**近似查询设计文档**

夏青扬 计14 2011011276

**设计思路：**

按照课程PPT上的算法思路一步一步地实现了近似查询的功能。在createIndex()函数中，使用STL中的map来做倒排列表，但是在倒排列表中没有删去重复的id，放在jaccard中实现。在searchED()函数中，使用scanCount算法实现初步的过滤统计，与课件中不同的是在统计的时候同时记录达到阈值要求的id，可以达到一些优化的作用。

**实现难点：**

由于一直没有考虑重复的gram，导致在DEADLINE前一直不能得到正确的结果，发觉问题修改后很快可以得出答案。