

# 高崇铭 (Chongming Gao)

HomePage: <https://chongminggao.me>

☎ (+86) 15008242302  
✉ [chongming.gao@gmail.com](mailto:chongming.gao@gmail.com)  
🌐 <https://chongminggao.me>  
WeChat: 619082231

## 研究领域

数据挖掘与人工智能, 主要致力于图挖掘与分析、时空数据挖掘、推荐系统、自然语言处理等方向研究.

## 实习经历

2020.07–至今

科研实习

快手.

实习导师: 雷文强, 陆逊, 何向南.

实习成果: 投出一篇 ICDM'20 (待出结果). 简介在下面.

2019.03–2019.09

科研实习

阿里巴巴 AI Labs.

实习导师: 王浩, 聂再清.

实习成果: 投出一篇 ICDM'20 (待出结果). 简介在下面.

## 教育背景

2020.09–至今

博士

中国科学技术大学, 信息科学技术学院.

科研导师: 何向南.

研究方向: 对话推荐系统, 因果推荐.

2016.09–2019.06

硕士

电子科技大学, 计算机科学与工程学院.

科研导师: 邵俊明.

毕设论文: 《轨迹语义表征与地点推荐研究》, 校级优秀毕设.

2012.09–2016.06

本科

电子科技大学, 英才实验学院 (电子科技大学实验项目学院).

GPA: 3.81, 排名: 9/72

科研导师: 邵俊明, 大三学年开始科研.

毕设论文: 《基于双同步聚类的双聚类算法及其在基因表达数据上的应用》, 校级优秀毕设.

## 论文发表

IS'20

**Semantic Trajectory Representation and Retrieval via Hierarchical Embedding.**

高崇铭, 张众, 黄晨, 杨勤丽, 邵俊明

Information Sciences (IS'20).

(中科院 SCI 期刊分区 2020 年大类: 1 区, 小类: 1 区, CiteScore: 6.90, 影响因子: 5.524).

工作简介: 轨迹数据通常不定长, 使得表征与挖掘都困难. 本文用一种动态聚类的方法将轨迹表征成为一个层次语义网络. 在此层次语义网络上, 用网络 Embedding 的方式来重新表征轨迹, 使得区域与轨迹之间的语义相似度被重新定义. 基于此方式的轨迹检索效果好于基于传统 DTW、LCSS、EDR 等方法.

DASFAA'19 **Towards Robust Arbitrarily Oriented Subspace Clustering.**

最佳论文!

张众, 高崇铭, 刘崇志, 杨勤丽, 邵俊明

International Conference on Database Systems for Advanced Applications (**DASFAA'19**),  
(会议等级: CCF B)。

工作简介: 传统子空间聚类的方法总是受到局部和全局噪声的干扰, 且运行效率低。本文提供一种全新的子空间搜索方法思路, 效果鲁班, 速度快。

DASFAA'19 **\*BLOMA: Explain Collaborative Filtering via Boosted Local Rank-One Matrix Approximation.**  
(Short Paper) 高崇铭, 袁帅, 张众, 阴红志, 邵俊明

International Conference on Database Systems for Advanced Applications (**DASFAA'19**),  
(会议等级: CCF B)

工作简介: 基于矩阵分解的推荐系统存在一个大问题——分解出来的隐向量没有意义, 不具有解释性。本工作提出一种全新的可解释性推荐算法, 能对一次推荐自动作出解释: “本次推荐满足了您对于中餐 (60%) 以及海鲜 (%30) 的喜好。” 方法原理: 利用秩一分解, 每次从用户-商品矩阵中采样出最大的“尚未解释”的分量, 利用 side information 将其解释并消去。

KBS'19 **\*Semantic Trajectory Compression via Multi-resolution Synchronization-based Clustering.**

高崇铭, 赵奕, 吴睿智, 杨勤丽, 邵俊明

Knowledge-Based Systems (**KBS'19**),  
(中科院 SCI 期刊分区 2020 年大类: 1 区, 小类: 1 区, CiteScore: 7.01, 影响因子: 5.101)。

工作简介: 轨迹数据量大、不定长、采样率不一致等特性使得其存储和表示成为难题。本工作利用一种合适的动态性距离算法, 巧妙地将轨迹全局地表示成一个多层次网络, 从而达到灵活表征、快速传输存取的目的。

ICDM'19 **Online Budgeted Least Squares with Unlabeled Data.**

黄晨, 李培炎, 高崇铭, 杨勤丽, 邵俊明

IEEE International Conference on Data Mining (**ICDM'19**),  
(会议等级: CCF B)。

工作简介: 在线的半监督聚类通常要求动态维护拉普拉斯矩阵, 这样复杂度非常高。本文提出一种在线 Budgeted 式的最小二乘法, 使得在线半监督聚类变得很高效。本文提供了理论证明, 这种在线方式的误差比起离线算法是有界的。

ICDM'19 **Generating Reliable Friends via Adversarial Training to Improve Social Recommendation.**

余俊良, 高敏, 阴红志, 李俊东, 高崇铭, 王覃泳

IEEE International Conference on Data Mining (**ICDM'19**),  
(会议等级: CCF B)。

工作简介: 在目前考虑社交网络的推荐系统中, 由于社交网络的极度稀疏, 很多理论上奏效的方法效果并不理想。本文用对抗生成的方式, 为每一个用户生成一些靠谱的朋友, 再基于这些靠谱的朋友继续用对抗生成的方式做出推荐。整个过程动态循环。该方式超出了所有 social recommendation 的 baseline 方法。

DASFAA'19 **SemiSync: Semi-supervised Clustering by Synchronization.**  
(Short Paper) 张众, 康迪迪, 高崇铭, 邵俊明

International Conference on Database Systems for Advanced Applications (**DASFAA'19**),  
(会议等级 CCF B)

工作简介: 在半监督数据中有一种特殊形式: 已知必须连 (Must-link) 的边与必不能连 (Cannot-link) 的边, 需要聚类得到最终簇。本文提出一种基于动态聚类的全新半监督聚类算法。利用“吸引或者排斥”的交互法则, 让所有点与周围邻居动态交互, 得到稳态即为聚类结果。

ICDM'17 **\*Synchronization-inspired Co-clustering and Its Application to Gene Expression Data.**

邵俊明, 高崇铭, 曾伟, 宋井宽, 杨勤丽

IEEE International Conference on Data Mining (ICDM'17),

(会议等级: CCF B).

工作简介: 在例如“用户-商品”(或“基因-蛋白质”)表示的矩阵数据上, 传统聚类算法只能对用户或者商品进行聚类。本文提出一种“双聚类”的算法, 同时对用户和商品进行聚类, 得到有意义且解释性强的子簇(子矩阵)。

---

## 获奖情况

- 2019 DASFAA'19 (CCF B) 最佳论文奖。
- 2019 电子科技大学校级优秀硕士毕业论文 (86/3744)。
- 2019 电子科技大学校级优秀硕士毕业生。
- 2016 电子科技大学校级优秀毕设答辩专场荣获最高分 (94 分)。
- 2016 电子科技大学校级优秀本科毕业生。英才实验学院第 (10/72) 名。
- 2014 美国数学建模大赛一等奖 (M 奖)。
- 2013 国家数学建模大赛四川省一等奖。
- 2012 唐立新奖学金, 60/25000.
- 2012 电子科技大学在云南省录取最高分 (614 分)。

---

## 擅长语言

- 编程 Python, MATLAB, JAVA, C/C++,  $\text{\LaTeX}$ , HTML5+CSS3+Javascript
- 排版 Adobe Illustrator, Adobe Photoshop

---

## 兴趣

读书, 跑步, 游泳.