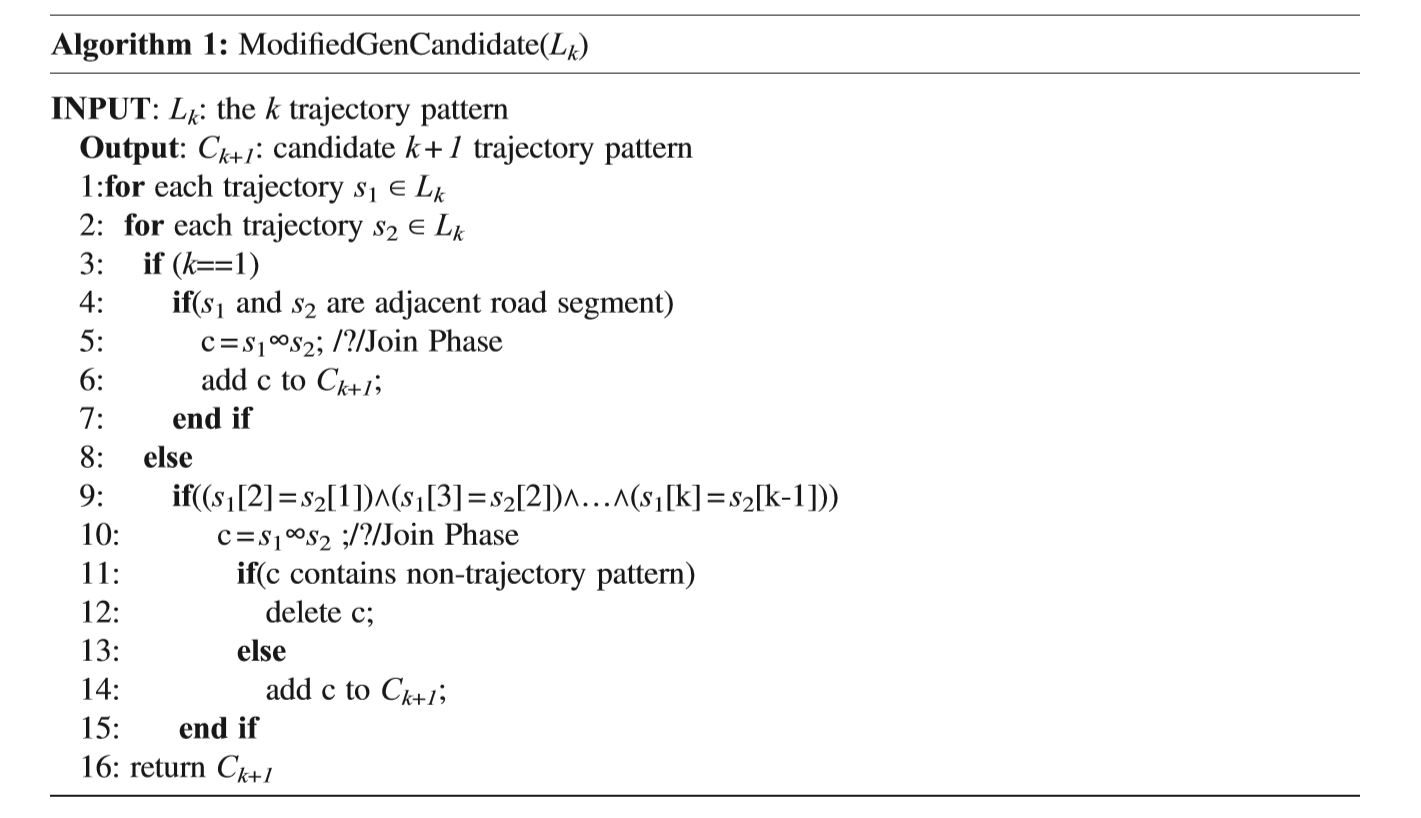
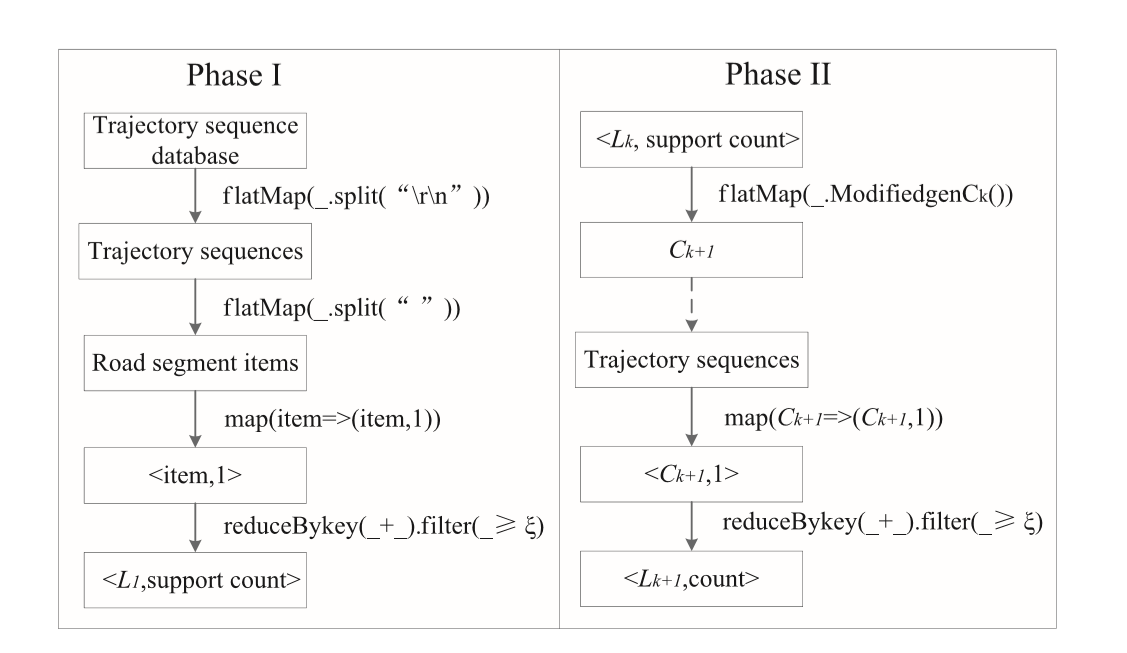
需求就是基于spark框架 使用java或者py语言，实现GSP和Prefixspan算法，对gsp进行连接限制，对Prefixspan进行连接限制

1 连接限制就是在生成gsp的候选集和prefixspan的前缀时，加上限制条件做链接时只有在原数据集里相邻的项做链接，相当于做了个剪枝

2 GSP连接限制伪代码：



GSP流程图：



3 Prefixspan在 spark的ml库（mllib/src/main/scala/org/apache/spark/mllib/fpm/PrefixSpan.scala）基础上改

4 输出运行时间