CCF走进高校

中国计算机学会 CCF

中國計算機學會

CHINA COMPUTER FEDERATION

融合图表示学习与最优传输的图对齐研究进展

刘钰

北京交通大学计算机科学与技术学院

北京交通大学

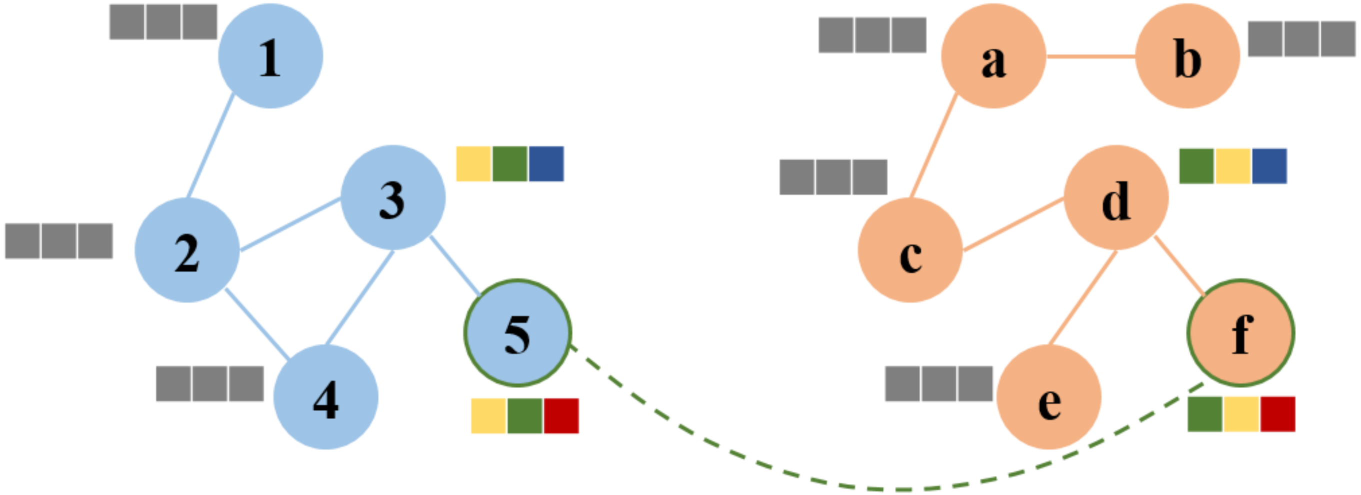
BEIJING JIAOTONG UNIVERSITY

汇报提纲

- 图对齐(Graph Alignment)问题与背景知识
- 现有代表性图对齐算法
- 我们的近期研究进展
  - 高鲁棒性的图对齐算法
  - 图对齐的表达能力研究
  - 融合大模型与最优传输的图对齐算法
- 图对齐的挑战和开放性问题

图对齐(Graph Alignment)问题

- 问题输入：
  - 两个图$G_s(V_s, E_s, X_s)$和$G_t(V_t, E_t, X_t)$，其中$X \in \mathbf{R}^{n\times d}$表示节点特征
  - 可能存在的一个锚链集(anchor links) S
- 问题输出：
  - $\{(u, v)|u \in G_s, v \in G_t, u和v匹配\}$
  - 一个对齐概率(alignment probability)矩阵$T \in \mathbf{R}^{|V_s|\times|V_t|}$，$T(u, v)$表示对齐概率



图对齐(Graph Alignment)问题

- 学习范式(Learning Paradigm)
  - 有监督(supervised)/半监督(semi-supervised)：有锚链，取决于有多少
  - 无监督(unsupervised)：无锚链
- 知识图谱(Knowledge Graph)上的实体对齐(Entity Alignment)
- 大图对齐
- 高异质图对齐(Highly Heterogeneous Entity Alignment)
- …

图对齐的应用

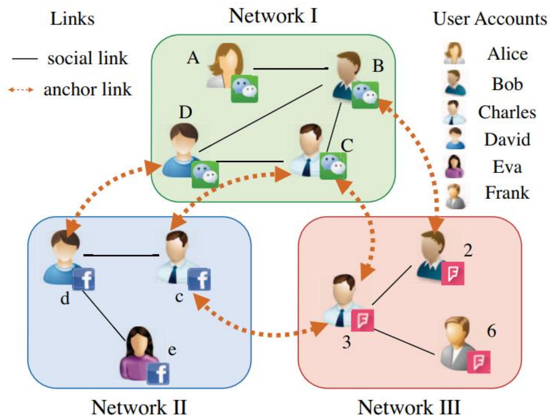
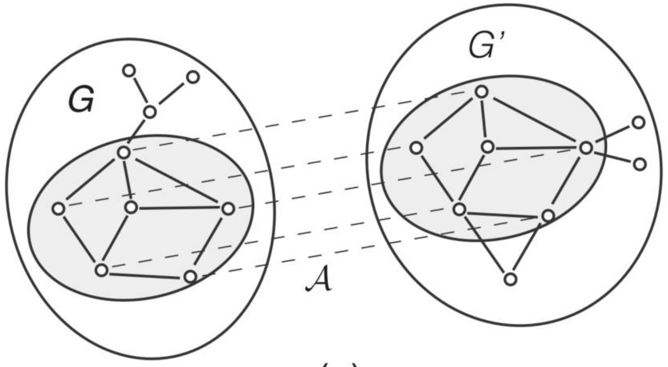


Fig. 1. An example of multiple anonymized partially aligned social networks.

社交网络账号对齐[ICDM’15]

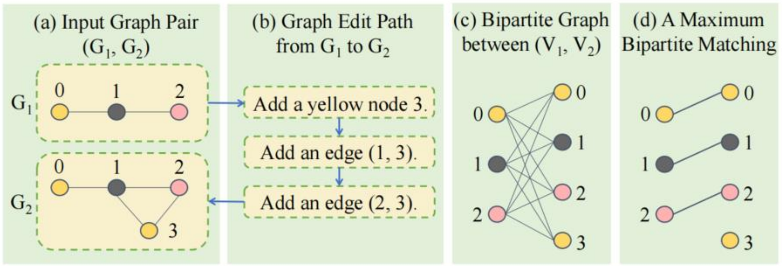


(a)

	gene co-expression networks	protein-protein interaction networks	metabolic networks
vertex	gene	protein	enzyme
edge	expression correlation of adjacent genes	adjacent proteins physically interact	adjacent enzymes share a metabolite
vertex similarity	sequence alignment score (*e.g.*, BLAST)	sequence alignment score (*e.g.*, BLAST)	functional similarity of adjacent enzymes
edge similarity	conservation of expression correlations	conservation of physical interactions	metabolite production by both species

(b)

生物网络中的应用[ICDM’15]



图编辑距离计算[VLDB’23]

- Jiawei Zhang, Philip S. Yu: Multiple Anonymized Social Networks Alignment. ICDM 2015: 599-608
- Michal Kolář, Jörn Meier, Ville Mustonen, Michael Lässig, Johannes Berg: GraphAlignment: Bayesian pairwise alignment of biological networks. BMC Syst. Biol. 6: 144 (2012)
- Chengzhi Piao, Tingyang Xu, Xiangguo Sun, Yu Rong, Kangfei Zhao, Hong Cheng: Computing Graph Edit Distance via Neural Graph Matching. Proc. VLDB Endow. 16(8): 1817-1829 (2023)

现有代表性图对齐算法（无监督）

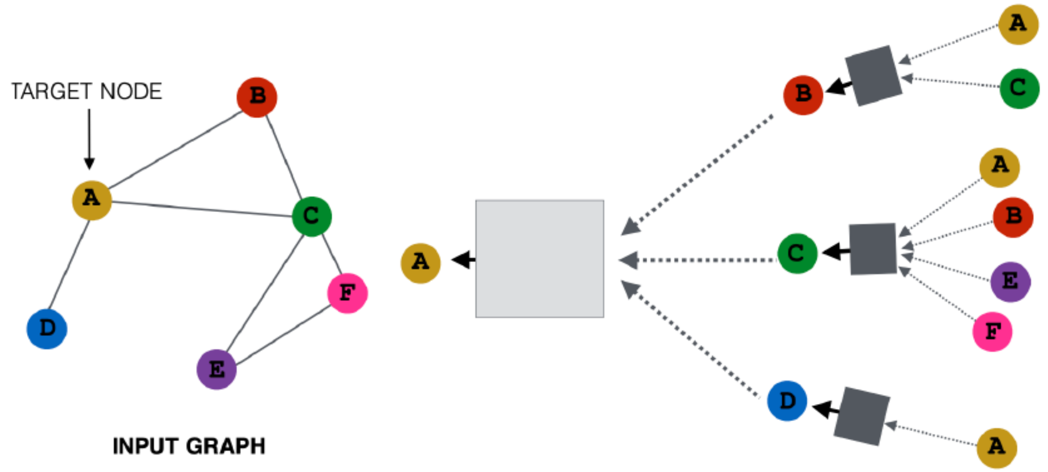
- 基于节点表示的(Embedding-based)方法
  - GTCAlign [TKDE’24]
- 基于最优传输(Optimal Transport-based)的方法
  - GWL [ICML’19]
  - SLOTAlign [ICDE’23]
- 其他代表性方法简介

基于节点表示的方法：GTCAlign [TKDE’24]

- 代表性的基于图表示学习(Graph Representation Learning)的方法：

节点表示学习+节点对齐(node alignment)

- 图神经网络(Graph Neural Network, GNN)回顾



两层图神经网络

非线性变换

节点v在第l层的向量表示

参数矩阵

$$\boldsymbol{h}_v^{(l+1)} = \sigma\left(\boxed{\sum_{u\in N(v)} \frac{\boldsymbol{h}_u^{(l)}}{|N(v)|} \boldsymbol{W}_{neigh}^{(l)}} + \boxed{\boldsymbol{h}_v^{(l)} \boldsymbol{W}_{self}^{(l)}}\right), \forall l \in \{0, \cdots, L-1\}$$

所有邻居节点u上一层表示的均值

参数矩阵

- Stanford CS224W: Machine Learning with Graphs. https://web.stanford.edu/class/cs224w/
- 《大规模图数据管理与分析》，高等教育出版社，邹磊，李友焕，刘钰，彭鹏著

基于节点表示的方法：GTCAlign [TKDE'24]

北京交通大学 BEIJING JIAOTONG UNIVERSITY

- 研究动机
  - "These methods focus on exploiting the local structures and node attributes for graph alignment and have yet to address cases where the local topology consistency principle does not apply."
- 解决办法：算一些全局拓扑指标(Global Topology Consistency)

Degree Centrality: $D_x = \frac{\deg(x)}{|E|}$

Closeness Centrality: $C_x = \frac{n-1}{\sum_{y \neq x} d(y,x)}$

Betweenness Centrality: $B_x = \sum_{z \neq y \neq x} \frac{\phi_{yz}(x)}{\phi_{yz}}$

Eigenvector Centrality: $\mathbf{A}\mathbf{c} = \lambda_{\max}\mathbf{c},\ E_x = \mathbf{c}[x]$

$$\Gamma_x = \frac{D_x + C_x + B_x + E_x}{4}$$

$$M_{uv} = \begin{cases} 1, & \text{if } \frac{\Gamma_u}{\Gamma_v} \leq L_1 \text{ or } \frac{\Gamma_u}{\Gamma_v} \geq 1/L_1; \\ 0, & \text{otherwise.} \end{cases}$$

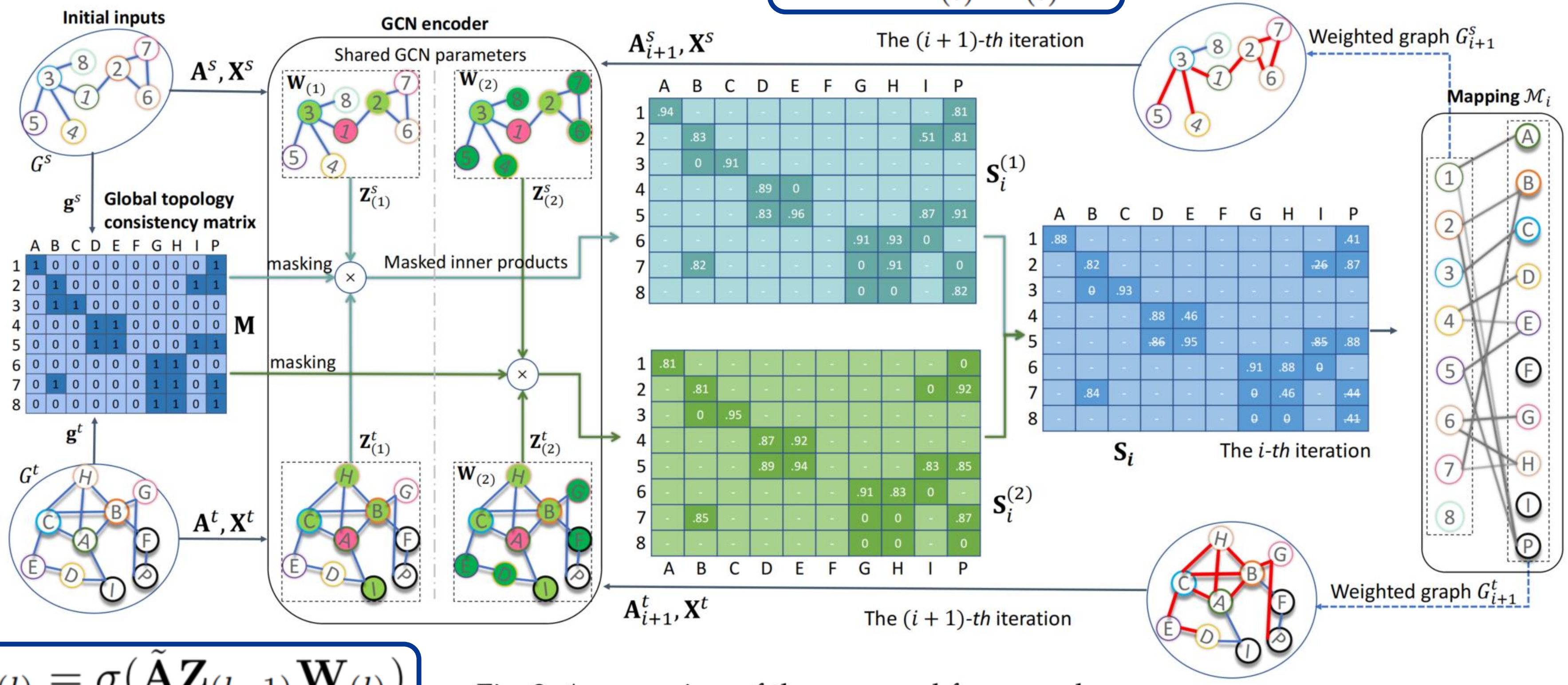
算出一个掩码矩阵

- Chenxu Wang, Peijing Jiang, Xiangliang Zhang, Pinghui Wang, Tao Qin, Xiaohong Guan: GTCAlign: Global Topology Consistency-Based Graph Alignment. IEEE Trans. Knowl. Data Eng. 36(5): 2009-2025 (2024)

基于节点表示的方法：GTCAlign [TKDE'24]

- 掩码矩阵配合图表示学习(GNN)
- 数据增强(Graph Augmentation)：修改图中部分边的权重
- 迭代优化

$$\mathbf{S}^{(l)} = \mathbf{Z}^{s}_{(l)}(\mathbf{Z}^{t}_{(l)})^{\top}$$



$$\mathbf{Z}_{(l)} = \sigma(\tilde{\mathbf{A}}\mathbf{Z}_{(l-1)}\mathbf{W}_{(l)})$$

Fig. 2: An overview of the proposed framework.

基于最优传输的方法：GWL [ICML'19]

- 研究动机
  - 图对齐可以自然的建模成GW学习(Gromov-Wasserstein Learning)和最优传输(Optimal Transport, OT)问题
- 最优传输(Optimal Transport, OT)回顾

In the case where the margins μ and ν are discrete, let μ_x and ν_y be the probability masses respectively assigned to $x \in \mathbf{X}$ and $y \in \mathbf{Y}$, and let γ_{xy} be the probability of an xy assignment. The objective function in the primal Kantorovich problem is then

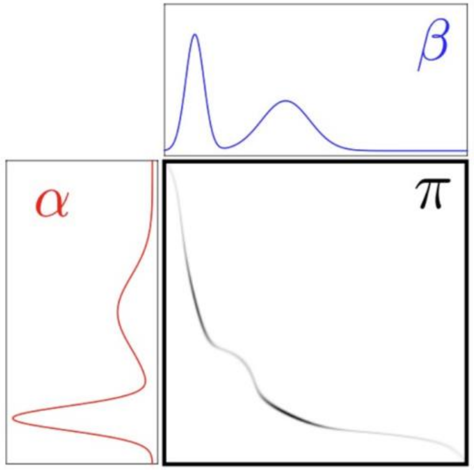
$$\sum_{x \in \mathbf{X}, y \in \mathbf{Y}} \gamma_{xy} c_{xy}$$

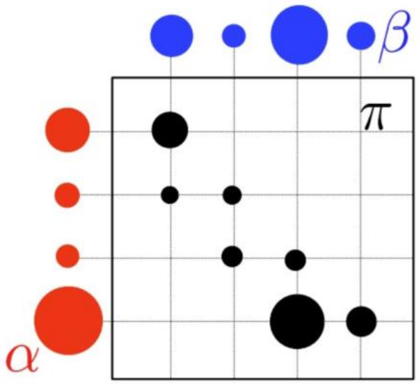
and the constraint $\gamma \in \Gamma(\mu, \nu)$ expresses as

$$\sum_{y \in \mathbf{Y}} \gamma_{xy} = \mu_x, \forall x \in \mathbf{X}$$

and

$$\sum_{x \in \mathbf{X}} \gamma_{xy} = \nu_y, \forall y \in \mathbf{Y}.$$





- Hongteng Xu, Dixin Luo, Hongyuan Zha, Lawrence Carin: Gromov-Wasserstein Learning for Graph Matching and Node Embedding. ICML 2019: 6932-6941

基于最优传输的方法：GWL [ICML’19]

- 研究动机
  - 图对齐可以自然的建模成GW学习(Gromov-Wasserstein Learning)和最优传输(Optimal Transport, OT)问题

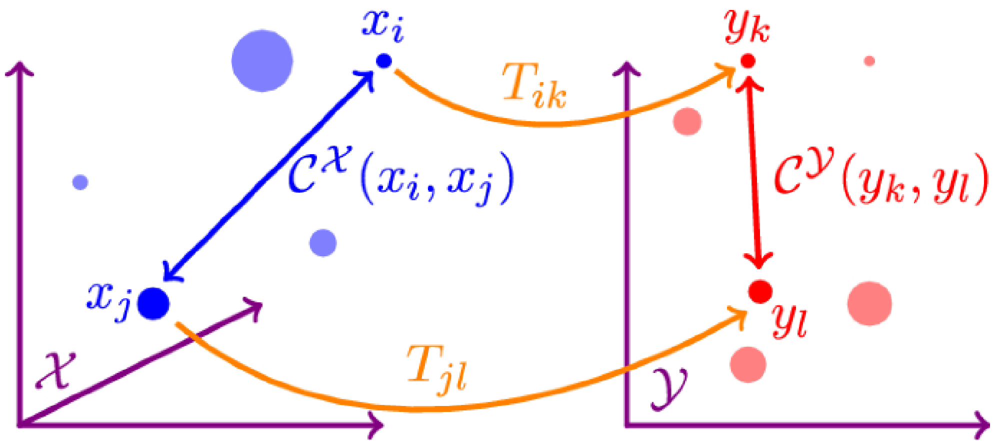
- Gromov-Wasserstein Learning回顾

Definition 2 (Gromov-Wasserstein Discrepancy (GWD) [17]). *Given the distribution $\boldsymbol{\mu}$ (resp. $\boldsymbol{\nu}$) over $\mathcal{V}_s$ (resp. $\mathcal{V}_t$), the GW discrepancy between $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$ is defined as*

$$\min_{\mathbf{T}\in\Pi(\boldsymbol{\mu},\boldsymbol{\nu})} \sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{l=1}^{n_2} |\mathbf{C}_s(i,j)-\mathbf{C}_t(k,l)|^2\mathbf{T}(i,k)\mathbf{T}(j,l), \quad (1)$$

$$s.t. \mathbf{T}\mathbf{1}_{n_2}=\boldsymbol{\mu}, \mathbf{T}^\top\mathbf{1}_{n_1}=\boldsymbol{\nu}.$$

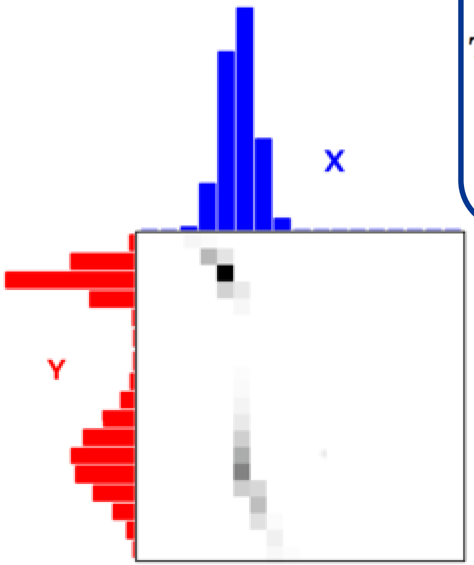
Here, $\mathbf{C}_s \in \mathbb{R}^{n_1\times n_1}$ and $\mathbf{C}_t \in \mathbb{R}^{n_2\times n_2}$ are the *intra-graph costs* for $\mathcal{G}_s$ and $\mathcal{G}_t$, respectively, which measure the similarity (or distance) of two nodes within each graph [21]. We have $\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\mathbf{T}(i,k)=1$ according to the constraints in Equation 1, i.e., $\mathbf{T}$ is the joint probability distribution over two node sets, and $\mathbf{1}_n$ denotes the all-ones vector in $\mathbb{R}^n$.



$$\mathcal{L}\left(C^{\mathcal{X}}(x_i,x_j), C^{\mathcal{Y}}(y_k,y_l)\right) T_{ik}T_{jl}$$

基于最优传输的方法：GWL [ICML’19]

- GWL算法
  - 设计cost的形式
  - 使用GW Learning和OT算法进行计算（Proximal Point Method, Sinkhorn Algorithm）
- 最优传输(OT)问题和Sinkhorn算法回顾

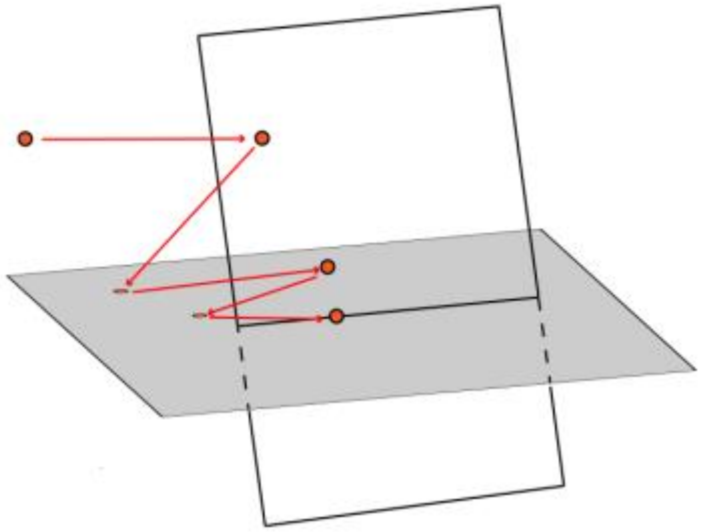


$$\min_{\mathbf{T}\in\Pi(\boldsymbol{\mu},\boldsymbol{\nu})} \sum_{i=1}^{n_1}\sum_{j=1}^{n_1}\sum_{k=1}^{n_2}\sum_{l=1}^{n_2} |\mathbf{C}_s(i,j)-\mathbf{C}_t(k,l)|^2\mathbf{T}(i,k)\mathbf{T}(j,l), \quad (1)$$

$$s.t.\mathbf{T}\mathbf{1}_{n_2}=\boldsymbol{\mu}, \mathbf{T}^\top\mathbf{1}_{n_1}=\boldsymbol{\nu}.$$

$$\mathbf{C}_{gwd}(i,k)=\sum_{j=1}^{n_1}\sum_{l=1}^{n_2}|\mathbf{C}_s(i,j)-\mathbf{C}_t(k,l)|^2\mathbf{T}(j,l)$$

$$\langle\mathbf{C}_{gwd},\mathbf{T}\rangle=\sum_{i=1}^{n_1}\sum_{k=1}^{n_2}\mathbf{C}_{gwd}(i,k)\mathbf{T}(i,k)$$



最优传输(OT)问题

Sinkhorn算法

- Hongteng Xu, Dixin Luo, Hongyuan Zha, Lawrence Carin: Gromov-Wasserstein Learning for Graph Matching and Node Embedding. ICML 2019: 6932-6941

基于最优传输的方法：SLOTAlign [ICDE'23]

- 研究动机
  - 只靠Embedding或OT方法，面对结构和特征不一致性时都不行

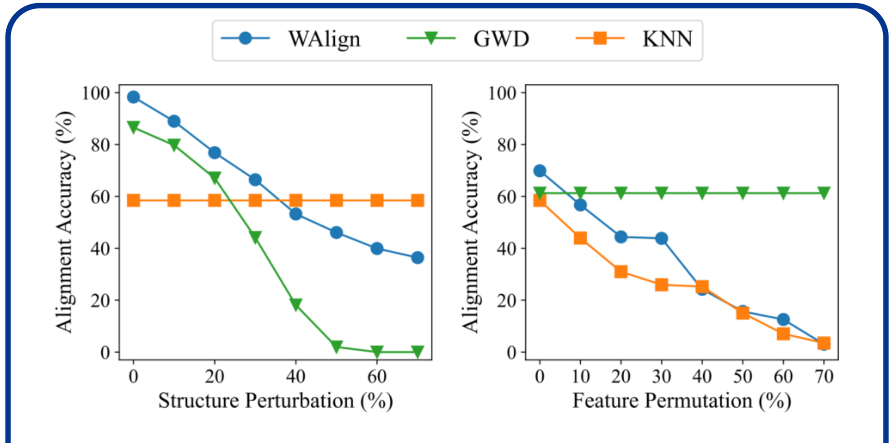


Fig. 3. Performance comparison of three graph alignment methods under different level of structure and feature inconsistency.

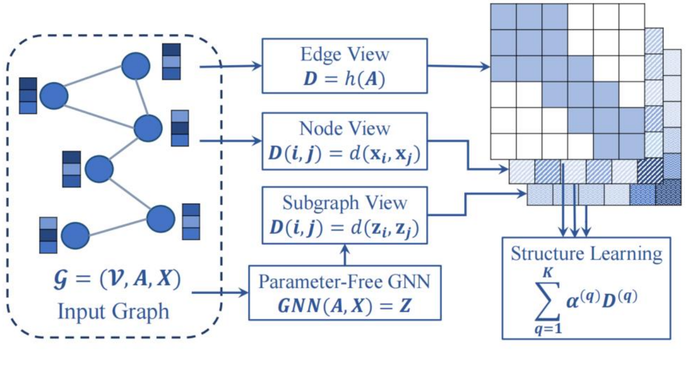


Fig. 5. The illustration of multi-view structure learning.

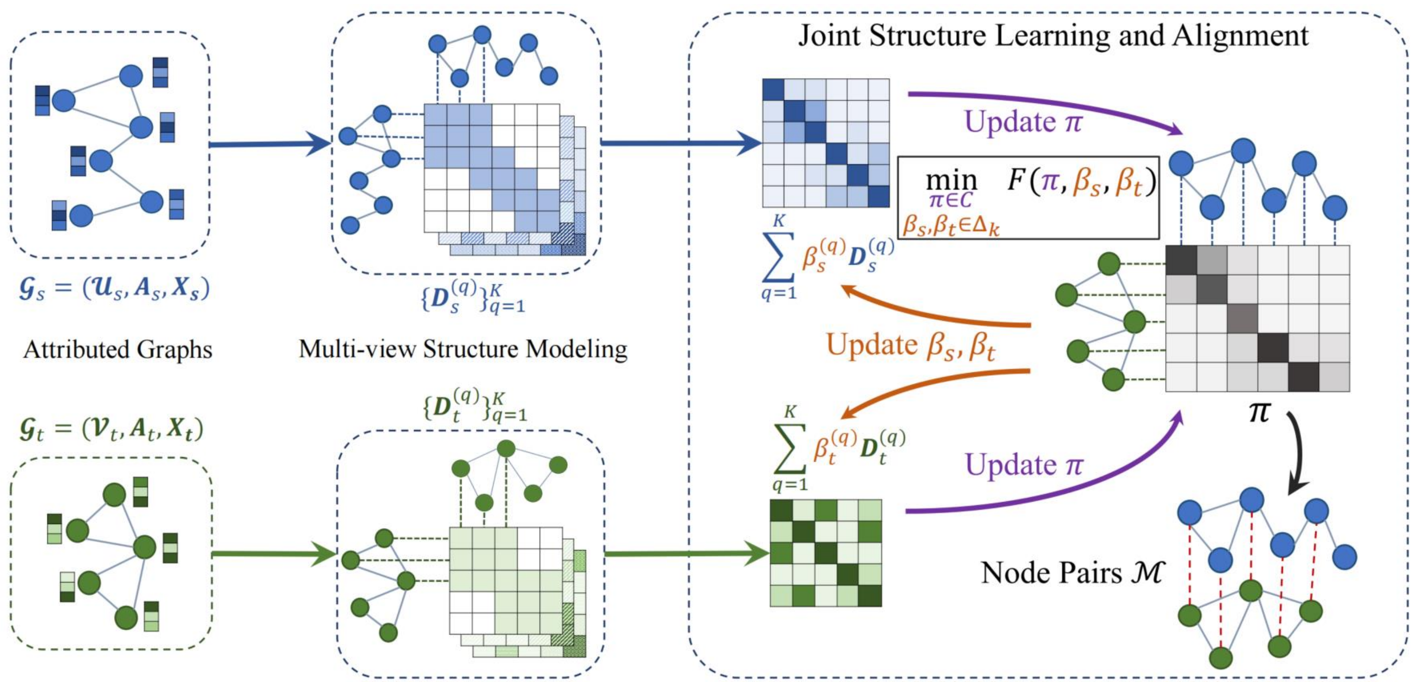


Fig. 4. The framework of SLOTAlign. Given two attribute graphs, it first constructs a set of candidate structure views for each graph, and then optimizes the integrated structure representations and node correspondences simultaneously in the joint structure learning and graph alignment framework.

- Jianheng Tang, Weiqi Zhang, Jiajin Li, Kangfei Zhao, Fugee Tsung, Jia Li: Robust Attributed Graph Alignment via Joint Structure Learning and Optimal Transport. ICDE 2023: 1638-1651

其他代表性方法简介

- 半监督(semi-supervised)图对齐
  - Zhichen Zeng, Si Zhang, Yinglong Xia, Hanghang Tong: PARROT: Position-Aware Regularized Optimal Transport for Network Alignment. WWW 2023: 372-382
- 知识图谱实体对齐(Entity Alignment)
  - Yunjun Gao, Xiaoze Liu, Junyang Wu, Tianyi Li, Pengfei Wang, Lu Chen: ClusterEA: Scalable Entity Alignment with Stochastic Training and Normalized Mini-batch Similarities. KDD 2022: 421-431
- 高异质图对齐(Highly Heterogeneous Entity Alignment)
  - Xuhui Jiang, Chengjin Xu, Yinghan Shen, Yuanzhuo Wang, Fenglong Su, Zhichao Shi, Fei Sun, Zixuan Li, Jian Guo, Huawei Shen: Toward Practical Entity Alignment Method Design: Insights from New Highly Heterogeneous Knowledge Graph Datasets. WWW 2024: 2325-2336

高鲁棒性的图对齐算法RESAlign

- 研究动机
  - 半监督(semi-supervised)图对齐对噪声的鲁棒性
- 解决办法
  - 可学习的传输代价设计(Learnable Transport Cost)
  - 多目标函数

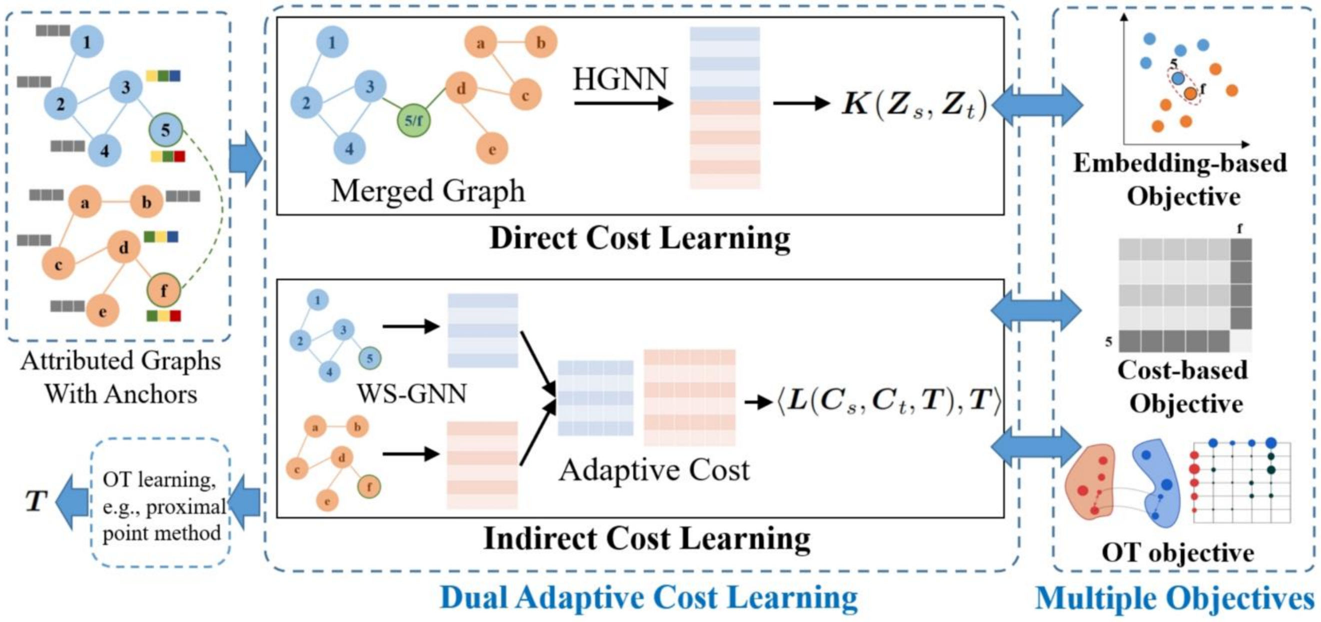


Fig. 1 Overview of the RESAlign framework

- Songyang Chen, Youfang Lin, Yu Liu*, Yuwei Ouyang, Zongshen Guo, Lei Zou: Enhancing robust semi-supervised graph alignment via adaptive optimal transport. World Wide Web (WWW) 28(2): 22 (2025)

高鲁棒性的图对齐算法RESAlign

BEIJING JIAOTONG UNIVERSITY

Table 3 Comparison of model performance in the semi-supervised setting.

Datasets	Metrics	FINAL	WAlign	NetTrans	BRIGHT	NeXtAlign	PARROT	RESAlign-WS	RESAlign
Douban Online-Offline	Hits@1	61.54	40.34	47.69	35.19	46.59	70.61	78.87	**80.67**
	Hits@5	91.68	62.68	61.32	52.29	77.54	93.63	94.32	**95.64**
	Hits@10	95.43	71.73	66.40	58.65	81.56	96.31	96.54	**97.21**
	MRR	74.33	51.05	58.91	43.51	58.50	80.59	87.63	**89.62**
	Time(s)	7.78	13.56	66.64	29.77	62.21	7.79	2.82	2.61
ACM-DBLP	Hits@1	39.72	62.69	70.36	47.88	51.68	72.11	74.38	**75.79**
	Hits@5	72.79	83.81	90.45	78.41	81.03	92.84	93.11	**94.07**
	Hits@10	83.34	89.74	94.28	85.63	87.87	96.01	96.33	**96.76**
	MRR	54.14	71.74	77.04	60.72	63.54	80.56	82.50	**83.18**
	Time(s)	21.75	238.27	586.54	751.29	1335.23	226.24	59.91	60.46
Allmv-Imdb	Hits@1	83.48	56.52	73.03	69.29	57.05	95.12	97.12	**97.25**
	Hits@5	95.87	72.58	82.46	80.72	67.56	96.57	97.44	**97.92**
	Hits@10	96.84	78.02	85.63	84.17	70.53	96.96	97.62	**98.24**
	MRR	89.06	64.22	77.00	74.52	61.44	95.79	97.42	**97.62**
	Time(s)	6.74	94.46	828.55	971.05	758.7	55.46	15.76	14.36
Cora	Hits@1	71.02	99.08	98.94	86.47	72.68	99.56	99.60	**99.63**
	Hits@5	86.25	99.95	99.86	98.84	84.45	**100**	**100**	**100**
	Hits@10	88.14	99.97	99.95	99.40	85.74	**100**	**100**	**100**
	MRR	77.34	99.47	99.35	92.23	77.68	99.71	99.75	**99.78**
	Time(s)	23.38	24.63	10.85	7.08	16.57	7.94	2.31	2.47
Citeseer	Hits@1	63.11	98.87	95.11	84.67	66.68	99.77	99.79	**99.86**
	Hits@5	69.08	**100**	97.10	99.84	80.58	**100**	**100**	**100**
	Hits@10	69.31	**100**	97.37	99.96	84.79	**100**	**100**	**100**
	MRR	65.76	99.42	96.09	91.92	72.61	99.86	99.89	**99.91**
	Time(s)	84.43	39.05	26.85	3.96	11.64	10.23	4.96	3.62
PPI	Hits@1	69.31	88.40	70.08	66.12	62.02	88.24	88.53	**88.68**
	Hits@5	75.18	91.64	77.93	77.08	78.43	91.32	**92.56**	92.43
	Hits@10	75.32	93.42	82.24	80.48	82.74	93.51	**93.71**	93.64
	MRR	71.89	90.16	76.64	71.23	69.02	90.40	90.47	**90.54**
	Time(s)	2.48	10.65	63.30	3.36	380.63	4.92	2.54	1.62

半监督场景与SOTA的对比

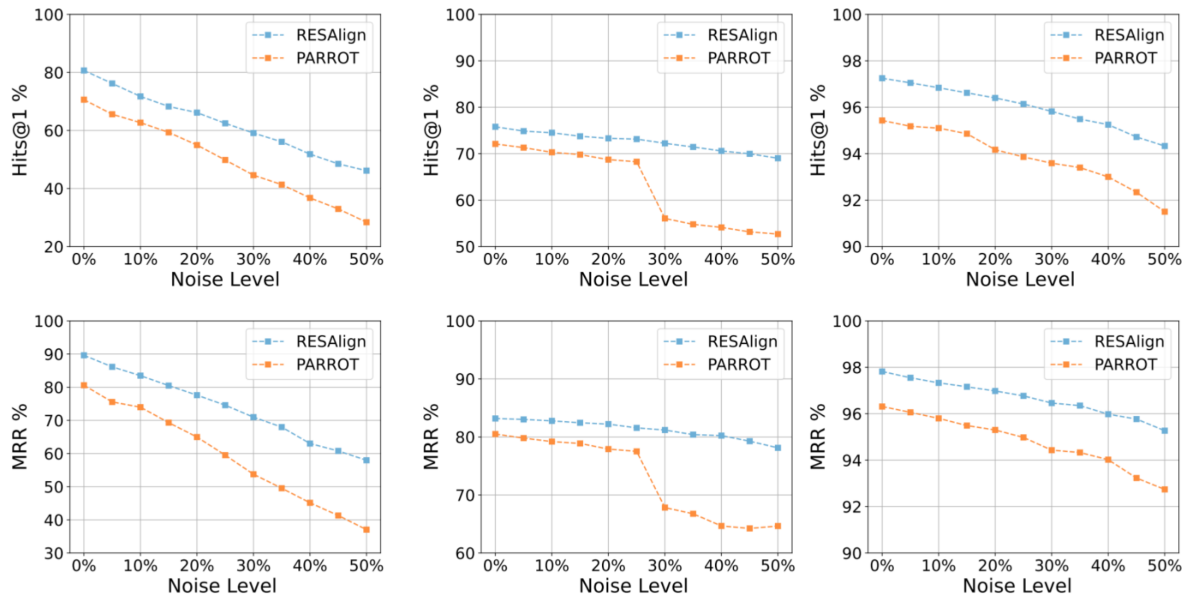
Table 4 Comparison of model performance in the unsupervised setting across five methods.

Datasets	Metrics	KNN	GAlign	GTCAlign	GWL	SLOTAlign	RESAlign-WS
Douban Online-Offline	Hits@1	27.55	45.26	61.79	3.29	51.43	**64.58**
	Hits@5	42.31	67.71	76.83	8.32	73.43	**82.56**
	Hits@10	49.28	78.00	82.29	9.93	77.73	**85.60**
	MRR	35.01	56.32	69.77	5.79	61.29	**76.70**
ACM-DBLP	Hits@1	36.25	70.20	60.92	56.36	66.04	**71.45**
	Hits@5	66.83	87.23	75.60	77.09	85.84	**89.91**
	Hits@10	76.22	91.36	79.97	82.18	87.76	**93.15**
	MRR	38.59	77.49	67.67	64.82	73.76	**79.53**
Allmv-Imdb	Hits@1	30.36	82.14	84.73	87.82	90.60	**95.47**
	Hits@5	47.14	86.35	89.89	92.31	92.75	**97.14**
	Hits@10	54.25	90.03	91.32	92.83	93.14	**97.41**
	MRR	38.59	84.96	87.12	89.64	91.61	**96.61**

无监督场景与SOTA的对比

- Songyang Chen, Youfang Lin, Yu Liu*, Yuwei Ouyang, Zongshen Guo, Lei Zou: Enhancing robust semi-supervised graph alignment via adaptive optimal transport. World Wide Web (WWW) 28(2): 22 (2025)

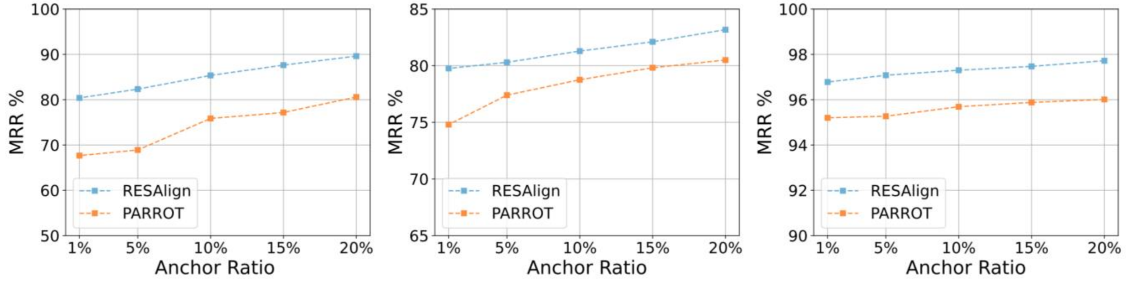
高鲁棒性的图对齐算法RESAlign



(a) Douban Online-Offline (b) ACM-DBLP (c) Allmv-Imdb

Fig. 2 Prediction accuracy vs. noise level

面对图结构扰动的鲁棒性



Douban Online-Offline ACM-DBLP Allmv-Imdb

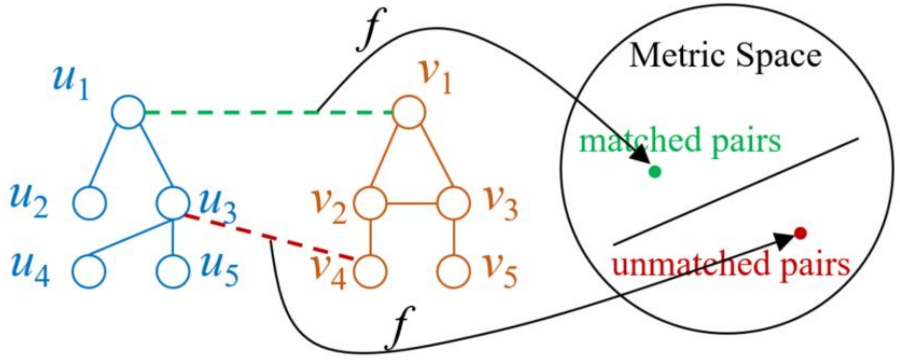
Fig. 4 MRR vs. anchor ratio

锚链比例对预测准确度的影响

- Songyang Chen, Youfang Lin, Yu Liu*, Yuwei Ouyang, Zongshen Guo, Lei Zou: Enhancing robust semi-supervised graph alignment via adaptive optimal transport. World Wide Web (WWW) 28(2): 22 (2025)

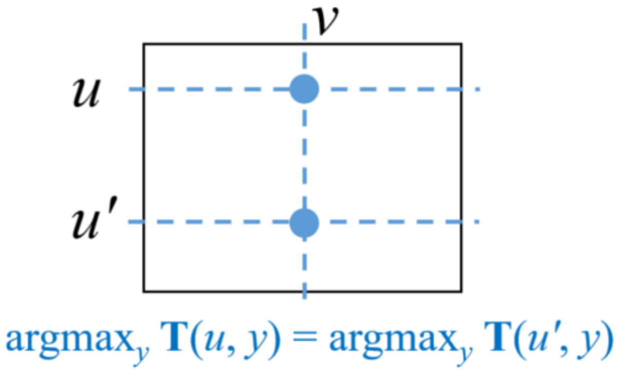
图对齐的表达能力研究（CombAlign算法）

- 现有（无监督）图对齐研究很少涉及模型的表达能力(expressive power)
- 区分匹配节点对和不匹配节点对的能力(discriminative power)

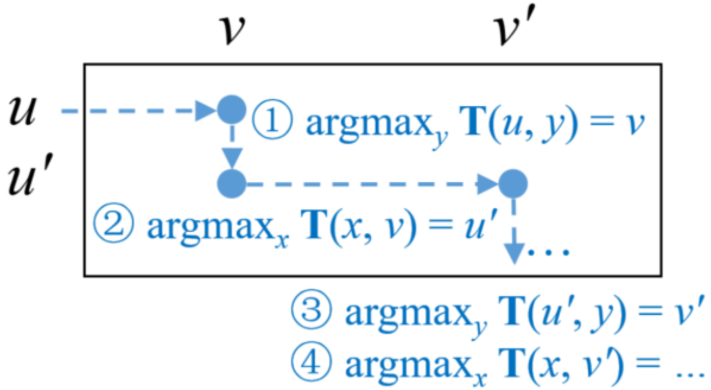


(a) Separating matched and unmatched node pairs

- 保证节点匹配性质(node matching property)的能力



(b) One-to-many prediction



(c) Mutual alignment

图对齐的表达能力研究（CombAlign算法）

- 解决办法：融合图表示学习、最优传输和传统图算法

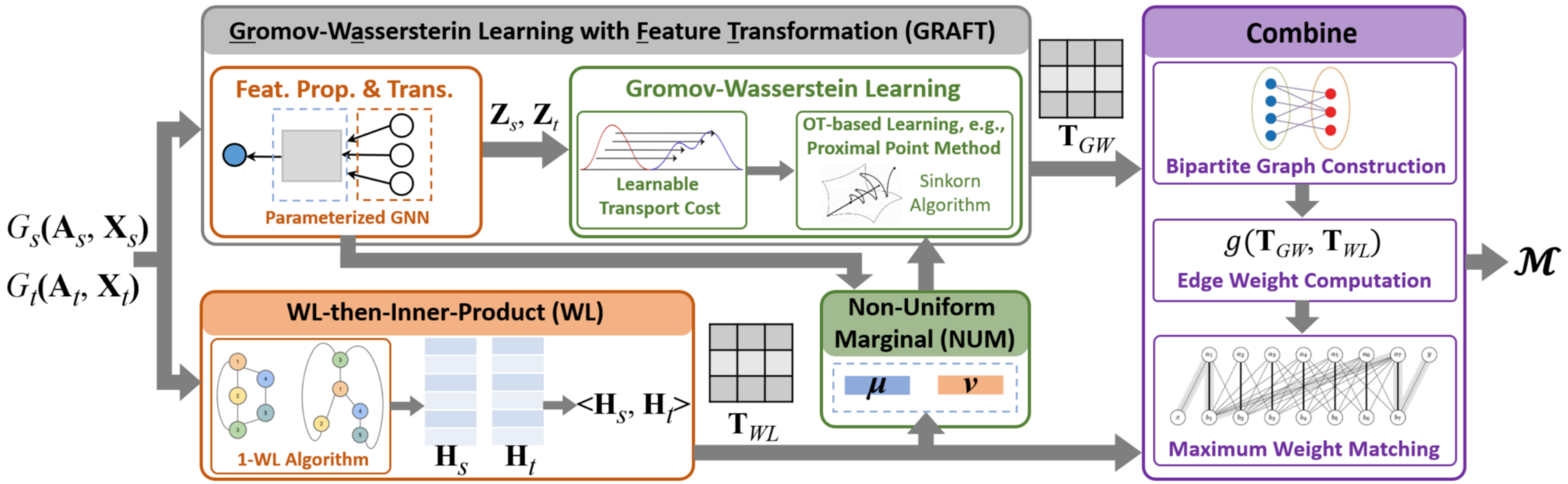


Fig. 2. The overall framework of CombAlign. Modules in orange, green, and purple belong to the embedding-based, OT-based, and traditional algorithm-based approaches, respectively.

Songyang Chen, Yu Liu*, Lei Zou, Zexuan Wang, Youfang Lin. CombAlign: Enhancing Model Expressiveness in Unsupervised Graph Alignment. Submitted to TKDE (Major Revision)

图对齐的表达能力研究（CombAlign算法）

- 解决办法：融合图表示学习、最优传输和传统图算法

只有特征传播
(feature propagation)
没有特征变换(GNN)
表达能力受限

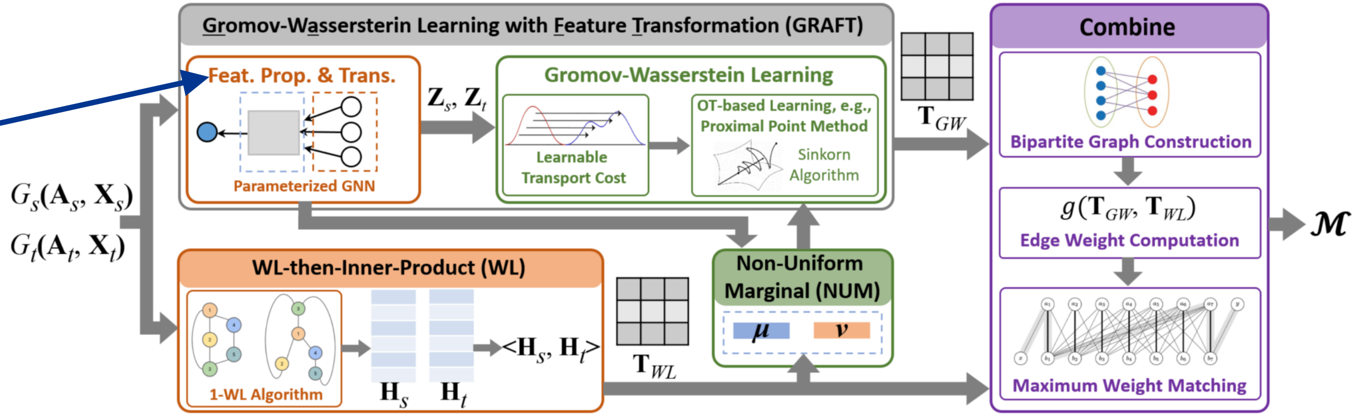


Fig. 2. The overall framework of CombAlign. Modules in orange, green, and purple belong to the embedding-based, OT-based, and traditional algorithm-based approaches, respectively.

Theorem III.1. *We are given the graph structures $\mathbf{A}_s, \mathbf{A}_t$ and node features $\mathbf{X}_s, \mathbf{X}_t$ as input. Denote feature propagation as $\mathbf{R}_p = g(\mathbf{A}_p)\mathbf{X}_p, p = s, t$, where $g(\cdot)$ is a function without learnable parameters. Denote the additional linear transformation as $\mathbf{Z}_p = \mathbf{R}_p\mathbf{W}$, where $\mathbf{W} \in \mathbb{R}^{d\times d}$ is the learnable matrix. Assume that we set the intra-graph cost matrices as $\mathbf{C}'_p = \mathbf{R}_p\mathbf{R}_p^\intercal$ and $\mathbf{C}_p = \mathbf{Z}_p\mathbf{Z}_p^\intercal$ for $p = s, t$, respectively, and let $(u_i, v_k), (u_j, v_l) \in \mathcal{M}^*$ and $(u_{j'}, v_l) \notin \mathcal{M}^*$ where $\mathcal{M}^*$ is the ground truth. Then, there exists a case that $|\mathbf{C}'_s(i,j) - \mathbf{C}'_t(k,l)| = |\mathbf{C}'_s(i,j') - \mathbf{C}'_t(k,l)|$ and $|\mathbf{C}_s(i,j) - \mathbf{C}_t(k,l)| \neq |\mathbf{C}_s(i,j') - \mathbf{C}_t(k,l)|$.*

Corollary III.2. *Assume that the intra-graph cost matrix $\mathbf{C}_p$ is the linear combination of graph structure $\mathbf{A}_p$, feature information $\mathbf{X}_p\mathbf{X}_p^\intercal$, and node embedding $\mathbf{Z}_p\mathbf{Z}_p^\intercal$, i.e.,*

$$\mathbf{C}_p = \beta_p^{(1)}\mathbf{A}_p + \beta_p^{(2)}\mathbf{X}_p\mathbf{X}_p^\intercal + \beta_p^{(3)}\mathbf{Z}_p\mathbf{Z}_p^\intercal, p = s, t, \quad (6)$$

where we use $\boldsymbol{\beta}_p$ to represent learnable coefficients. If feature transformation is applied to compute $\mathbf{Z}_p$, the costs have more discriminative power in separating the matched and unmatched node pairs.

Corollary III.3. *If the f_{GNN} function has more expressive power in distinguishing node embeddings, e.g. [42], the intra-graph cost has more discriminative power in separating matched and unmatched pairs.*

图对齐的表达能力研究（CombAlign算法）

• 解决办法：融合图表示学习、最优传输和传统图算法

在图上做最优传输，假设节点独立同分布导致表达能力受限

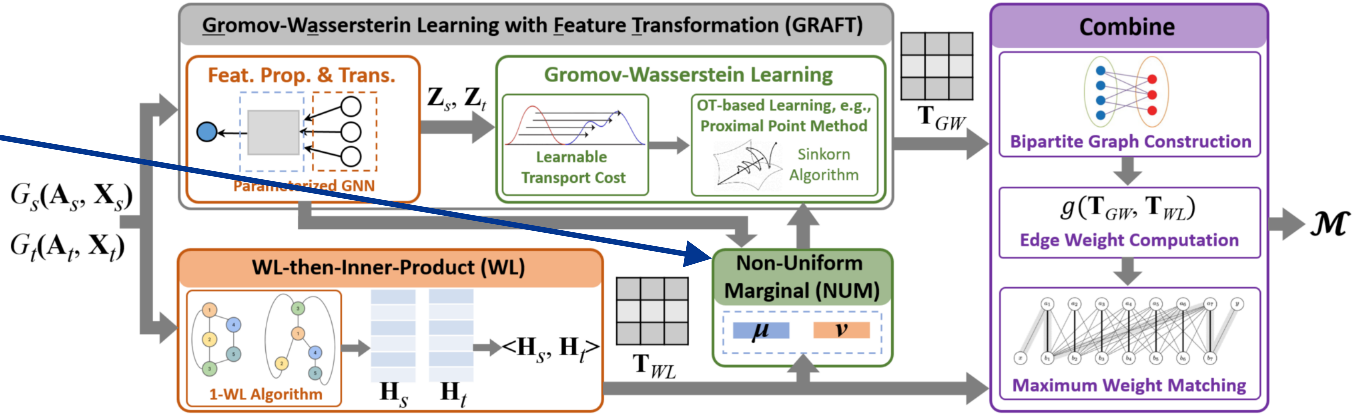
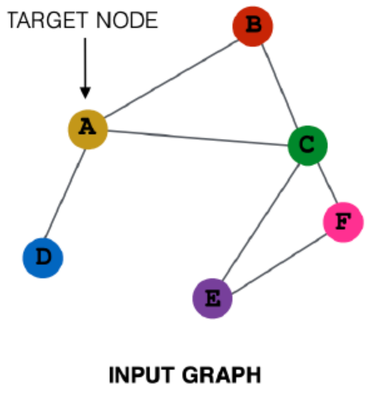


Fig. 2. The overall framework of CombAlign. Modules in orange, green, and purple belong to the embedding-based, OT-based, and traditional algorithm-based approaches, respectively.



图是非独立同分布的

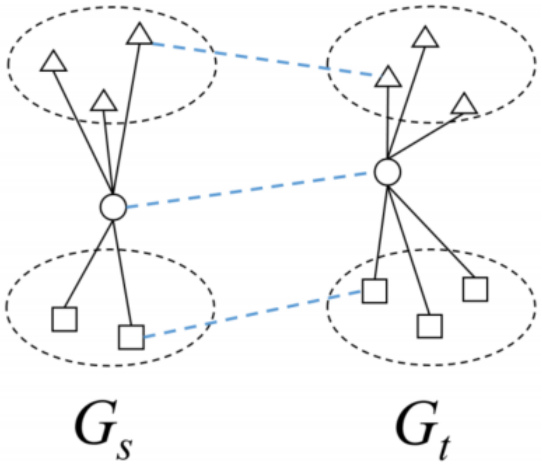
Theorem III.5. *Consider the case that* $(u_i, v_k) \in \mathcal{M}^*$ *and* $(u_i, v_{k'}) \notin \mathcal{M}^*$. *For GW learning (e.g., GRAFT), under the mild assumption of the intra-graph cost, i.e.,*

$$\mathbf{C}_s(i,i) = a, \forall u_i, \mathbf{C}_t(k,k) = b, \forall v_k, \tag{14}$$

$$\mathbf{C}_s(i,j) = \mathbf{C}_s(j,i), \mathbf{C}_t(k,l) = \mathbf{C}_t(l,k), \forall u_i, u_j, v_k, v_l, \tag{15}$$

$$\mathbf{C}_t(k,l) = \mathbf{C}_t(k',l), \forall v_l \in \mathcal{V}_t \backslash \{v_k, v_{k'}\}, \tag{16}$$

with uniform marginals $\boldsymbol{\mu} = (1/n_1, \ldots, 1/n_1)^\intercal$ *and* $\boldsymbol{\nu} = (1/n_2, \ldots, 1/n_2)^\intercal$, *the first iteration of the GW learning process with* $\mathbf{T}^{(0)} = \boldsymbol{\mu}\boldsymbol{\nu}^\intercal$ *cannot determine whether* u_i *is matched to* v_k *or* $v_{k'}$.



图对齐的表达能力研究（CombAlign算法）

- 解决办法：融合图表示学习、最优传输和传统图算法

如果要保证一些节点匹配性质，需要融入传统图算法

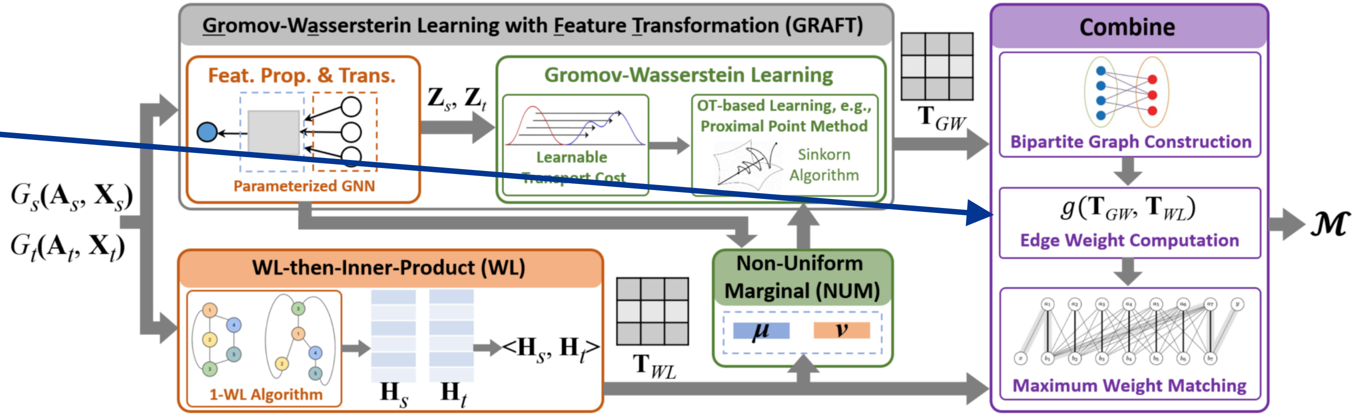


Fig. 2. The overall framework of CombAlign. Modules in orange, green, and purple belong to the embedding-based, OT-based, and traditional algorithm-based approaches, respectively.

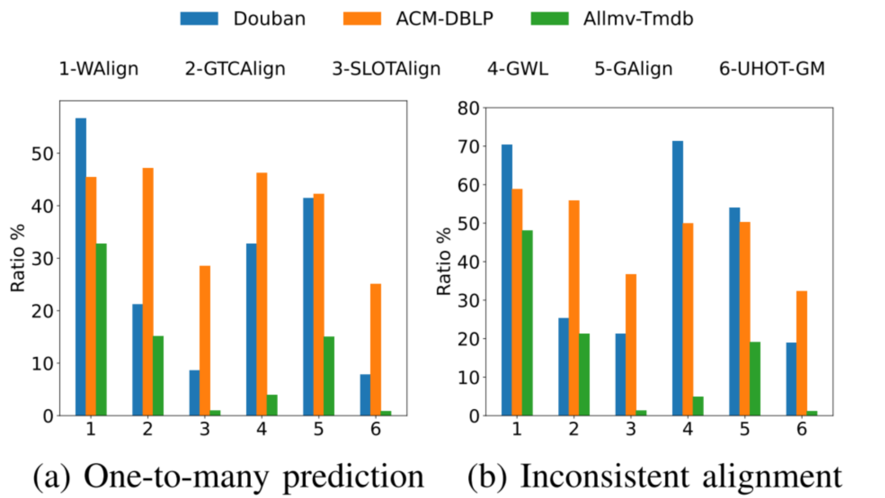


Fig. 4. Percentage of one-to-many predictions and inconsistencies in terms of mutual alignment for representative baselines.

图对齐的表达能力研究（CombAlign算法）

北京交通大学
BEIJING JIAOTONG UNIVERSITY

- 模型预测精度

TABLE IV
COMPARISON OF MODEL PERFORMANCE ON SIX DATASETS.

Datasets	Metrics	KNN	GAlign	WAlign	GTCAlign	GWL	SLOTAlign	UHOT-GM	CombAlign w/o C
Douban	Hits@1	27.82	45.26	39.45	61.79	3.29	51.43	60.23	**68.52**
	Hits@5	45.53	67.71	62.35	76.83	8.32	73.43	71.36	**87.84**
	Hits@10	52.68	78.00	71.47	82.29	9.93	77.73	76.91	**91.41**
	MAP	36.08	56.32	46.22	69.77	5.79	61.29	67.35	**77.08**
ACM-DBLP	Hits@1	36.35	70.20	63.43	60.92	56.36	66.04	69.89	**72.18**
	Hits@5	66.83	87.23	83.18	75.60	77.09	85.84	87.12	**88.98**
	Hits@10	76.22	91.36	86.58	79.97	82.18	87.76	90.65	**92.63**
	MAP	50.11	77.49	70.76	67.67	64.82	73.76	77.18	**79.55**
Allmovie-Imdb	Hits@1	32.39	82.14	52.61	84.73	87.82	90.60	91.73	**96.25**
	Hits@5	51.57	86.35	70.91	89.89	92.31	92.75	94.36	**97.66**
	Hits@10	58.79	90.03	76.52	91.32	92.83	93.14	94.96	**97.89**
	MAP	41.50	84.96	61.17	87.12	89.64	91.61	92.74	**97.31**
Cora	Hits@	95.01	99.45	98.45	99.35	86.19	99.48	99.48	**99.56**
	Hits@5	100	**100**	**100**	**100**	93.61	**100**	**100**	**100**
	Hits@10	100	**100**	**100**	**100**	94.57	**100**	**100**	**100**
	MAP	98.66	99.69	99.18	99.69	89.71	99.71	99.72	**99.75**
Citeseer	Hits@1	89.72	99.73	97.81	99.68	57.05	99.25	99.47	**99.82**
	Hits@5	100	**100**	**100**	**100**	65.04	**100**	**100**	**100**
	Hits@10	100	**100**	**100**	**100**	65.95	**100**	**100**	**100**
	MAP	94.91	99.84	98.88	99.89	61.31	99.62	99.69	**99.91**
PPI	Hits@1	84.96	89.20	88.51	89.25	86.76	89.30	89.33	**89.70**
	Hits@5	89.10	90.64	93.10	92.81	88.06	92.53	92.97	**93.15**
	Hits@10	92.17	94.16	94.17	94.07	88.62	93.49	93.52	**94.85**
	MAP	87.65	90.72	89.02	90.80	87.74	90.76	90.81	**91.12**

不保证节点匹配性质时的精度对比

TABLE V
COMPARISON OF HITS@1 WITH ONE-TO-ONE MATCHING CONSTRAINT.

Datasets	Douban	ACM-DBLP	Allmovie-Imdb
KNN	23.88	31.11	28.79
MJV	31.03	56.89	35.14
GAlign	20.84	66.18	80.28
WAlign	15.56	60.33	51.29
GTCAlign	57.15	56.96	83.91
GWL	3.22	52.09	87.22
SLOTAlign	49.19	64.32	90.31
UHOT-GM	58.31	67.78	91.26
CombAlign	**70.75**	**74.19**	**96.57**

保证节点匹配性质时的精度对比

Songyang Chen, Yu Liu*, Lei Zou, Zexuan Wang, Youfang Lin. CombAlign: Enhancing Model Expressiveness in Unsupervised Graph Alignment. Submitted to TKDE (Major Revision)

图对齐的表达能力研究（CombAlign算法）

北京交通大學
BEIJING JIAOTONG UNIVERSITY

- 模型各模块分析

TABLE VI
THE PERFORMANCE (HITS@1) OF GRAFT WITH DIFFERENT GNNS.

Hits@1	Douban	ACM-DBLP	Allmovie-Imdb
w/ LGCN	57.07	68.87	93.39
w/ GCN	61.45	69.68	94.57
w/ GIN	63.15	71.64	95.61
w/ SGFormer	65.38	72.73	96.02

图表示学习模块

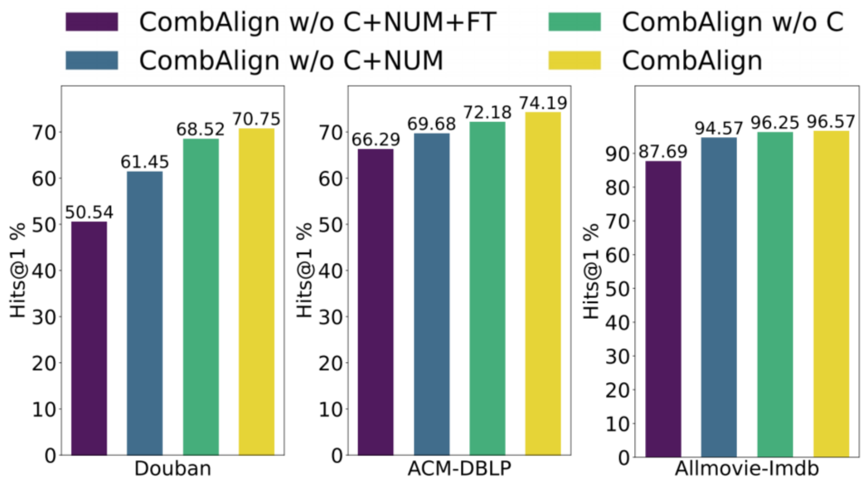


Fig. 5. Ablation study on three real-world datasets.

各模块消融实验

TABLE VII
IMPROVING OTHER OT-BASED MODELS W/ COMBINE (HITS@1).

Methods	Douban	ACM-DBLP	Allmovie-Imdb
GWL	3.29	56.36	87.82
GWL+C	5.81	63.46	90.07
SLOTAlign	51.43	66.04	90.60
SLOTAlign+C	60.01	69.20	91.03
UHOT-GM	60.23	69.89	91.73
UHOT-GM+C	63.17	71.32	92.45

把CombAlign框架应用到SOTA工作

TABLE VIII
THE RATIO OF THE REMAINING SET OVER GROUND TRUTH AFTER ELIMINATING ONE-TO-MANY PREDICTIONS.

CombAlign	Douban	ACM-DBLP	Allmovie-Imdb
- w/ C	**100%**	**100%**	**100%**
- w/o C	94.36%	88.36%	98.54%
- w/o C+NUM	94.18%	86.50%	97.60%
- w/o C+NUM+FT	91.32%	83.27%	96.25%

是否保证了节点匹配性质

Songyang Chen, Yu Liu*, Lei Zou, Zexuan Wang, Youfang Lin. CombAlign: Enhancing Model Expressiveness in Unsupervised Graph Alignment. Submitted to TKDE (Major Revision)

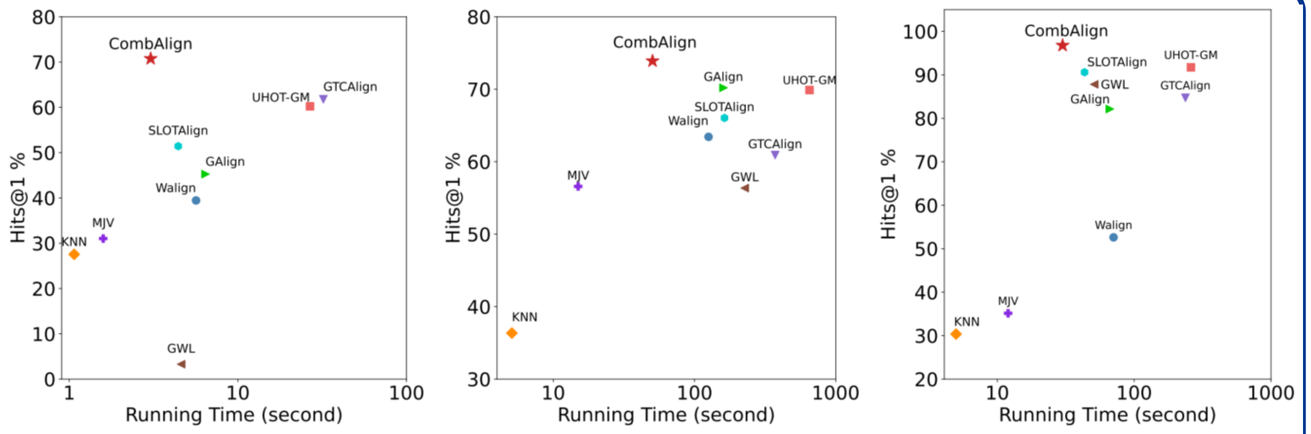
图对齐的表达能力研究（CombAlign算法）

- 模型预测精度

TABLE II
COMPARISON OF MODEL COMPLEXITY WITH STATE OF THE ART.

Method	Time Complexity
`GTCAlign` [16] (Emb.)	$O(IK(md+n^2d))$
`SLOTAlign` [3] (OT)	$O(I(K(md+nd^2)+n^2d+I_{ot}n^3))$
`CombAlign`	$O(I(K(md+nd^2)+n^2d+I_{ot}n^3))$
`CombAlign` (Optimized)	$O(I(K(md+nd^2)+n^2(d+I_{ot}\mathrm{log}n)))$

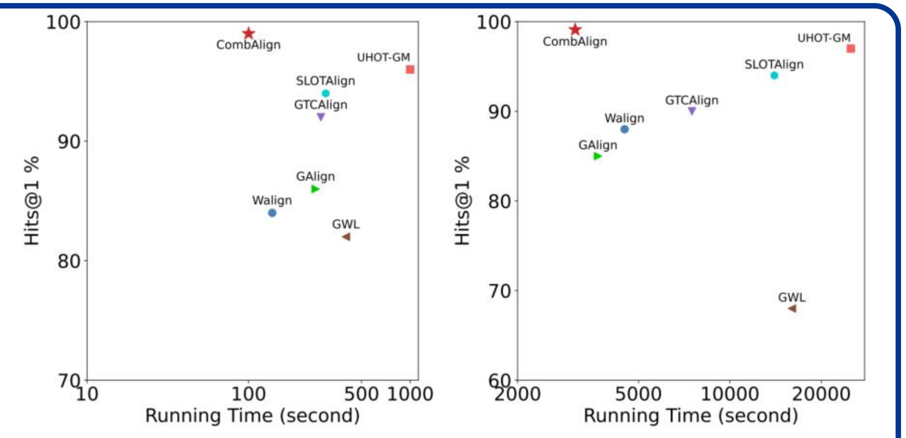
计算复杂度



(a) Douban (b) ACM-DBLP (c) Allmovie-Imdb

Fig. 7. Running time vs. Hits@1 of different methods.

运行时间vs.准确度



(a) CS (b) Physics

Fig. 8. Scalability analysis.

可扩展性对比

图对齐的表达能力研究（GlobAlign算法）

- 研究动机
  - 图对齐：全局比对
  - 基于图表示学习/最优传输的方法：局部信息计算表示或传输代价，不匹配

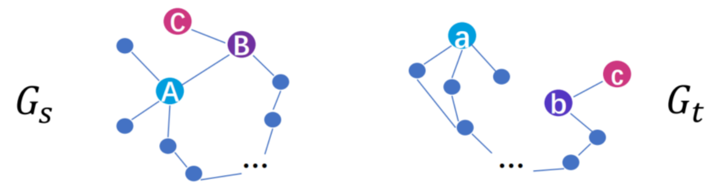
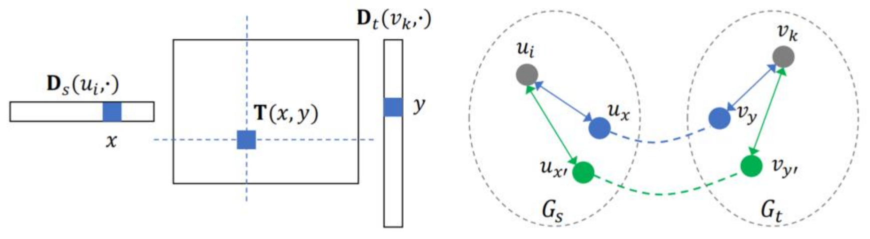


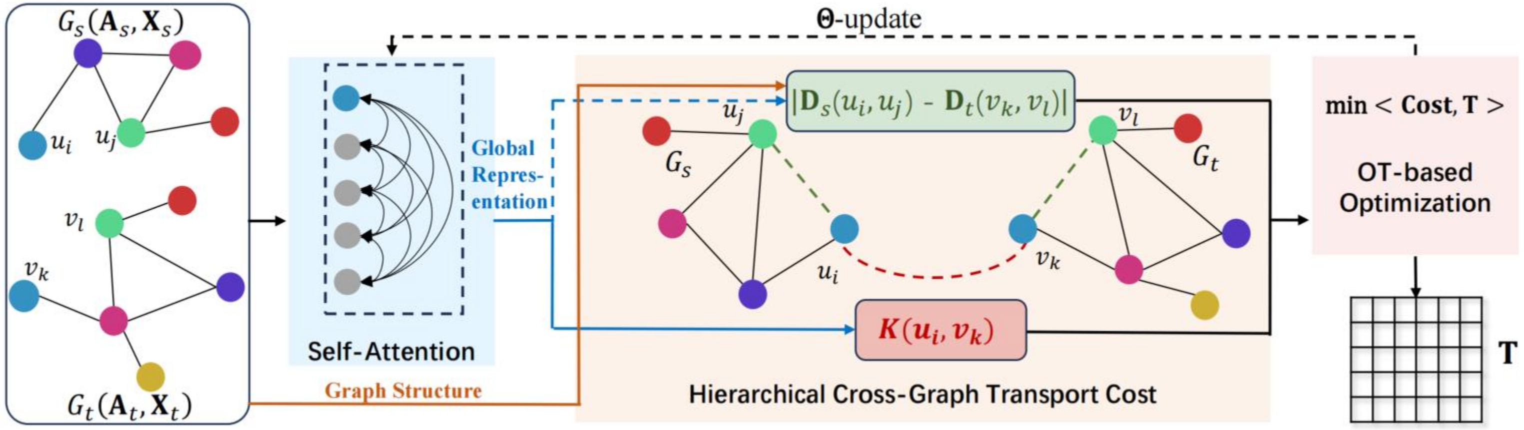
Figure 2: A toy example to show the limitation of local representation for graph alignment.



(a) Matrix formation (b) Graph formation

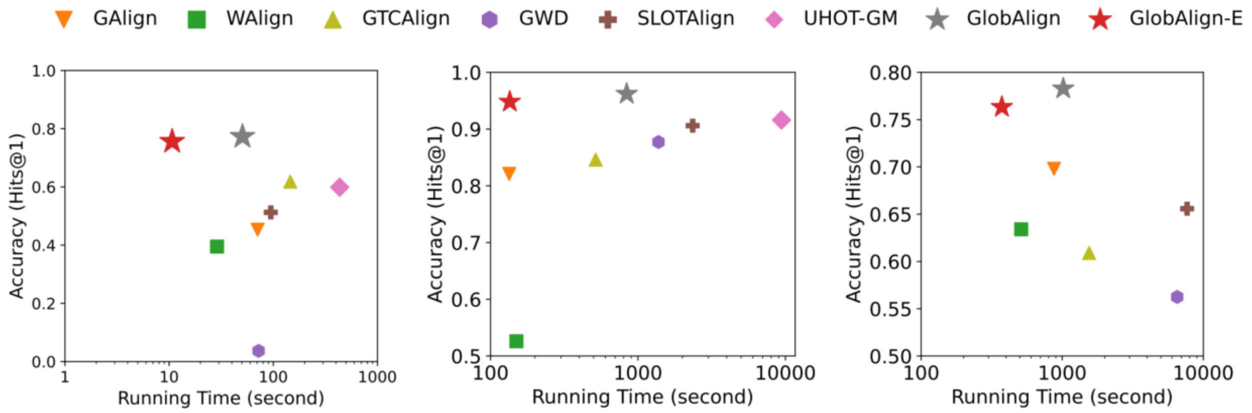
Figure 3: Illustration of node alignment from OT perspective.

- 解决办法



Songyang Chen, Youfang Lin, Yu Liu*, Shuai Zheng, Lei Zou. Towards Effective and Efficient Graph Alignment without Supervision. Submitted to WWWJ

图对齐的表达能力研究（GlobAlign算法）



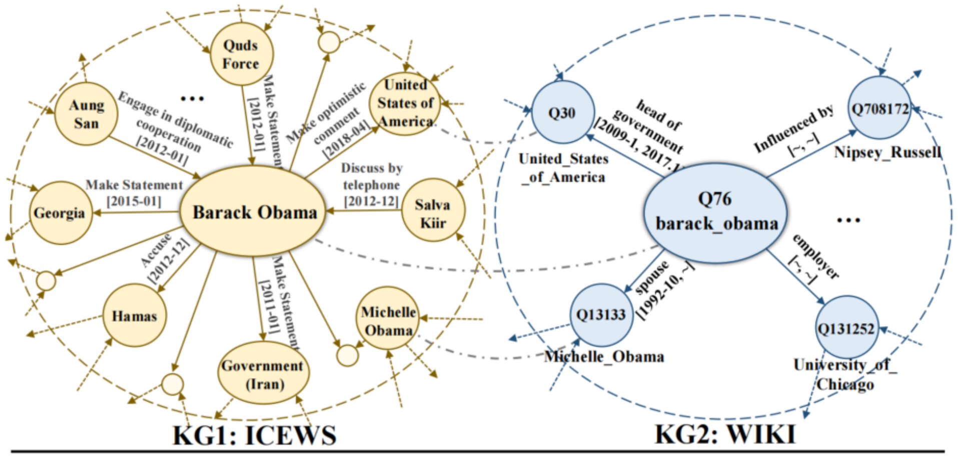
(a) Douban (b) Allmv-Imdb (c) ACM-DBLP

Figure 1: Running time (s) vs. accuracy (Hits@1) on three widely-adopted datasets [5, 11, 9, 10, 6]. Existing embedding and OT-based methods show similar performance in terms of efficiency-accuracy tradeoff. Our `GlobAlign` model significantly surpasses existing solutions in accuracy, while our `GlobAlign-E` model achieves up to one order of magnitude speedup with comparable performance.

运行时间vs.准确度

融合大模型与最优传输的ULOT-HHEA算法

- 高异质图对齐(Entity Alignment on Highly Heterogeneous Knowledge Graphs)
  - 直接使用GNN/Knowledge Graph Representation Learning的方法失效



Dataset		#Entities	#Relations	#Facts	Density	#Anchors	Overlapping
ICEWS-WIKI	ICEWS	11,047	272	3,527,881	319.352	5,058	45.79%
	WIKI	15,896	226	198,257	12.472		31.82%

Figure 1: An example of the HHKGs (ICEWS and WIKI) and their statistics. The scale and density of the KGs are very different, and the overlapping ratios (45.79% and 31.82%) indicate that the two KGs are far from the 1-to-1 assumption.

HHKG特性

Table 2: The settings of baselines, and the main experiment results on DBP15K(EN-FR), DBP-WIKI, ICEWS-WIKI, and ICEWS-YAGO. *Bold*: the best result; *Underline*: the runner-up result. "*Struc.*, *Name.*, *Temporal.*" denotes whether methods utilize structure, name, and temporal information, respectively; "*Semi.*" denotes whether methods adopt the semi-supervised strategy; Baselines are separated according to the groups described in Section A.5.2.

	Models	Settings				DBP15K(EN-FR)			DBP-WIKI			ICEWS-WIKI			ICEWS-YAGO		
		Struc.	*Name.*	*Temp.*	*Semi.*	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR
Trans.	MTransE					0.247	0.577	0.360	0.281	0.520	0.363	0.021	0.158	0.068	0.012	0.084	0.040
	AlignE	✓				0.481	0.824	0.599	0.566	0.827	0.655	0.057	0.261	0.122	0.019	0.118	0.055
	BootEA	✓			✓	0.653	0.874	0.731	0.748	0.898	0.801	0.072	0.275	0.139	0.020	0.120	0.056
GNN	GCN-Align	✓				0.411	0.772	0.530	0.494	0.756	0.590	0.046	0.184	0.093	0.017	0.085	0.038
	RDGCN	✓	✓			0.873	0.950	0.901	0.974	0.994	0.980	0.064	0.202	0.096	0.029	0.097	0.042
	Dual-AMN(*basic*)	✓				0.756	0.948	0.827	0.786	0.952	0.848	0.077	0.285	0.143	0.032	0.147	0.069
	Dual-AMN(*semi*)	✓			✓	0.840	0.965	0.888	0.869	0.969	0.908	0.037	0.188	0.087	0.020	0.093	0.045
	Dual-AMN(*name*)	✓	✓			0.954	0.994	0.970	0.983	0.996	0.991	0.083	0.281	0.145	0.031	0.144	0.068
	TEA-GNN	✓		✓		-	-	-	-	-	-	0.063	0.253	0.126	0.025	0.135	0.064
	TREA	✓		✓		-	-	-	-	-	-	0.081	0.302	0.155	0.033	0.150	0.072
	STEA	✓		✓	✓	-	-	-	-	-	-	0.079	0.292	0.152	0.033	0.147	0.073
Other	BERT		✓			0.937	0.985	0.956	0.941	0.980	0.963	0.546	0.687	0.596	0.749	0.845	0.784
	FuAlign	✓	✓			0.936	0.988	0.955	0.980	0.991	0.986	0.257	0.570	0.361	0.326	0.604	0.423
	BERT-INT	✓	✓			**0.990**	**0.997**	**0.993**	**0.996**	**0.997**	**0.996**	0.561	0.700	0.607	0.756	0.859	0.793
	PARIS	✓	✓			0.902	-	-	0.963	-	-	0.672	-	-	0.687	-	-
Ours	Simple-HHEA		✓	✓		0.948	0.991	0.960	0.967	0.988	0.979	**0.720**	**0.872**	**0.754**	**0.847**	**0.915**	**0.870**
	Simple-HHEA+	✓	✓	✓		0.959	0.995	0.972	0.975	0.991	0.988	0.639	0.812	0.697	0.749	0.864	0.775

Simple-HHEA实验效果对比

Xuhui Jiang, Chengjin Xu, Yinghan Shen, Yuanzhuo Wang, Fenglong Su, Zhichao Shi, Fei Sun, Zixuan Li, Jian Guo, Huawei Shen: Toward Practical Entity Alignment Method Design: Insights from New Highly Heterogeneous Knowledge Graph Datasets. WWW 2024: 2325-2336

融合大模型与最优传输的ULOT-HHEA算法

- 解决办法
  - 虽然embedding-based的方法失效，但是OT-based方法仍然有效

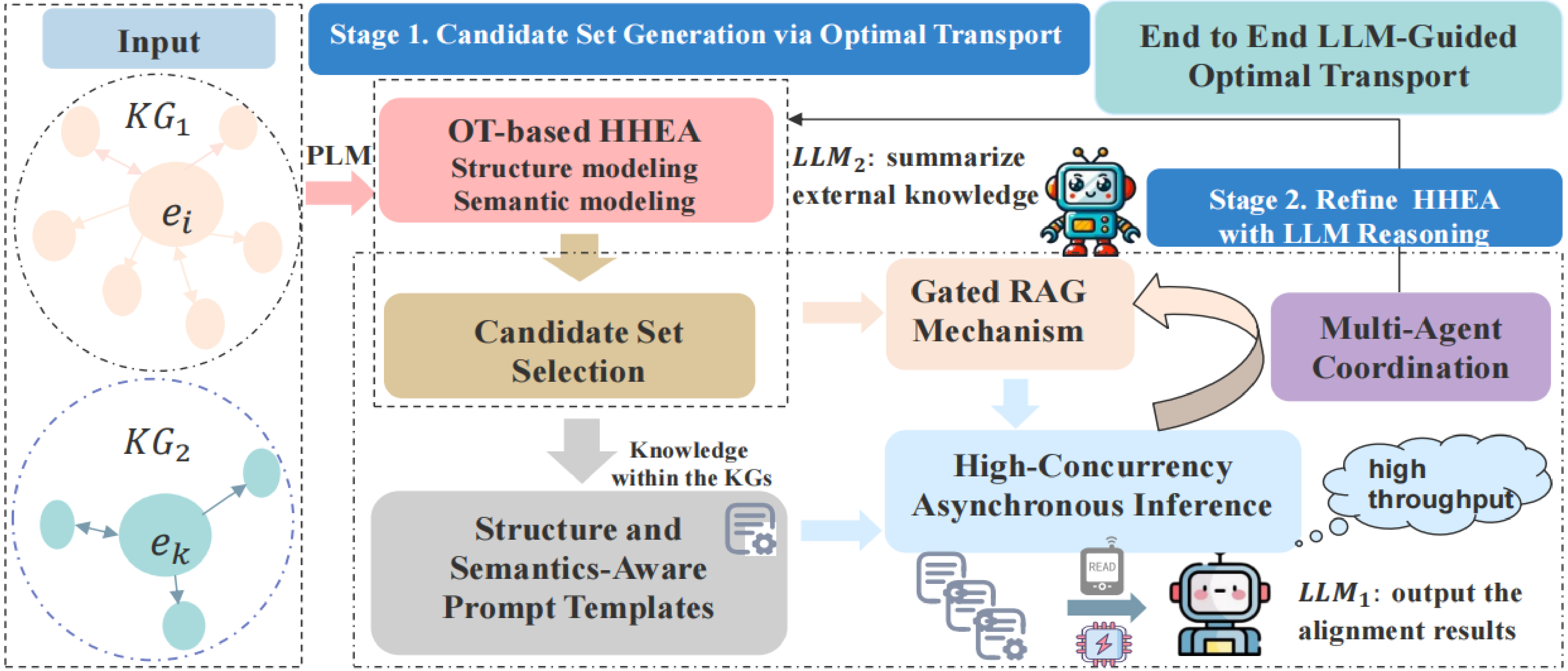


Figure 1: The architecture of ULOT-HHEA.

When Large Language Models Meet Optimal Transport: Efficient Alignment of Highly Heterogeneous Knowledge Graphs. Submitted to AAAI

融合大模型与最优传输的ULOT-HHEA算法

- 实验效果

Method		ICEWS-WIKI			ICEWS-YAGO			DBP15K(ZH-EN)			DBP15K(JA-EN)			DBP15K(FR-EN)		
Type	Model	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR
Semi-Supervised	MTransE	0.021	0.158	0.068	0.012	0.084	0.04	0.308	0.614	0.364	0.279	0.575	0.349	0.247	0.577	0.360
	AlignE	0.057	0.261	0.122	0.019	0.118	0.055	0.472	0.792	0.581	0.448	0.789	0.563	0.481	0.824	0.599
	GCN-Align	0.046	0.184	0.093	0.017	0.085	0.038	0.413	0.744	0.549	0.399	0.745	0.546	0.411	0.772	0.531
	RDGCN	0.064	0.202	0.096	0.029	0.097	0.042	0.708	0.846	0.746	0.767	0.895	0.812	0.873	0.952	0.901
	BootEA	0.072	0.275	0.139	0.022	0.123	0.056	0.629	0.848	0.703	0.622	0.854	0.701	0.653	0.874	0.731
	Dual-AMN	0.083	0.281	0.145	0.031	0.144	0.068	0.861	0.964	0.901	0.892	0.978	0.925	0.954	0.994	0.970
	BERT-INT	0.561	0.704	0.607	0.756	0.859	0.793	0.968	0.992	0.977	0.964	0.991	0.975	0.990	0.985	0.993
	Simple-HHEA	0.722	0.872	0.754	0.847	0.915	0.872	-	-	-	-	-	-	0.959	0.995	0.972
Unsupervised	SEU	0.002	0.003	0.002	0.001	0.007	0.007	0.902	0.965	0.924	0.956	0.991	0.969	0.988	0.999	0.992
	EASY	-	-	-	-	-	-	0.898	0.979	0.931	0.943	0.991	0.960	0.981	0.998	0.994
	CPL-OT	0.001	0.008	0.005	0.026	0.076	0.043	0.927	0.964	0.944	0.956	0.983	0.971	0.991	0.994	0.992
	LightEA	0.005	0.034	0.016	0.002	0.013	0.006	0.952	0.984	0.964	0.981	0.997	0.987	0.995	0.998	0.996
	ChatEA	0.883	0.945	0.912	0.935	0.955	0.944	-	-	-	-	-	-	0.991	**1.000**	0.995
	AdaCoAgentEA	0.954	0.973	0.961	0.937	0.958	0.944	-	-	-	-	-	-	0.972	0.993	0.986
	ULOT-HHEA	**0.986**	**0.995**	**0.989**	**0.980**	**0.988**	**0.983**	**0.984**	**0.996**	**0.989**	**0.991**	**0.999**	**0.994**	**0.998**	**1.000**	**0.999**
	Improve (absolute)	0.032	0.022	0.028	0.043	0.030	0.039	0.016	0.004	0.011	0.010	0.002	0.007	0.003	0	0.003

Table 2: Main experiment results on the five datasets. Bold: the best result; Underline: the runner-up result.

When Large Language Models Meet Optimal Transport: Efficient Alignment of Highly Heterogeneous Knowledge Graphs. Submitted to AAAI

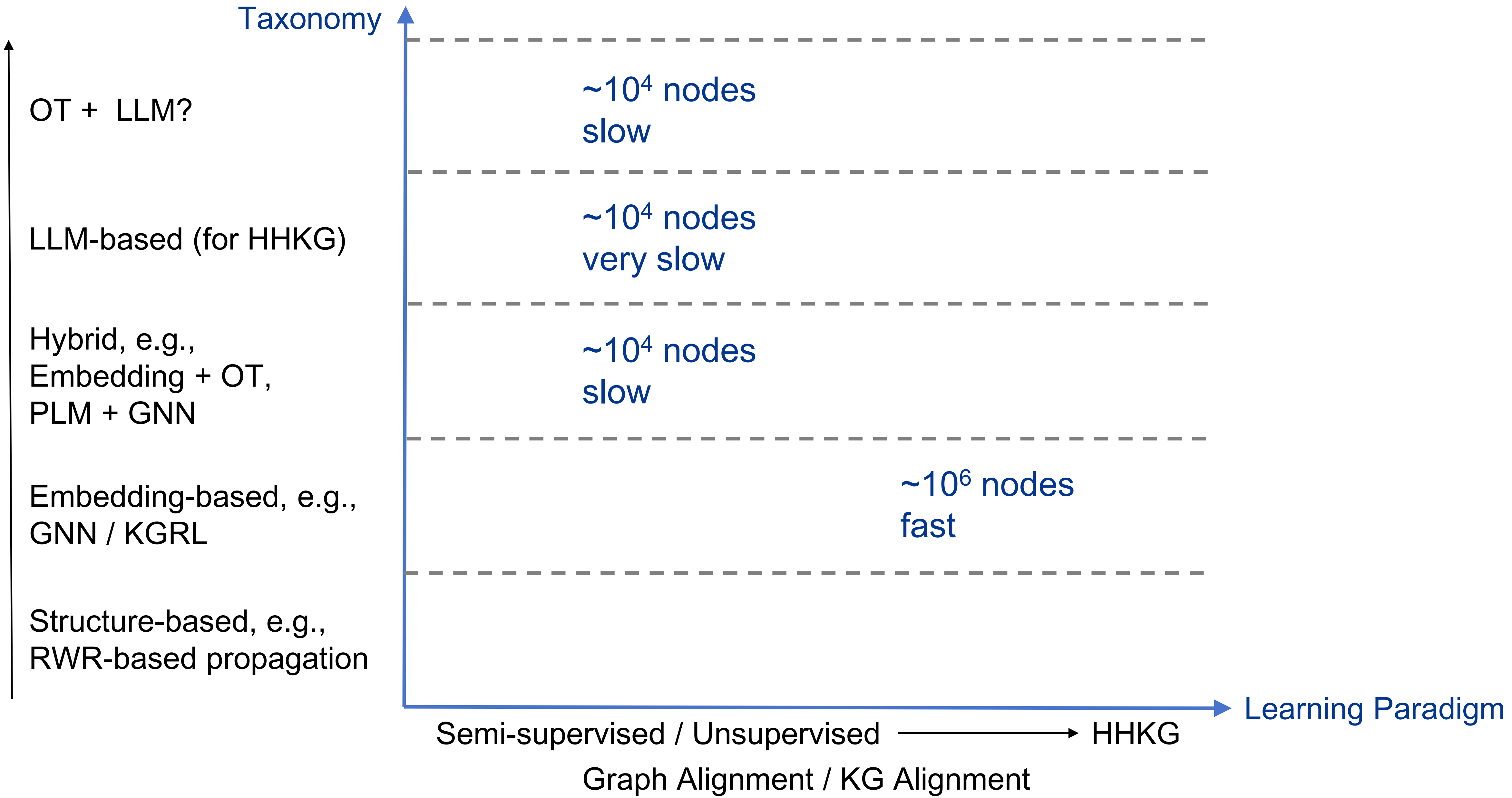
图对齐(Graph Alignment)问题

- 效率对比

Model	Tokens / $min(\|\mathcal{E}_1\|, \|\mathcal{E}_2\|)$	Time / $min(\|\mathcal{E}_1\|, \|\mathcal{E}_2\|)$
ChatEA	17032	230.1
AadCoAgentEA	3091	7.8
ULOT-HHEA	927	0.13

Table 4: Comparative efficiency analysis of entity alignment models across ICEWS-WIKI: average token consumption and execution time per aligned entity. lower values indicate better efficiency.

图对齐的挑战和开放性问题



谢谢
Q&A

Gromov-Wasserstein Learning算法

BEIJING JIAOTONG UNIVERSITY

Algorithm 3: Unsupervised Gromov-Wasserstein Learning with Feature Transformation (`GRAFT`)

Input: $\mathcal{G}_s$ and $\mathcal{G}_t$, marginal distributions $\boldsymbol{\mu}$ and $\boldsymbol{\nu}$

Output: The OT-based alignment probability $\mathbf{T}_{GW}$

$\mathbf{T}_{GW} \leftarrow \boldsymbol{\mu}\boldsymbol{\nu}^{\intercal}$, $\boldsymbol{\beta}_s, \boldsymbol{\beta}_t \leftarrow (1,1,1)^{\intercal}$;

for $i = 1$ *to* I **do**

  $\mathbf{Z}_s \leftarrow \texttt{FeatProp\&Trans}(\mathbf{A}_s, \mathbf{X}_s, \mathbf{W}^{(1,\ldots,K)})$;

  $\mathbf{Z}_t \leftarrow \texttt{FeatProp\&Trans}(\mathbf{A}_t, \mathbf{X}_t, \mathbf{W}^{(1,\ldots,K)})$;

  `// intra-graph cost (Eq. 6)`

  $\mathbf{C}_p \leftarrow f_{\boldsymbol{\beta}}(\mathbf{A}_p, \mathbf{X}_p, \mathbf{Z}_p), p = s, t$;

  `// inter-graph cost (Eq. 1)`

  $\mathbf{C}_{gwd} \leftarrow f_{gwd}(\mathbf{C}_s, \mathbf{C}_t, \mathbf{T}_{GW})$;

  Minimize $\langle \mathbf{C}_{gwd}, \mathbf{T}_{GW} \rangle$ by updating $\boldsymbol{\Theta} = \{\boldsymbol{\beta}, \mathbf{W}^{(1,\ldots,K)}\}$;

  `// the proximal point method`

  Initialize $\mathbf{T}^{(0)} \leftarrow \mathbf{T}_{GW}$ and $\mathbf{a} \leftarrow \boldsymbol{\mu}$;

  for $i' = 0$ *to* $I_{ot} - 1$ **do**

    Set $\mathbf{G} \leftarrow \exp\left(-\frac{\mathbf{C}_{gwd}}{\tau_T}\right) \odot \mathbf{T}^{(i')}$;

    `// Sinkhorn-Knopp algorithm`

    for $j = 1$ *to* J **do**

      $\mathbf{b} \leftarrow \frac{\boldsymbol{\nu}}{\mathbf{G}^{\intercal}\mathbf{a}}$;

      $\mathbf{a} \leftarrow \frac{\boldsymbol{\mu}}{\mathbf{G}\mathbf{b}}$;

    $\mathbf{T}^{(i'+1)} \leftarrow \mathrm{Diag}(\mathbf{a})\mathbf{G}\mathrm{Diag}(\mathbf{b})$;

  $\mathbf{T}_{GW} \leftarrow \mathbf{T}^{(I_{ot})}$;

return $\mathbf{T}_{GW}$;