



浙江大學
ZHEJIANG UNIVERSITY

Simultaneous Matrix Orderings for Graph Collections

Jiacheng Pan

02/06/2022



作者介绍

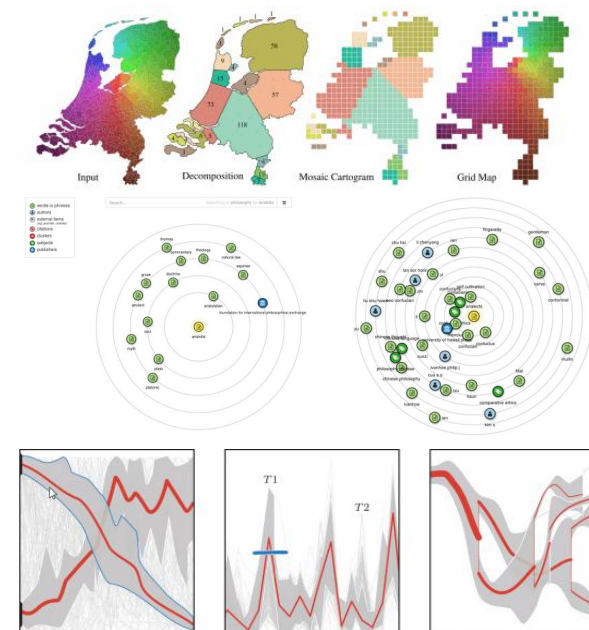
1. Nathan van Beusekom: Ph.D.
2. Wouter Meulemans: Assistant Professor
3. Bettina Speckmann: Full Professor



荷兰埃因霍温理工大学 (Eindhoven University of Technology)

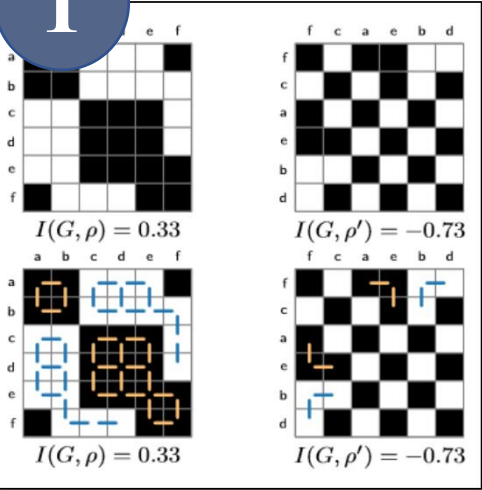
图数据、地理信息数据、时序数据

- A Simple Pipeline for Coherent Grid Maps
- SolarView: Low Distortion Radial Embedding with a Focus
- Multi-Granular Trend Detection for Time-Series Analysis



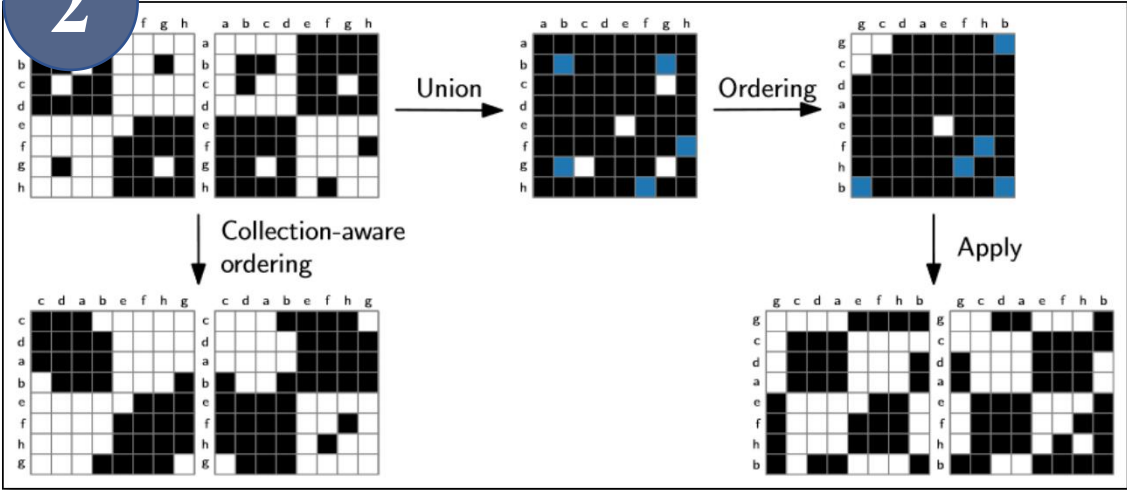
报告结构

1



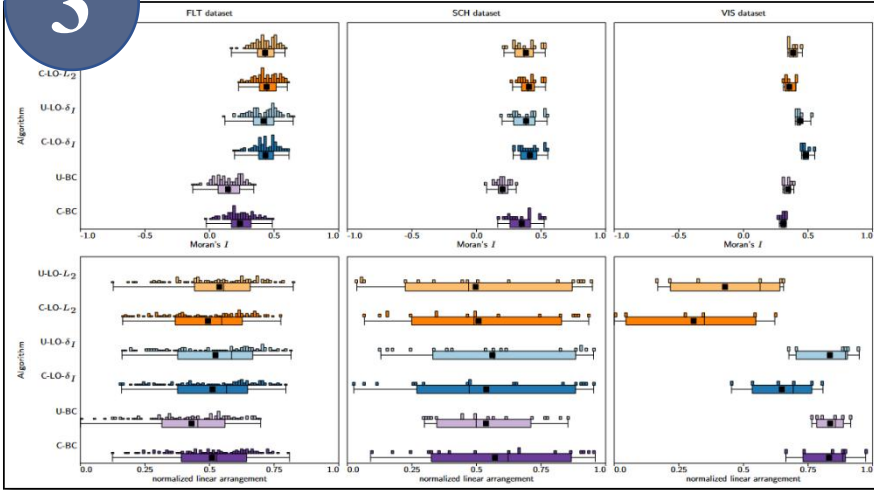
新质量度量方法

2



集合图数据的多矩阵排序

3



实验



图数据的可视化方法

- 节点链接图

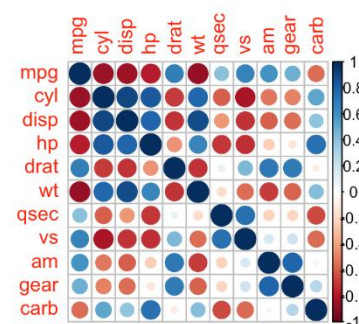
- 优点：直观
- 缺点：交叉、遮挡

- 矩阵

- 优点：无遮挡
- 缺点：理解困难



节点链接图



矩阵

- 矩阵可视化的关键问题：排序

1. 如何定义排序的质量？
2. 如何自动计算一个高质量排序？

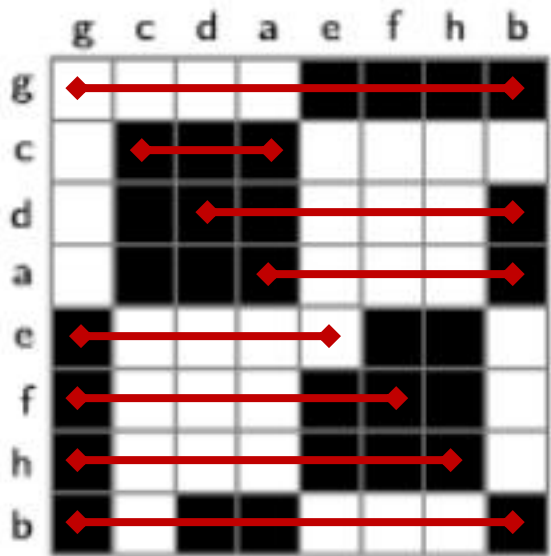


重排序算法



排序的质量 -- 传统度量方法

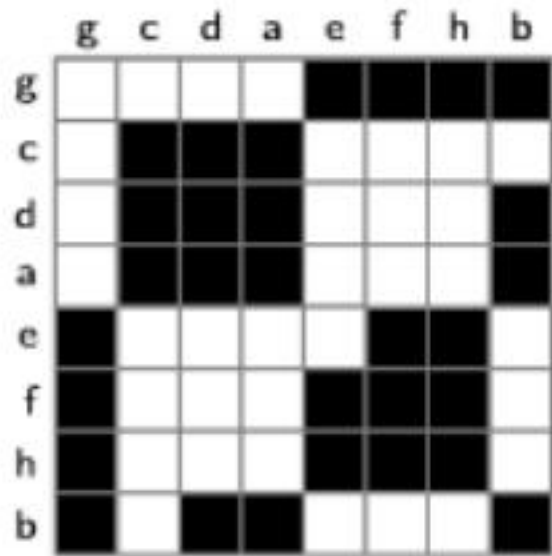
Linear Arrangement (LA)



每一行离对角线最远距离

$$\sum_{i=1}^n \max_{j < i \wedge \rho(j,i) \in E} \lambda(\rho(j), \rho(i))$$

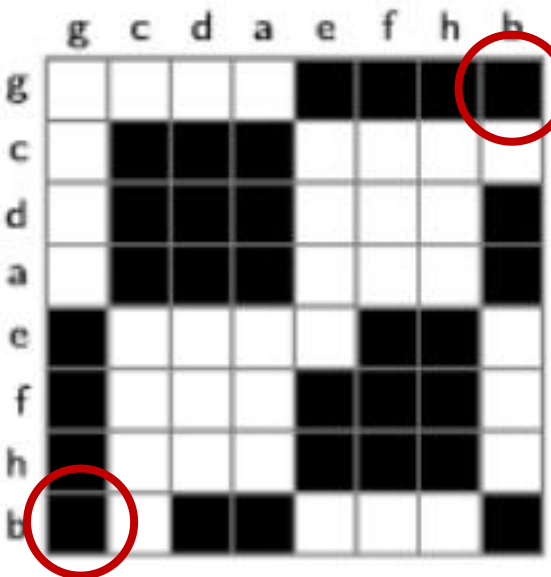
PRofile (PR)



每一格到对角线的距离

$$\sum_{(u,v) \in E} \lambda(u, v)$$

BandWidth (BW)



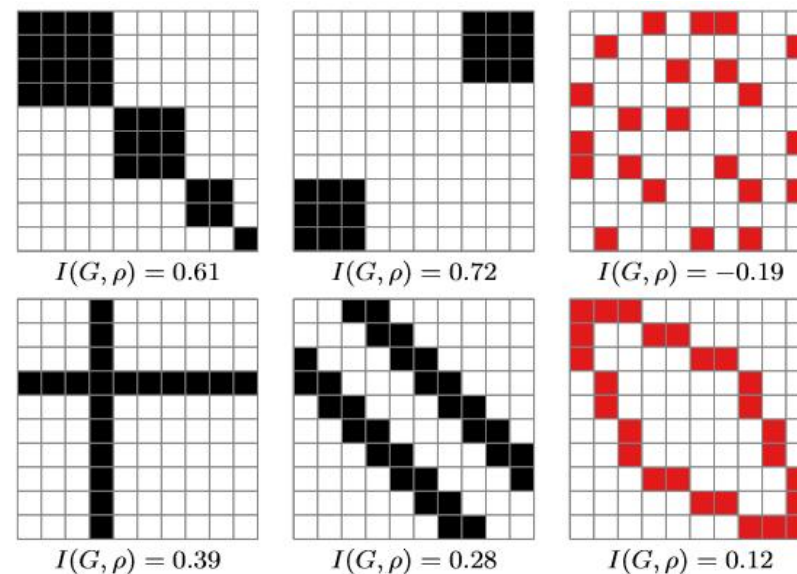
整个矩阵到对角线最远

$$\max_{(u,v) \in E} \lambda(u, v)$$



排序的质量 -- 传统度量方法的问题

- 关注对角线（下图左上角）而忽略其它有意义的模式（下图黑色）
- 更好的描述：邻接关系
 - measure of effectiveness [1,2]
 - stress measure [3]
 - Lenstra and Kan [4]



[1] W. McCormick, S. B. Deutsch, J. Martin, and P. Schweitzer. Identification of data structures and relationships by matrix reordering techniques. 1969.

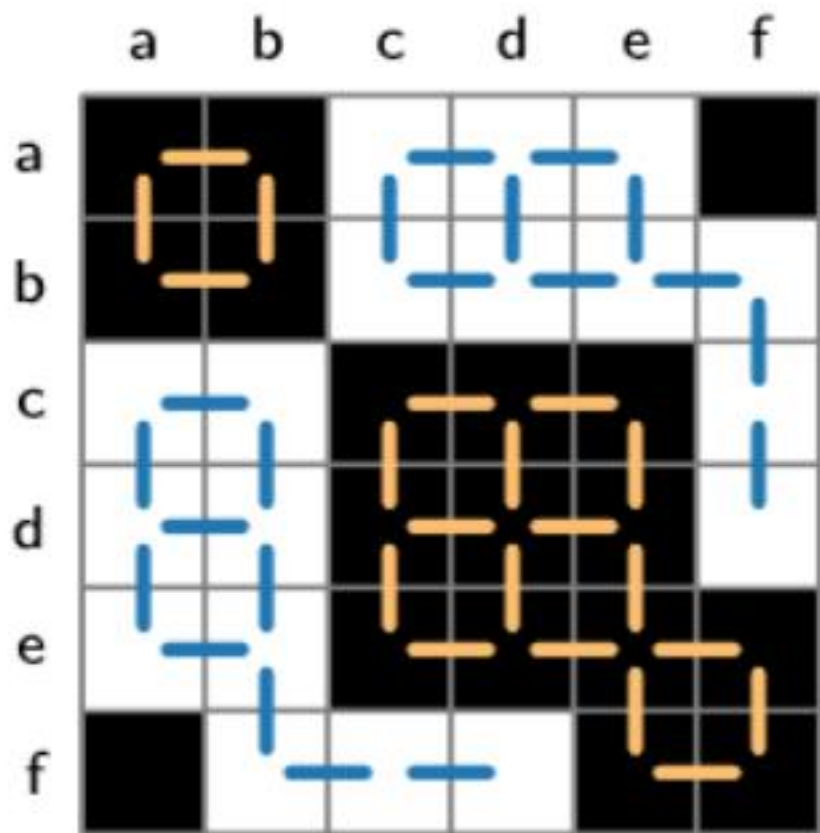
[2] W. T. McCormick, P. J. Schweitzer, and T. W. White. Problem decomposition and data reorganization by a clustering technique. *Operations Research*, 20(5):993–1009, 1972.

[3] S. Niermann. Optimizing the ordering of tables with evolutionary computation. *The American Statistician*, 59(1):41–46, 2005. doi: 10.1198/ 000313005X22770

[4] J. K. Lenstra and A. H. G. R. Kan. Some simple applications of the travelling salesman problem. *Operational Research Quarterly* (19701977), 26(4):717–733, 1975



排序的质量 -- 新的度量方法: Moran's I



$$I(G, \rho) = 0.33$$

Moran's I: 空间自相关度量

Notes on continuous stochastic phenomena, 1950

$$I = \frac{r}{t} \cdot \frac{\sum_{a=1}^r \sum_{b=1}^r T_{ab} (x_a - \bar{x})(x_b - \bar{x})}{\sum_{a=1}^r (x_a - \bar{x})^2}$$

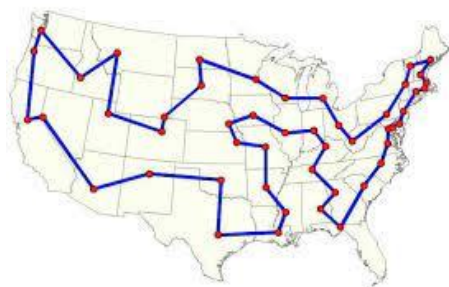
$$I(G, \rho) = c_B(G) \cdot \boxed{B(G, \rho)} + c_W(G) \cdot \boxed{W(G, \rho)} - 1$$

黑色连续块 白色连续块



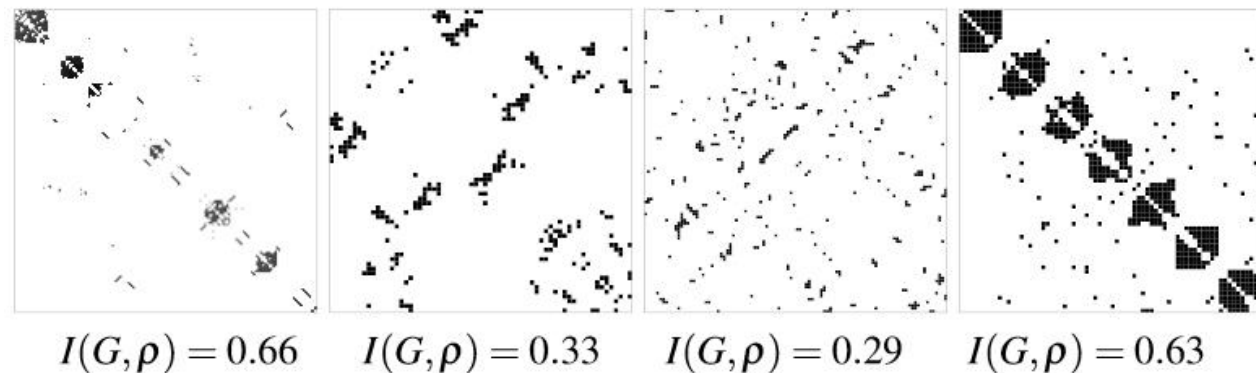
如何最大化Moran's I?

- 旅行商问题：最短距离路线

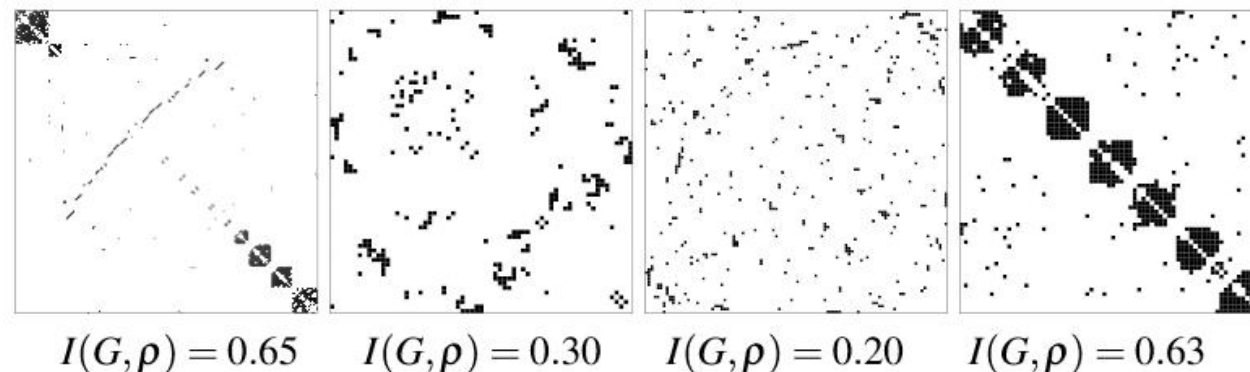


- 目标优化算法

- NN-2OPT算法, 1958 [1]



NN-2OPT算法，距离度量：Moran's I



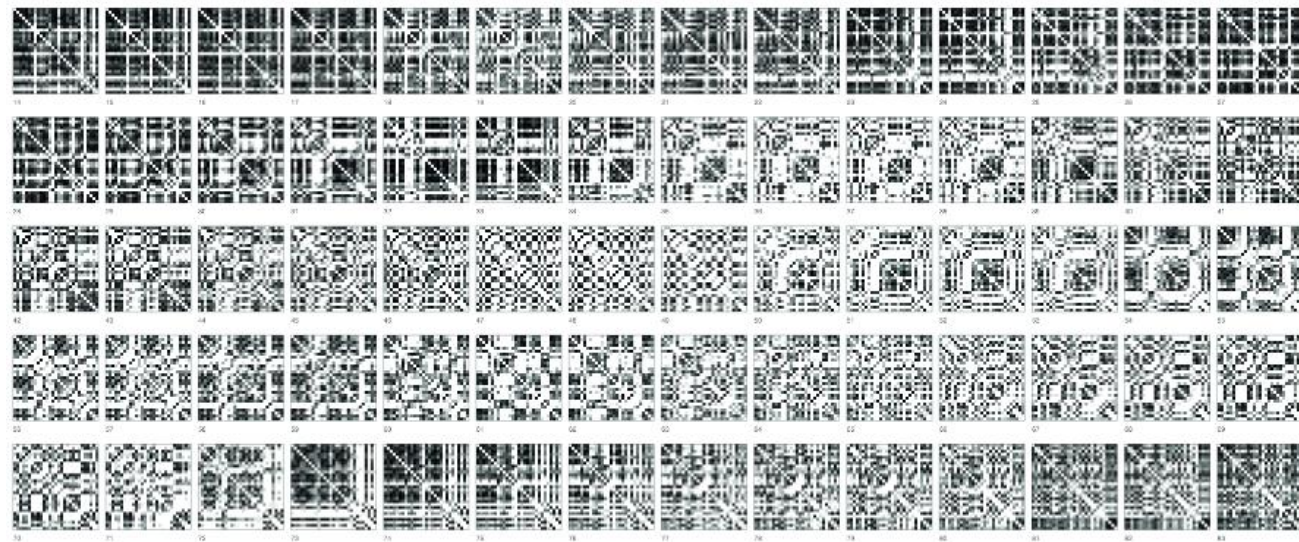
Leaf ordering算法，距离度量：L2

[1] G. A. Croes. A method for solving traveling-salesman problems. Operations Research, 6(6):791–812, 1958.

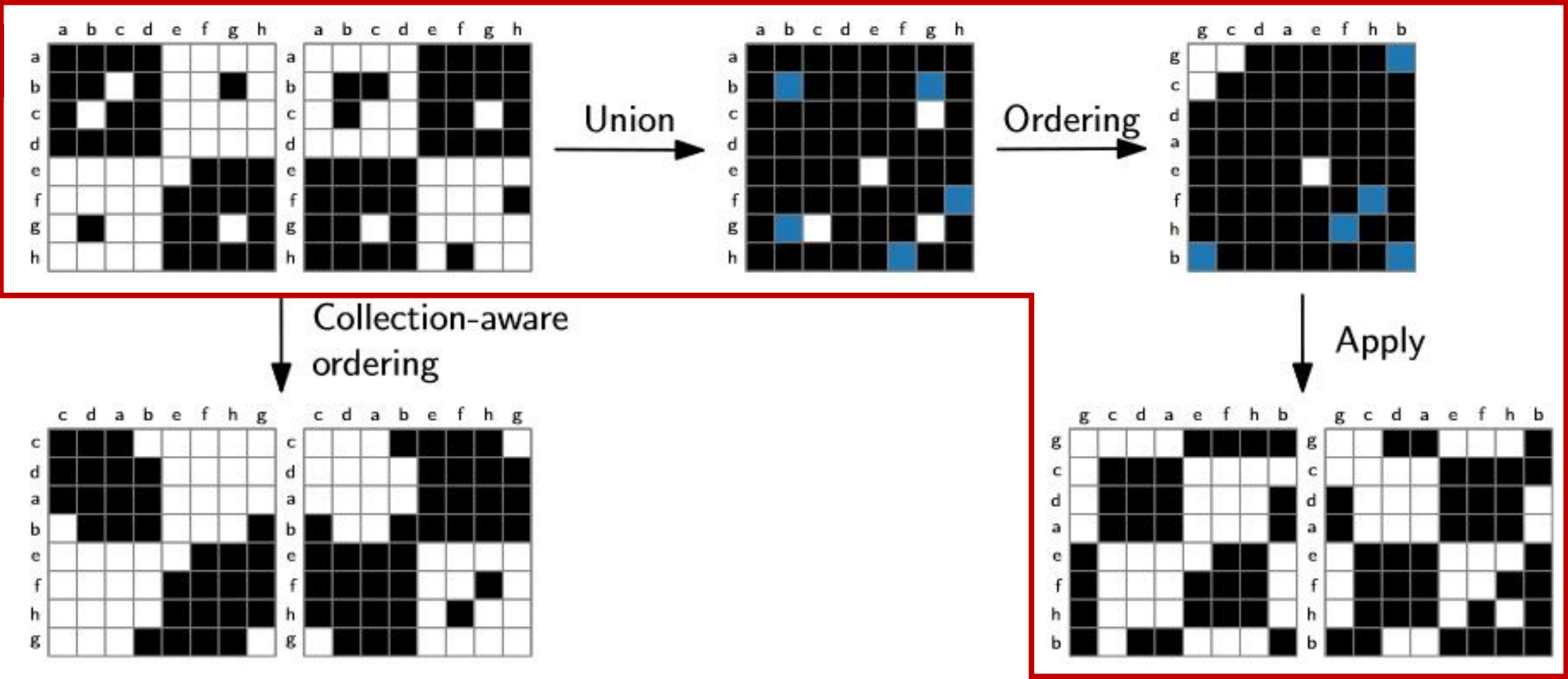


图数据集合的矩阵可视化

- 图数据往往以**集合**形式存在
 - 不同人、不同时间的大脑图像扫描
 - 学生和学生之间不同时间段的交流
 - 不同时段的可视化合作网络

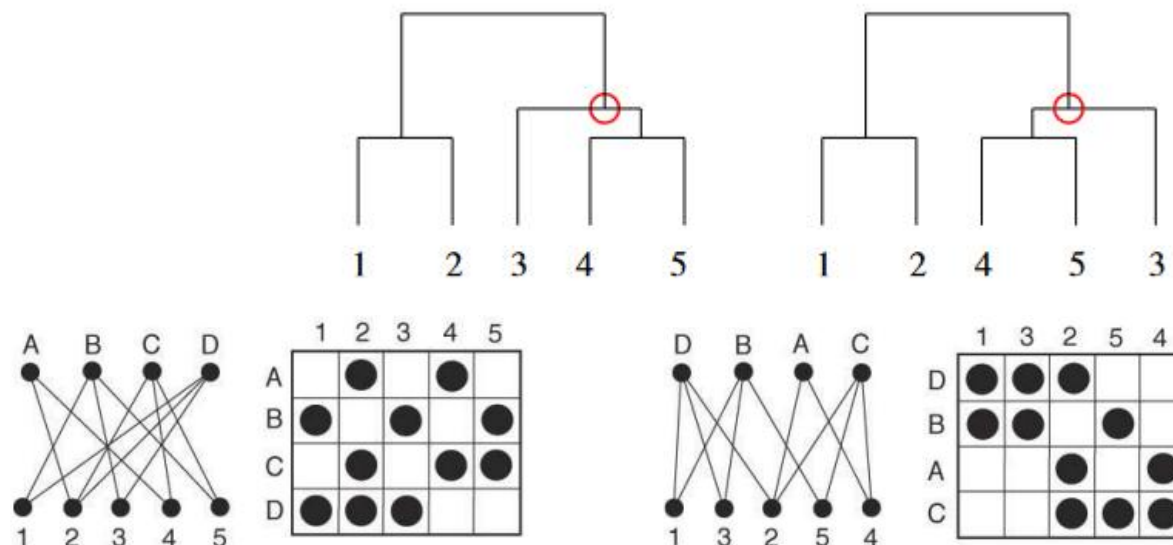


传统的矩阵集合重排序算法： Union, Ordering, then Apply



Weighted Matrix 重排序算法

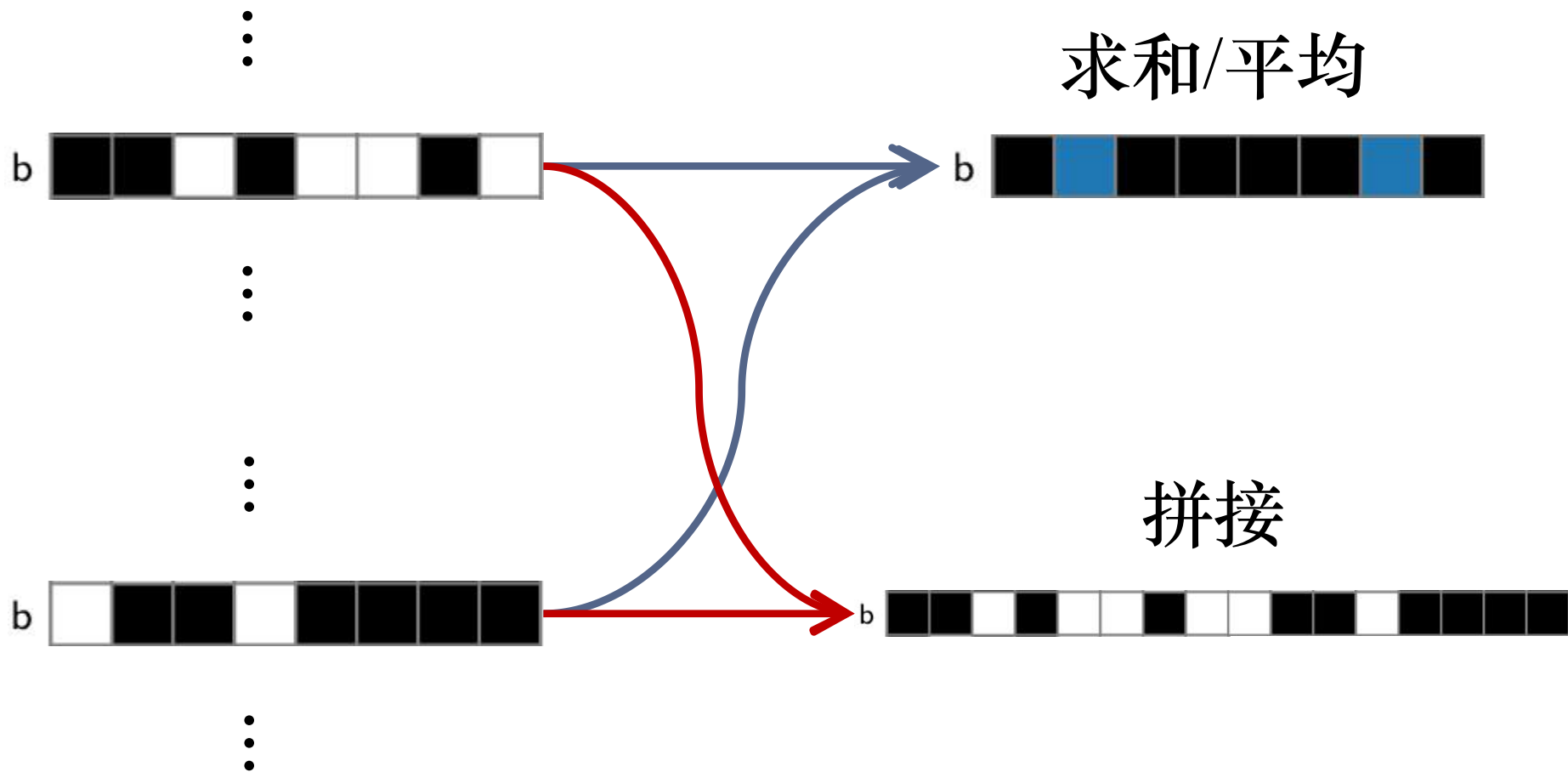
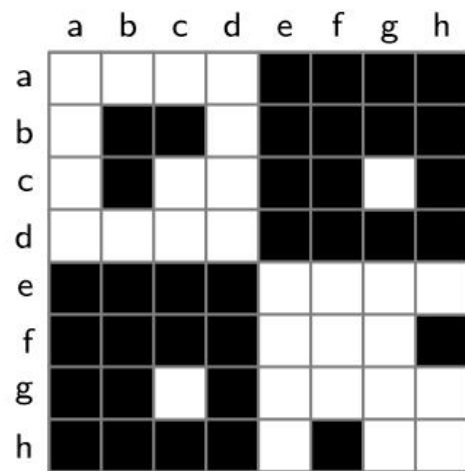
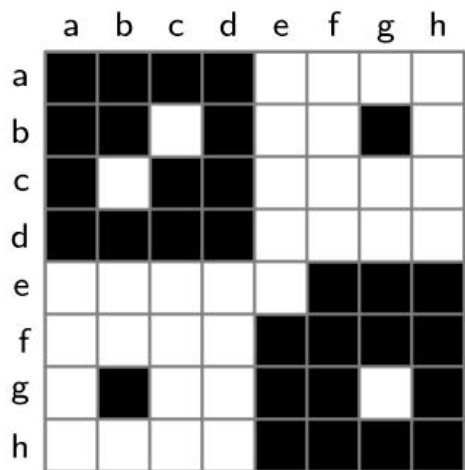
- Leaf order [1] (MultiPiles [2] 推荐): 用层次聚类树解决旅行商问题
 - 动态规划问题
- Barycenter method [3, 4, 5]:
 - 二部图的边交叉最小化
 - 贪心算法



- [1] Z. Bar-Joseph, D. K. Gifford, and T. S. Jaakkola. Fast optimal leaf ordering for hierarchical clustering. *Bioinformatics*, 17(suppl 1):S22–S29, 2001.
- [2] B. Bach, N. Henry-Riche, T. Dwyer, T. Madhyastha, J.-D. Fekete, and T. Grabowski. Small MultiPiles: Piling time to explore temporal patterns in dynamic networks. *Computer Graphics Forum*, 34:31–40, 05 2015. doi: 10.1111/cgf.12615
- [3] P. Eades and N. C. Wormald. Edge crossings in drawings of bipartite graphs. *Algorithmica*, 11(4):379–403, 1994.
- [4] E. R. Gansner, E. Koutsofios, S. C. North, and K.-P. Vo. A technique for drawing directed graphs. *IEEE Transactions on Software Engineering*, 19(3):214–230, 1993.
- [5] E. Mäkinen and H. Siirtola. The barycenter heuristic and the reorderable matrix. *Informatica (Slovenia)*, 29(3):357–364, 2005.



Collection-aware的矩阵集合重排序方法



实验

- **算法：**

- **集合重排序机制：(U) Union v.s. (C) Collection-aware**
- **重排序算法：(LO) leaf order 和 (BC) barycenter**
- **算法中使用的距离度量：欧式距离 L_2 和 莫兰指数 δ_I**

- **数据集：**

- **帕金森病患者大脑连通性 (96 graphs)**
- **小学学生和老师的社会交往联系 (17 graphs)**
- **15年到20年的VIS合作网络 (6 graphs)**

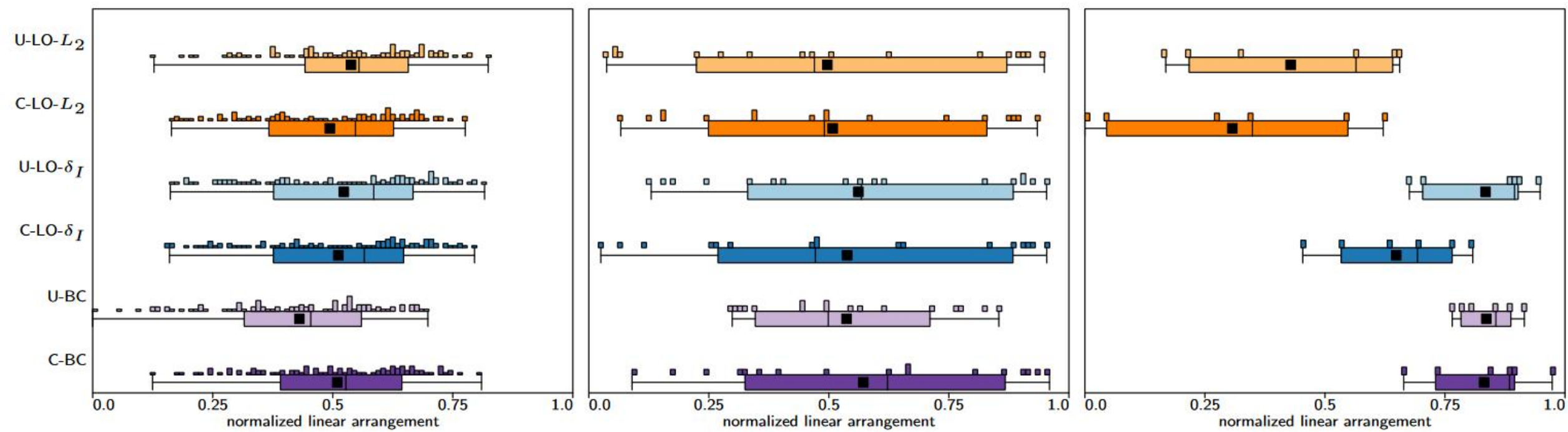
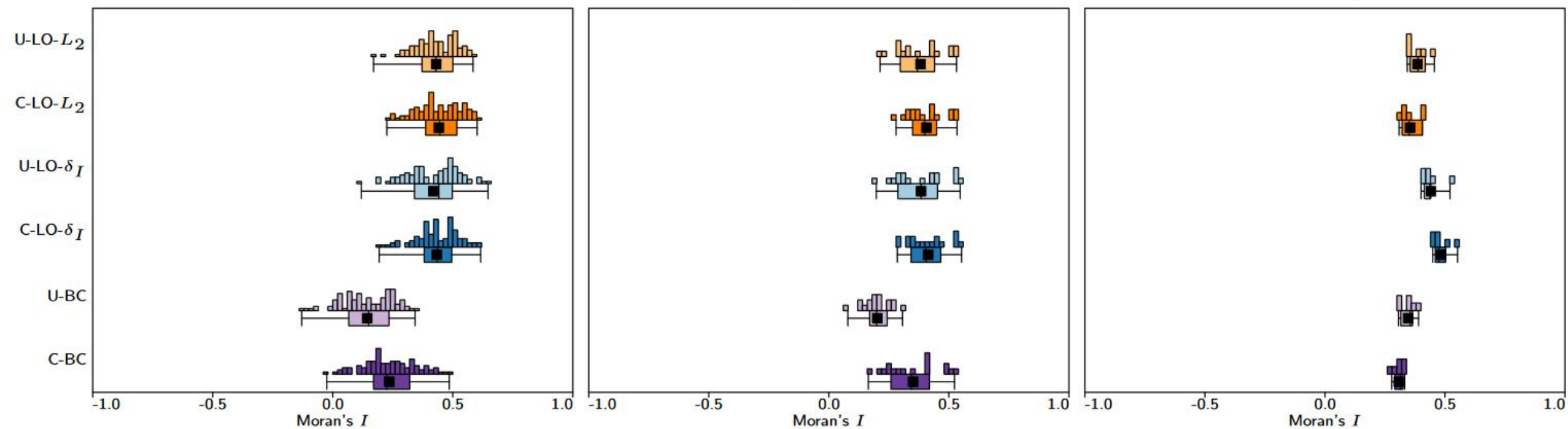
- **质量度量：Moran's I**



FLT dataset

SCH dataset

VIS dataset



讨论和未来工作

- 感知研究
 - 目前定量的度量方式与人的感知之间的关系
- 更多算法
 - 直接修改Leaf ordering算法的细节
- 不稳定的排序
 - 在某些矩阵上做出一些小的改变，让它的结构更明显
- 个人评论
 - “既做裁判又做运动员”
 - 简单的算法解决一个冷门的问题





浙江大學
ZHEJIANG UNIVERSITY

Thanks & Questions

Jiacheng Pan

02/06/2022

