



硕士生导师：**刘钰**（个人主页：<https://faculty.bjtu.edu.cn/9759/>）

联系方式：18701517219（微信同手机号），yul@bjtu.edu.cn

1. 让我们更感兴趣的你，是怎样的？

- **我们看重的：较好的科研习惯&一定的科研经历**
 - ★有**野心**参与高水平的研究、接受高水平的训练、收获高水平的成果
 - ★CCF推荐的A/B/C类会议期刊**投稿**经历（★一作更赞！）
 - 读过数据管理/数据挖掘/机器学习/人工智能**顶级**会议期刊论文（CCF-A）
 - 有过**进组实习**经历，独自承担和完成过无论大小的一个问题
 - 自认为在**思维直觉/理论分析/动手能力（各种比赛）**等某一方面略有所长

2. 我们能为你提供什么？

我们专注图（Graph）相关的研究，包括Graph Algorithms/Learning等...

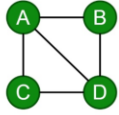
- **高水平的科研能力&合作团队**
 - ★小组常年在国际顶级会议期刊（CCF-A类）保持论文发表
 - ★导师主持**CCF-腾讯犀牛鸟科研基金/腾讯犀牛鸟专项研究计划/北京大学医信交叉种子基金**等高水平项目，参与多个**自科重点基金**和**北京通用人工智能研究院/华为2012/MSRA Research Grant**等多个高水平项目
 - ★导师获得中国工业与应用数学学会**应用数学落地成果认证**、CCF-腾讯犀牛鸟**科研基金优秀专利奖**等学会和企业认可
 - ★服务国家重大工程任务，面向具体技术应用落地
 - ★与北京大学PKUMOD实验室**紧密耦合**，与北大化院、人大高瓴、人大信息学院、中科院计算所等团队紧密合作
- **全方位的培养策略**
 - 在小组内/外派科研实习中做高水平的**科学研究**（顶会论文&高水平研究类项目）
 - 通过参加国际顶会/跨校合作交流/大量专家讲座/高水平研究类项目等**提升能力**
 - 支持2段分别为期3个月的**实习经历**（自己解决/导师推荐到大厂相关部门）
- **良好的科研环境条件**
 - ★**同学们都很优秀！**（近年来多数招生为本校保研/外校推免）
 - 小组及合作团队都具有良好的科研氛围和高水准的研究品味
 - 科研经费充足，硬件环境支撑到位
 - 良好的工位配置，人均面积>2.5m²（★**交大留学生创业园北206**，欢迎参观☺）

3. 我们期待你毕业后能收获什么？

- **心仪的工作**
 - ★拥有进入互联网大厂/外企/优质国企事业单位所需要的**科研能力、研究成果、实习经验和人脉关系**
- **较为优秀的科研素质**
 - ★以第一作者身份发表CCF-A论文/互联网大厂**菁英人才计划**...
 - ★我们的目标：**毕业时你的科研水平超过绝大部分入学时与你能力平齐的同学。**

图算法研究

SimRank节点相似度



	A	B	C	D
A		0.47	0.47	0.36
B	0.47		0.47	0.47
C	0.47	0.47		0.36
D	0.36	0.47	0.36	

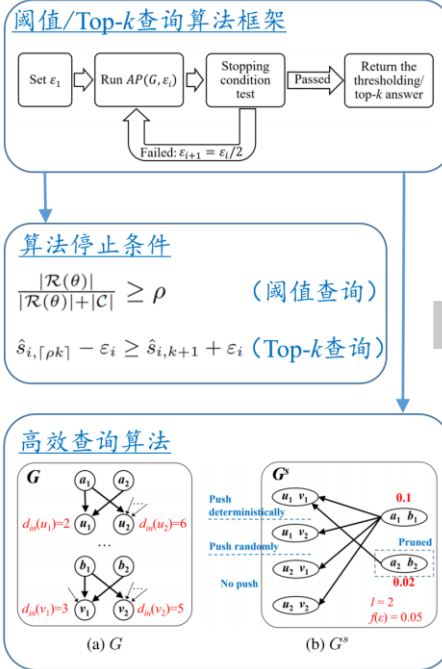
阈值查询

$\theta = 0.4 \rightarrow (A, B), (A, C), (B, C), (B, D)$

Top-k查询

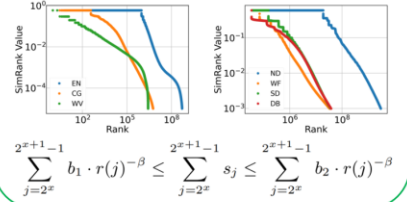
$k = 4 \rightarrow (A, B), (A, C), (B, C), (B, D)$

算法设计



理论分析

问题建模



理论复杂度结论

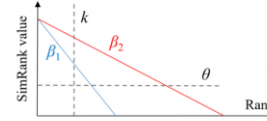
$$\tilde{O}\left(\sum_{u \in V} \sum_{v \in V} \sqrt{d_{in}(u)d_{in}(v)} s(u, v)\right)$$

阈值查询:

$$\epsilon_t = O\left(\max\left(\frac{(b_1 - \rho^\beta b_2)\theta}{b_1 + \rho^\beta b_2}, \epsilon_{min}\right)\right)$$

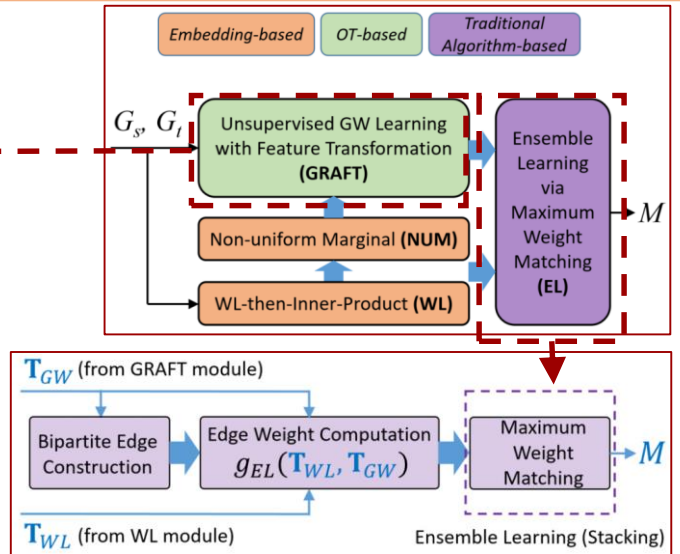
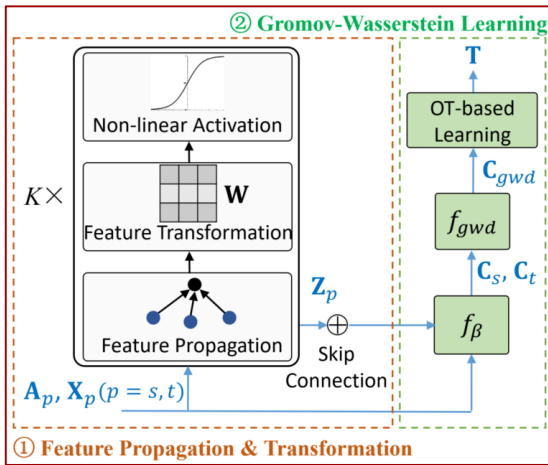
Top-k查询:

$$\epsilon'_t = O\left(\max(b_1(x_1 - t)^{-\beta} - b_2(x_2 - t)^{-\beta}, \epsilon_{min})\right)$$



- [Qian Ge, Yu Liu, et al.: Efficient and Accurate SimRank-based Similarity Joins: Experiments, Analysis, and Improvement. Proc. VLDB Endow. 17\(4\): 617-629 \(2023\) \(CCF-A\)](#)

图机器学习



Theorem 3.1. We are given the graph structures A_s, A_t and node features X_s, X_t as input. Denote feature propagation as $R_p = g(A_p)X_p, p = s, t$, where $g(\cdot)$ is a function without learnable parameters. Denote the additional linear transformation as $Z_p = R_p W$, where $W \in \mathbb{R}^{d \times d}$ is the learnable matrix. Assume that we set the intra-graph cost matrices as $C'_p = R_p R_p^T$ and $C_p = Z_p Z_p^T$ for $p = s, t$, respectively, and let $(u_i, v_k), (u_j, v_l) \in M^*$ and $(u_{j'}, v_{l'}) \notin M^*$ where M^* is the ground truth. Then, there exists a case that $|C'_s(i, j) - C'_t(k, l)| = |C'_s(i, j') - C'_t(k, l)|$ and $|C_s(i, j) - C_t(k, l)| \neq |C_s(i, j') - C_t(k, l)|$.

Theorem 3.5. Consider the case that $(u_i, v_k) \in M^*$ and $(u_i, v_{k'}) \notin M^*$. For GW learning (e.g., GRAFT), under the mild assumption of the intra-graph cost, i.e.,

$$C_s(i, i) = a, \forall u_i, C_t(k, k) = b, \forall v_k, \quad (16)$$

$$C_s(i, j) = C_s(j, i), C_t(k, l) = C_t(l, k), \forall u_i, u_j, v_k, v_l, \quad (17)$$

$$C_t(k, l) = C_t(k', l), \forall v_l \in \mathcal{V}_t \setminus \{v_k, v_{k'}\}, \quad (18)$$

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

- [Songyang Chen, Yu Liu, et al.: Combining Optimal Transport and Embedding-Based Approaches for More Expressiveness in Unsupervised Graph Alignment. arXiv \(2024\)](#)