool [61]	-	-	76.25 (9)	80.64 (4)	-	-	6.50
gPool [15]	-	-	77.68 (3)	82.43 (2)	-	-	2.50
EigenPool [29]	-	-	76.60 (5)	78.60 (7)	77.00 (12)	74.90 (8)	8.00
SAGPool [23]	-	-	71.86±0.97 (14)	76.45±0.97 (11)	67.45±1.11 (14)	74.06±0.78 (9)	12.00
SUGAR (Ours)	96.74±4.55(1)	77.53±2.82(1)	81.34±0.93(1)	84.03±1.33(1)	84.39±1.63(1)	84.82±0.81(1)	1.00

SUGER在不同node-encoder和子图大小下的任务表现

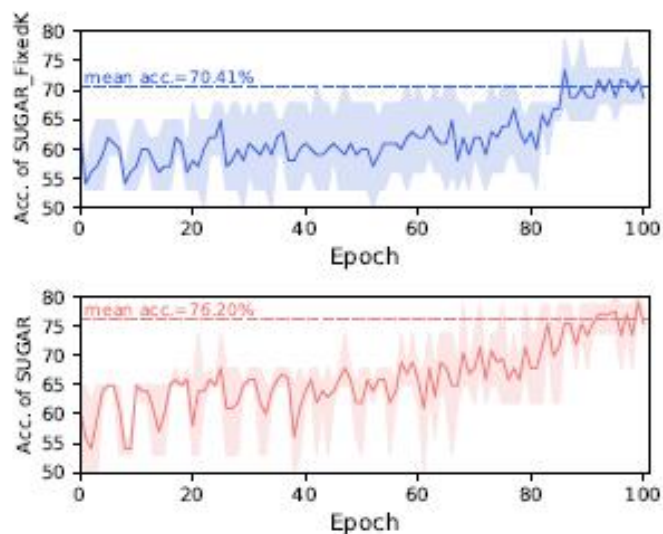




模型12：增强的显著性子图挖掘框架

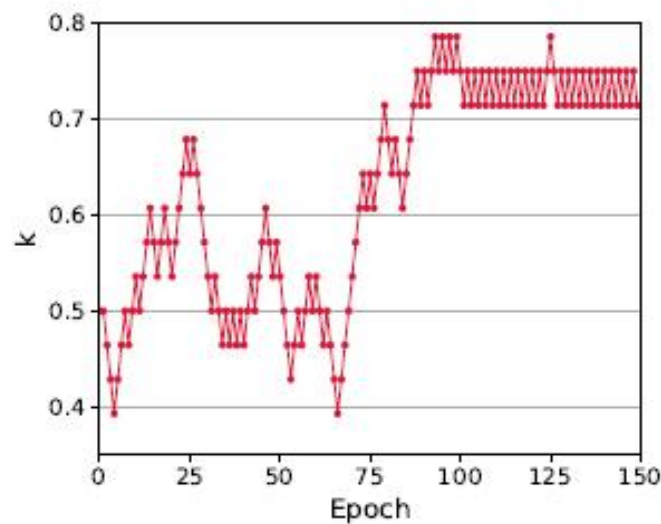
强化学习模块实验

固定k-自适应k模型表现



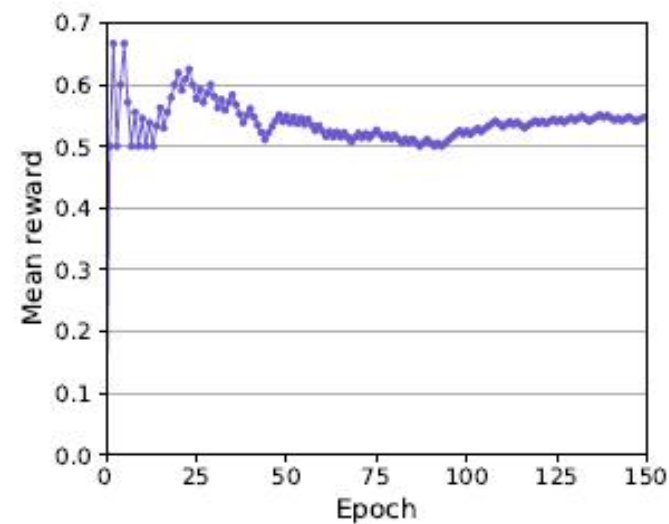
(a) Training process of SUGAR-FixedK and SUGAR on PTC.

K变化曲线



(b) Updating process of k on PTC.

平均奖励变化曲线



(c) Learning curve of RL on PTC.

模型13：结构熵优化的群体对抗

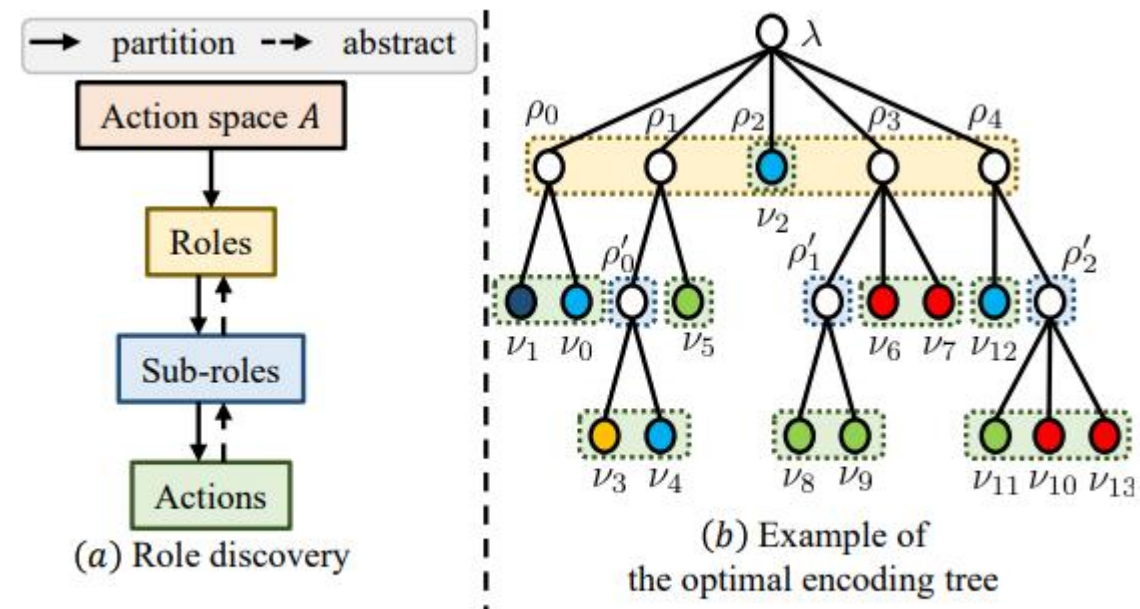
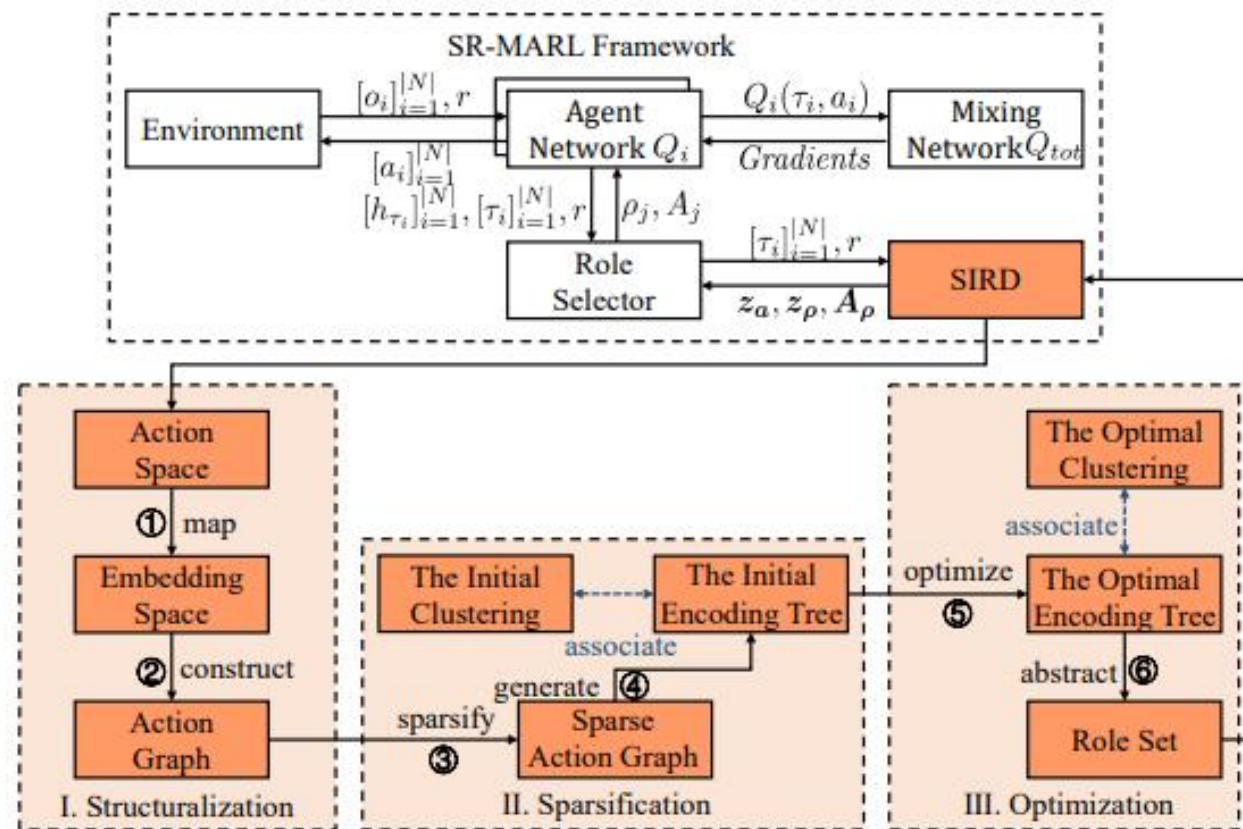


Figure 3: The role discovery on the optimal encoding tree.

Figure 1: The overall framework of the SR-MARL.

模型13：结构熵优化的群体对抗

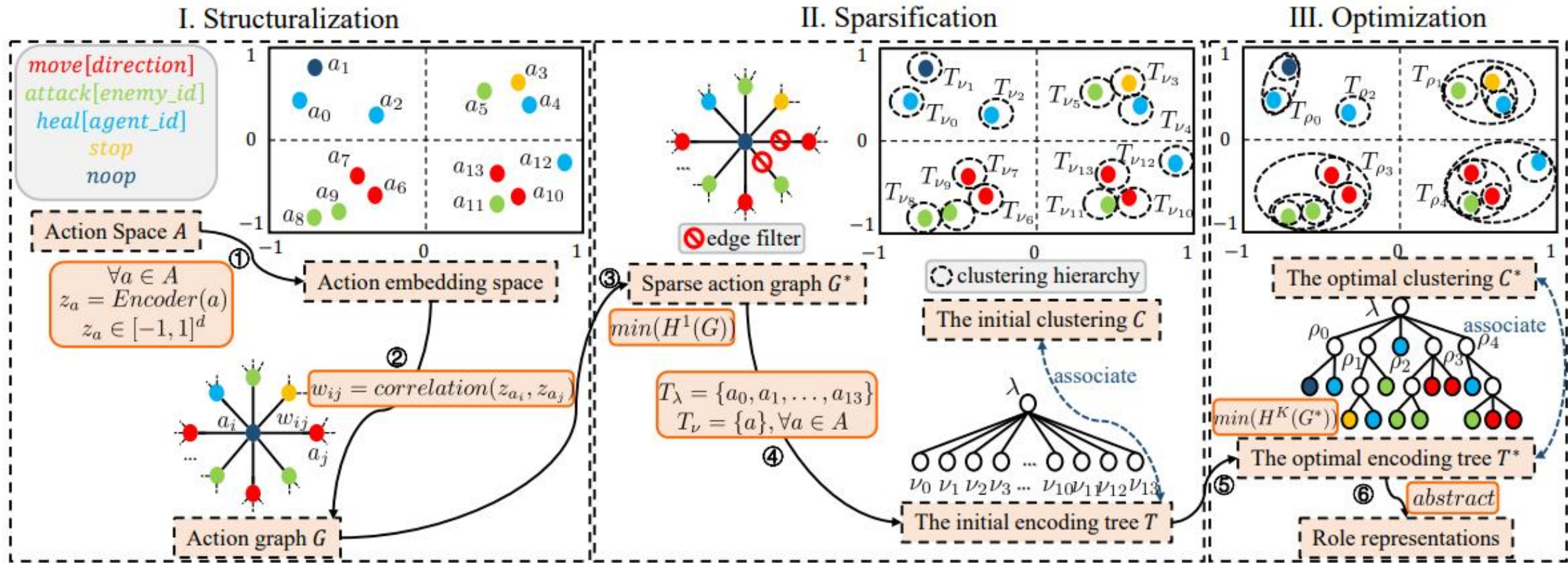


Figure 2: The structural information principles-based role discovery.



模型13：结构熵优化的群体对抗

Categories	Easy	Hard	Super Hard
COMA	16.67 $\pm$ 22.73	4.51 $\pm$ <u>9.58</u>	-
IQL	52.50 $\pm$ 40.69	73.44 $\pm$ 24.85	10.55 $\pm$ 18.49
VDN	85.01 $\pm$ 17.22	71.49 $\pm$ 18.78	71.10 $\pm$ 27.23
QMIX	<u>98.44 $\pm$ 2.10</u>	87.11 $\pm$ 18.58	70.31 $\pm$ 38.65
QTRAN	64.69 $\pm$ 36.79	58.20 $\pm$ 45.37	16.80 $\pm$ 20.61
QPLEX	96.88 $\pm$ 5.04	<u>89.85 $\pm$ 11.35</u>	84.77 $\pm$ 10.76
RODE	93.47 $\pm$ 10.19	88.44 $\pm$ 20.96	<u>92.71 $\pm$ 9.20</u>
SR-MARL	98.61 $\pm$ 1.75	95.31 $\pm$ 6.63	95.71 $\pm$ 3.10
Improvements(%)/Reductions(%)			
Average	$\uparrow$ 0.17	$\uparrow$ 6.08	$\uparrow$ 3.24
Deviation	$\downarrow$ 16.67	$\downarrow$ 30.80	$\downarrow$ 66.30

Table 1: Summary of the test win rates under different map categories: “average value $\pm$ standard deviation” and “improvements/reductions” (%). Bold: the best performance under each category, underline: the second performance.

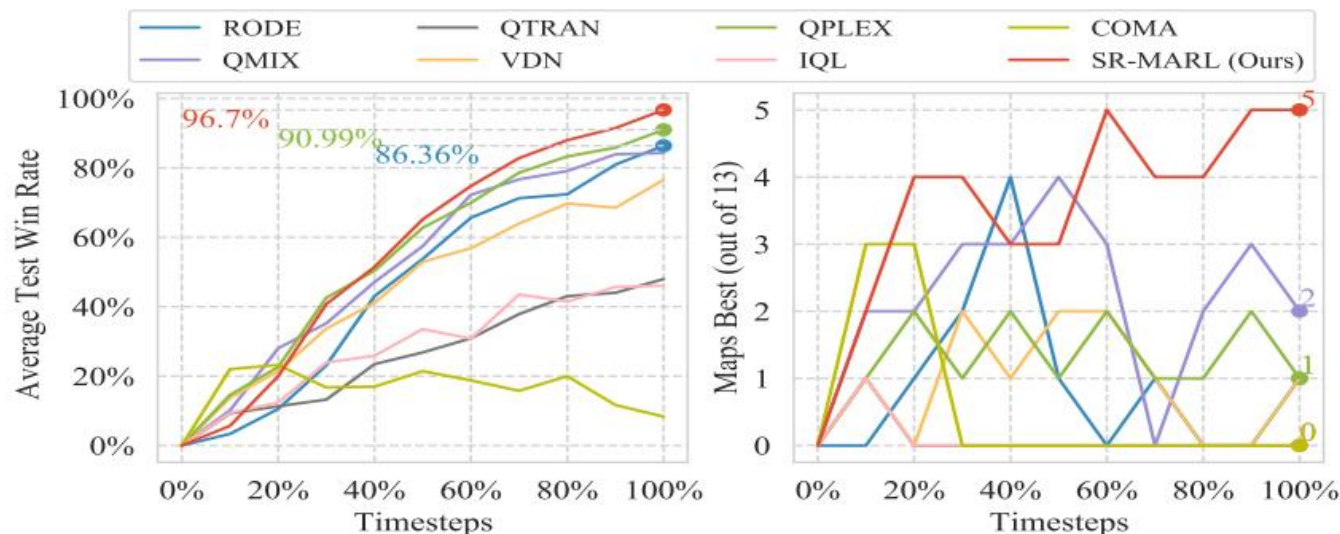
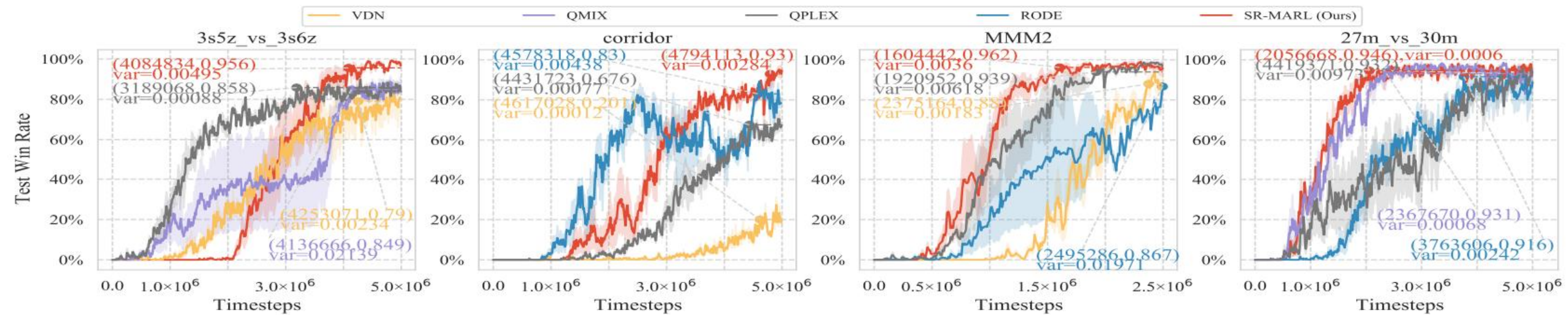


Figure 4: (left) The average test win rates across all 13 maps; (right) the number of maps (out of 13) where the algorithm’s average test win rate is the highest.



模型13：结构熵优化的群体对抗



系统应用：社交网络水军监测平台

社交机器人账号自动识别系统

Welcome,admin

平台: Twitter 是否人工标注: 请选择 标签: 请选择 机器得分: 请输入账号 查询 + 账号 + 批量导入

QuandtStefan

US选举 东南亚 俄乌冲突 欧盟问题

粉丝: 5563 关注: 756 发帖: 1569 分析状态: 分析成功 机器人账号相似度: 3.7 / 5.0 查看 更新 标注 删除

CarloMasala1

US选举 东南亚 俄乌冲突 欧盟问题

粉丝: 25137 关注: 805 发帖: 10449 分析状态: 分析成功 机器人账号相似度: 2.6 / 5.0 查看 更新 标注 删除

rollysms

US选举 欧盟问题

粉丝: 54 关注: 80 发帖: 1682 分析状态: 分析成功 机器人账号相似度: 3.0 / 5.0 查看 更新 标注 删除

chefzhenya

粉丝: 21930 关注: 299 发帖: 4111 分析状态: 分析成功 机器人账号相似度: 2.5 / 5.0 查看 更新 标注 删除

johnsweeneyroar

俄乌冲突 欧盟问题

粉丝: 65954 关注: 3684 发帖: 122314 分析状态: 分析成功 机器人账号相似度: 2.5 / 5.0 查看 更新 标注 删除

danistrubetskoy

共 64932 条 10条/页 1 2 3 4 5 6 ... 6494 前往 1 页

系统应用：社交网络水军监测平台

社交机器人账号自动识别系统

Welcome,admin

目标详情

好友列表

粉丝列表

发帖列表

返回上级

目标名称: QuandtStefan

主页地址:

发帖量: 1569

主页地址: <https://twitter.com/QuandtStefan>

点赞量: 5563 (最新200条帖子)

注册时间: 2016-02-29 06:03:06

转评量: 37 (最新200条帖子)

最近发帖时间: 2023-03-30 22:54:37

转评占比: 23.7% (最新200条帖子)

语种: 德语

简介: #StaatsbürgerInUniform de aus Überzeugung • Grundsatzreferent @TFK_Bundeswehr • Ex-@BAKS_Bund + PersRef @BAKS_President • Aus
#SachsenInBerlin • Hier privat

发帖规律统计(最新200条帖子)

发帖量

按小时

按周



词频统计

seiner
politisch fe
vad. dort
ohne
ckgabe
humanit
bw
uhr



系统应用：社交网络水军监测平台

☰ 社交机器人账号自动识别系统

Welcome,admin ▾

👁

📊

⚙

目标详情

好友列表

粉丝列表

发帖列表

返回上级

■ 基本信息





Stefan Quandt

粉丝 5563 关注 756

分析账号: QuandtStefan

发帖量: 1569

点赞量: 5563 (最新200条帖子)

转评量: 37 (最新200条帖子)

转评占比: 23.7% (最新200条帖子)

地址:

主页地址: <https://twitter.com/QuandtStefan>

注册时间: 2016-02-29 06:03:06

最近发帖时间: 2023-03-30 22:54:37

语种: 德语

简介: #StaatsbürgerInUniform de aus Überzeugung • Grundsatzreferent @TFK_Bundeswehr • Ex-@BAKS_Bund + PersRef @BAKS_President • Aus #SachsenInBerlin • Hier privat

系统应用：社交网络水军监测平台

≡ 社交机器人账号自动识别系统

Welcome,admin



目标详情

好友列表

粉丝列表

发帖列表

返回上级

共116条数据



Steffen Seibert

GerAmbTLV

链接 <https://twitter.com/GerA...>

简介 Ambassador of Germany to Israel. 2010-21 Spokesman of Chancellor Merkel and the German Federal Government. A TV journalist before that.

+ 加入分析



Botschaft Israel

IsraelinGermany

链接 <https://twitter.com/Israeli...>

简介 Shalom und willkommen beim offiziellen Twitter der Botschaft des Staates Israel in Berlin! - Welcome to the official Twitter of the Embassy of Israel in Berlin!

+ 加入分析



Michael Stempfle

mijoste

链接 <https://twitter.com/mijoste>

简介 Sprecher Verteidigungsministerium, privater account/personal views only

+ 加入分析



Der Postillon

Der_Postillon

链接 https://twitter.com/Der_P...

简介 Ehrliche Nachrichten - unabhängig, schnell, seit 1845

+ 加入分析



Der Gepardkommandant

gepardtatze

链接 <https://twitter.com/gepar...>

简介 Air Defence/Flugabwehr/Drohnen/Privater Account und meine persönliche Meinung.

+ 加入分析



Helge Fuhst

HelgeFuhst

链接 <https://twitter.com/Helge...>

简介 Journalist, Zweiter Chefredakteur ARD-aktuell, tagesschau & tagesthemen

+ 加入分析



Team Luftwaffe

Team_Luftwaffe

链接 https://twitter.com/Team_...

简介 Offizieller Auftritt der Luftwaffe

+ 加入分析



Heinrich Brauss

BraussHeinrich

链接 <https://twitter.com/Braus...>

简介 German Council on Foreign Relations; German Atlantic Association; former NATO Assistant Secretary General for Defence Policy and Planning (2013-2018)

+ 加入分析



Grünfink

Grnfink2

链接 <https://twitter.com/Grnfink...>

简介 Globale Konflikte, Sicherheit, Politik, Marine, Amphibik, Marineinfanterie und Ostsee #wirsindmarines #MOV (private Meinung), him/his

+ 加入分析



Marine im DBwV

DBwV_Marine

链接 <https://twitter.com/DBwV...>

简介 Moin, an dieser Stelle berichtet der Vorstand Marine im DBwV - und stellt Fragen und kommentiert!

+ 加入分析



Verteidigungsminis



Juliane Seifert



Thorsten Kodalle



Kersten Lahl



Heer

系统应用：社交网络水军监测平台

≡ 社交机器人账号自动识别系统

Welcome,admin ▾



目标详情 好友列表 粉丝列表 发帖列表

返回上级



QuandtStefan

Schon ein äußerst fragwürdiges #Demokratie-Verständnis: #Wagenknecht und andere @dieLinke fordern mehr Verständnis für + Verhandlungen mit #Putin und #Russland - gleichzeitig

原帖 2023-03-30 22:54:37 | 160 | 23

全文



QuandtStefan

□ Mal wieder 🇩🇪 @Team_Luftwaffe. #CharlesIII

原帖 2023-03-30 03:05:22 | 75

全文



QuandtStefan

📖 „Eine Deutsche Partei “ 🗨 Wie funktioniert die #NoAfD? Anschauen + verstehen. #Demokratie #Freiheit <https://t.co/FpsAXKB5mm>

原帖 2023-03-30 03:01:27 | 16 | 1

全文



QuandtStefan

RT @Panzermann7: Führungskreis des Territorialen Führungskommandos der Bundeswehr berät sich in Potsdam zur Landes- und Bündnisverteidigung...

转发 2023-03-27 21:12:08

全文



QuandtStefan

#Ukraine: Gert Gawellek erklärt im Interview mit der @NZZ den Deutschen, wie Russen kämpfen: „Wir müssen aufhören mit unseren Schönwetterübungen, Soldaten müssen wissen,

原帖 2023-03-25 17:45:30 | 181 | 18

全文



未来工作

- 下一步希望能够完成的工作包括但不限于：

1

多模态语义融合的舆情事件早期检测

2

人机交互的互联网水军对抗平台研制

3

社会网络突发舆情事件传播预警预测



谢谢各位老师和同学，敬请批评指正！

THANKS