**服务计算前沿技术**

课程实验：Skyline 服务查询

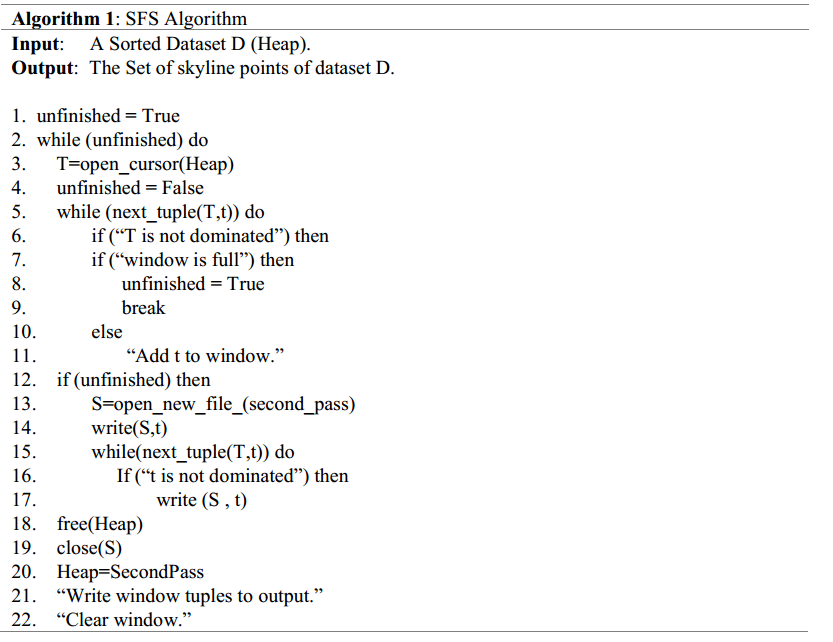
**一、实验算法**

Skyline 服务查询是在进行服务推荐的一个重要步骤，其查询速度和准确性关系到服务推荐的效果。

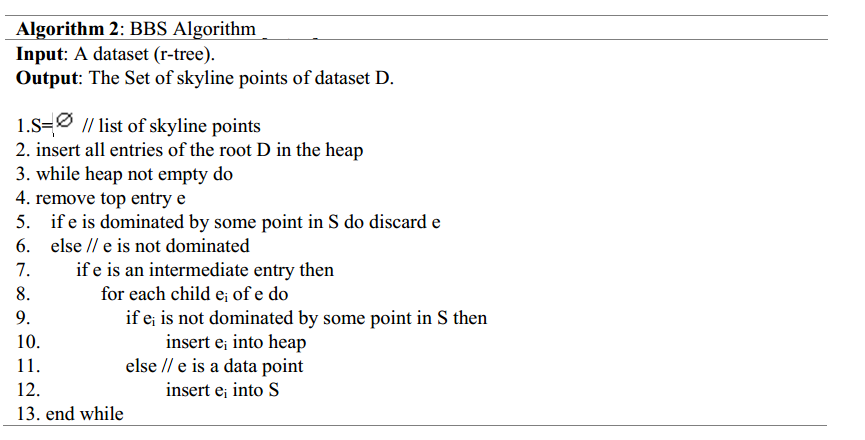
BNL（Borzsony, S.; Kossmann, D.; Stocker, K （2001）. “The Skyline Operator”，in proceedings of 17th International Conference on ICDE:421-430）是最早提出的skyline查询基本算法之一。

在BNL的基础上，又有以下两种常用的改进skyline算法：

1．SFS （Jan Chomicki, Parke Godfrey, Jarek Gryz, and Dongming Liang. (2003) “Skyline with presorting.” In ICDE: 717–719.）



2．BBS（Dimitris Papadias, Yufei Tao, Greg Fu, and Bernhard Seeger (2003). “An optimal and progressive algorithm for skyline queries.” In Proceedings of the 2003 ACM SIGMOD international conference on Management of data, SIGMOD ’03: 467–478.  
Dimitris Papadias, Yufei Tao, Greg Fu, and Bernhard Seeger (2005). “Progressive skyline computation in database systems.” ACM Trans. Database Syst., 30(1): 41–82.）



**二、实验要求：**

1 请在公开数据集QWS上，实现BNL方法，以及BBS或SFS方法，以验证算法实现的正确性。

2 请在仿真工具randdataset产生服务质量QoS数据集（包括：独立，正相关、反相关，数据集至少需要10K\*10维，20K\*10维，50K\*10维）上，实现BNL的skyline运算。  
3 在同样数据集上，采用BBS或SFS查找Skyline服务（二者选择一种实现即可），并对结果和运算速度与BNL方法进行比较。  
4若选择BBS，其中r-tree可以使用现成的包（可以在GitHub上寻找适合的r-tree库自己修改）。  
5 实验没有语言限制，推荐使用python或java，以便于验收。

注：群文件中的归一化QWS数据集，属性值都是越大越好

**三、实验报告要求**

1. 运算结果正确性分析；
2. 运算速度分析比较；
3. 实验结果的启示（如：有什么方法能够更好的提高速度）。

报告格式自定，必须包括以上内容。