

## CoSTUR:面向用户评级的空间文本竞争选址研究

李晨伟<sup>1</sup>, 王蒙<sup>1\*</sup>, 李辉<sup>2</sup>, 崔江涛<sup>2</sup>, 王西汉<sup>1</sup>

(1 西安工程大学, 2 西安电子科技大学)

[wangmeng@xpu.edu.cn](mailto:wangmeng@xpu.edu.cn)

**摘要:** 随着 GPS 定位技术和移动互联网的发展, 各类 LBS(Location-based services)应用积累了大量带有位置和文本信息的空间文本数据, 这些数据广泛应用于市场营销、城市规划等设施选址决策场景中。代表性空间文本选址方法借鉴反向最近邻思想, 从候选位置集合中挖掘出最佳位置新建设施, 以期影响最多具有空间文本属性的对象, 如用户或车辆等, 其中空间愈邻近且文本越相似则影响力越大。然而, 现有方案不仅未考虑现实场景中普遍存在的同行竞争, 也忽略了用户对竞争设施的评级(例如通过在线消费平台点评)。此外, 传统模型假设空间文本对象只被对其施加影响力最大的设施吸引, 这也与真实世界中用户可能同时受多个设施影响的情况相背离。为更合理地在同行竞争环境结合用户评级因素进行选址决策, 提出新的空间文本竞争选址问题 CoSTUR。为解决不同评级设施间竞争量化、用户同时受多个设施影响建模、以及数据量较大导致计算代价急剧上升等问题, 构建了新的解决框架。首先, 借鉴现有竞争均分模型, 结合用户评级完善竞争影响力定义。其次, 利用可权衡的影响力阈值摆脱用户只受单个设施影响的局限。最后, 构建基于 aR-tree 思想的空间索引结构, 实现空间与文本联合检索, 并结合阈值设计了基于用户所受影响范围的新剪枝策略。在真实和合成数据集上的实验结果表明, 利用所设计树索引的两种解决方案, 相比基线算法能够提升近一个量级的计算效率。

**关键词:** 空间文本数据; 选址问题; 空间文本索引; 竞争影响; 多设施影响; 用户评级