实现的热点用户打压的余弦相似度协同过滤算法公式：

Sim(I\_i,I\_j)=

# 算法流程

* 数据输入

User1: item1,item2,item3…

User2: item2,item5,item8,item7…

……

* 统计每个user的行为次数，记RDD为user\_times
* 统计每个item的行为次数，记RDD为item\_times
* 生成item1,item2的pair：对每一条有n次行为的用户数据，可以生成个i1i2-pair（同时加入分子上的用户热度），记RDD为i1i2
* 将i1i2按key为(i1,i2)的形式reduceByKey，累加得分，得到i1,i2:score，变换为i1:i2,score后再join item\_times，并除以item1的权重，输出i2：i1,score再与item\_times join，并除以item2的权重，最后输出i1:i2,score
* 对i1，按score得分降序排列i2，取topN的i2作为推荐列表。