

## 对毕设课题的理解

11612126 李可明

## 问题:

Cover Ratio Maximization (user distribution  $\theta$ ):

Given a dataset  $D$ , product group  $P = \{p_1, p_2, \dots, p_m\}$ , product creation budget  $B$ , a positive integer  $k$  and a creation cost function  $C$ , CRM introduce a new product  $p$  such that  $C(p) \leq B$  and  $cp(p, P, k) = \frac{|\{w | \forall w \in W, \{P \cup \{p\}\} \cap \text{TopK}(w) \neq \emptyset\}|}{|W|}$  is maximized

对问题的理解:

我对  $D$  的理解是 product 的一个集合, 是离散的

$P$  是另一个 product 的集合, 也是离散的

某些 product in  $P$  在  $P \cup D$  respect to some user preference vector 能排前  $k$ , 覆盖了一部分 user preference vector

我们要找一个新的满足约束  $C(p) < B$  的 product  $p$ , 使得  $P \cup \{p\}$  覆盖最多的 user preference vector

解决问题路线:

可以结合 k-hit query 的思路对 cell tree 修改, 这里应该可以有一个 baseline, 原先希望这周能想出 baseline 的思路但还有些问题。

因为 k-hit query 是 top1, 用到一个 k-hit query 自己定义凸包 (convex hull), 只有凸包上的点才有可能排 top 1。

我们的要做的是 topk, 这意味着我们

1. 首先要找到哪些点有可能会排 top k(或者说哪些点根本排不了 tok), 找到这个解集  
这里可能会修改 k-hit query 的某些部分, 由于是 topk 不是 top1 还会用到修改的 cell tree
2. 根据约束  $C(p) < B$  挑选合适的  $p$   
这里一般来说就是在约束边界上找解, 这意味着第 1 步根据这个约束能有一些剪枝
3. 由于  $P$  已经覆盖了一部分 user preference space, 找到 cover ratio maximization 的新的  $p$  并不简单, 但一般来说解就在约束条件上

但是在第 2 步中由于找到的解集肯定是连续的区域解集, 这意味着我们要在连续的 product space 找到一个解, 问题已经不是单纯的离散的 product space

Baseline 暂时还有点问题, 需要点时间给出来

计划是希望

1. 寒假能将 baseline 的思路， 每一步的证明完善以及代码基本框架做出来
2. 回来后跟老师讨论 baseline 的理论正确性， 哪些地方要改
3. 回来第一个月(2 月)根据跟老师讨论的结论得到正确的 baseline 或者更好的算法
4. 回来第二、三个月(3 月， 4 月)是代码实现跟测试
5. 4 月、5 月跑实现， 写论文

## 问题 2

Given a dataset  $D$ , a user preference set  $W$ , a cover percentage  $\tau$ , a positive integer  $k$  and a creation cost function  $C$ ,  $CCM$  introduces a new product  $p$  such that the cover ratio

$$cp(p, k) = \frac{|\{w | \forall w \in W, p \text{ is in } TopK(w)\}|}{|W|} > \tau$$

and the creation cost  $C(p)$  is minimized

对问题的理解：

$D$  是 product space 的数据集， 是离散的

$W$  是服从某种分布的 user preference vector 的集合

Product space(这个是连续的空间) 上找到所有满足覆盖率不低于  $\tau$  的  $p$ ， 在这些  $p$  中筛选代价 creation cost 最小的一个解