Truth Inferences based Combinatorial Multi-armed Bandit Scheme with some Workers Failure for Mobile Crowdsensing

Ke Fang, Anfeng Liu\*

*Abstract*—我们想做的大致的想法是这样的. Gao G, Wu J, Xiao M, et al. Combinatorial multi-armed bandit based unknown worker recruitment in heterogeneous crowdsensing[C]//IEEE INFOCOM 2020-IEEE Conference on Computer Communications. IEEE, 2020: 179-188. 是说, 有多个任务, platform招募多个workers来感知数据. 这些 workers感知的数据的质量不相同. 而高的质量则能够获得高的收益, 因而, 应该选择高质量的workers来感知数据. 但是, workers感知数据的质量是不知道的. 这个论文是假设如果workers提交了数据的话, 则platform就能够知道workers的数据的质量. 这样, 他就采用多臂赌博机问题(Multi-armed bandit problem, K-armed bandit problem, MAB)的方法来解决这个问题. 他这论文的关键是公式9, 10, 11. 公式10是选取workers的概率, 这是通用的MBA问题的选取概率的计算公式. 公式10有2项, 一是workers的质量高, 则其选取的概率高, 而第2项, 如果workers选取的次数多, 则其选取的概率下降. 这样做的目的是: exploitation, exploration. Exploitation是workers提交了数据, 就能够知道其质量, 因而, 当有一些workers提交了一些数据后, 我们就能够得到其质量, Exploitation显然就是利用的信息, 从已经知道workers质量的这些workers中选取质量高的workers来做任务. 但是, 已经知道的workers是一部分. 还有一部分workers的质量未知, 这些质量未知的workers中也有可能存在质量非常高的workers, 因而, 如果仅从已经知道质量的workers中选取, 则后面那些高质量的workers就不会选到. 因而, 还要exploration, 就是还是要从未知质量的workers去选取一些, 如果取到高的质量的workers就能够获得更优秀的结果. 因而, MBA是exploitation与 exploration结合．

其选取的公式10就是这样的意思, 高质量的workers选取一定次数后，其被选取的概率越减少，从而会选取其它workers，就是探索．这个论文然后把公式10放入到公式11得到收益, 公式12选取能够给系统带来单位成本收益最大的workers就会被选取.

但是, 上面的研究中假设工人一提交数据, 那么就知道数据质量, 但是, 实际上工人的数据报上来了后, 平台也是不知道质量的. 因而, 我们就不能象上面那样的做. 我们就要想办法推测获得工人的质量. 我们怎么获得呢. 我们假设有一小部分的工人的信任度是可信的, 就是每次提交真实的数据. 比如10000个工人我们事先知道100个工人是可信的.

其次, 我们还要检验工人的真实信任度是多少? 我们选择工人的时候, 一边选取高信任度的工人, 同时也选取一部分信任度未知的工人来确定其信任度. 如果待检验的工人报的数据与可信的工人一致, 则其信任度就提高, 反之, 就下降. 如果一个工人反复得到检验, 信任度上升到一定度后, 就是可信的工人, 可信的工人就会不断的探索出来,增大, 这个过程一直下去, 就能够获得更多的工人的信任度. 我们记工人个人信任度为. 这们我们把公式10个的用来代替, 这样, 在开始的时候, 知道信任度的工人有确定的值, 不知道的工人的个人信任度设为0.5. 这样, 在选取的时候,个人信任度都高的就会先选取. 为每一个人选取高可信的工人来做.

我们为了选取一部分workers来检验, 就改造公式10, 检验也能够带来收益. 把检验的收益加到公式10中. 就是说, 我们先取这个workers能够检验到其它工人的信任度, 也会对系统有收益, 因而, 加上这部分收益, 如果没有能够得到检验信任度, 则这部分收益为0. 公式10就有3项了.

我们还要做的是. 在上面, 我们假设工人执行任务都会做, 而且会成功, 但是有些工人会失败的. 因而, 我们采用下面的方法来补全工人失败的地方. 第二: 有很多地方是没有可信的工人的, 工人得到检验. 因而, 我们如果得到部分地方的数据(有可能工人的地方), 我们采用下面这个论文的把其它没可信工人的数据补全, 用补的数据来检验工人的信任度. 这样就能够加快探索工人信任度. 能够获得更好质量.

Marchang N, Meitei G M, Thakur T. Task reduction using regression-based missing data imputation in sparse mobile crowdsensing[J]. The Journal of Supercomputing, 2022: 1-34.

*Index Terms*—Multi-armed bandit

# Introduction

————————————————

1. Manuscript received xx xx, 2022; revised xx xx, 2022. This work was supported by the National Natural Science Foundation of China 61772554. (\*Corresponding author: Anfeng Liu).
2. K. Fang and A. Liu are with School of Computer Science and Engineering, Central South University, ChangSha 410083 China. (e-mail: \*\*\*\*\*\*\*@csu.edu.cn, Afengliu@mail.csu.edu.cn).
3. K.).

Consider such a scenario where a requester wants to recruit workers to collect the traffic data (e.g., traffic photos or videos) at some urban intersections for a period of time. The whole data collection is divided into multiple rounds. In each round, it consists of many location-related sensing tasks, each of which corresponds to a traffic intersection, as shown in Fig. 1. Here, each task is attached with a weight to indicate its importance. Each worker can complete (a.k.a., cover) one or more tasks. The tasks that each worker can deal with might be different, i.e., the sets of tasks covered by different workers are heterogeneous. All workers will tell the platform the tasks they want to perform and the costs they expect to charge. Each worker can provide multiple options, composed of different task combinations and costs, but at most one option will be selected. Moreover, each worker has a sensing quality, following an unknown distribution. Our objective is to design a worker recruitment strategy that can maximize the total task completion quality under a given budget. The main challenge lies in that the platform does not know workers’ sensing qualities in advance, so it needs to learn their quality values by tentatively recruiting workers to complete some tasks and then selects the best group of workers according to the learned results. Generally, the two processes are called exploration and exploitation [18], [19], respectively. We need to balance the two processes so as to maximize the total task completion quality under a given budget. To address this challenge, we model the unknown worker recruitment process as a novel Combinatorial Multi-Armed Bandit (CMAB) problem, where each worker is seen as an arm, its sensing quality is seen as the corresponding reward, and recruiting workers is equivalent to pulling arms. Moreover, we let a fixed number of arms (i.e., K) be pulled in each round. Our CMAB model has two novel characteristics, different from all the existing CMAB models. First, each arm has multiple options, each of which corresponds to a set of covered tasks and a cost. The platform needs to not only select arms but also determine the option for each arm. Second, it contains a budget-limited maximum weighted coverage problem (i.e., maximizing the total task completion quality, which involves a weighted sum function on some maximum sensing qualities), making it very challenging.

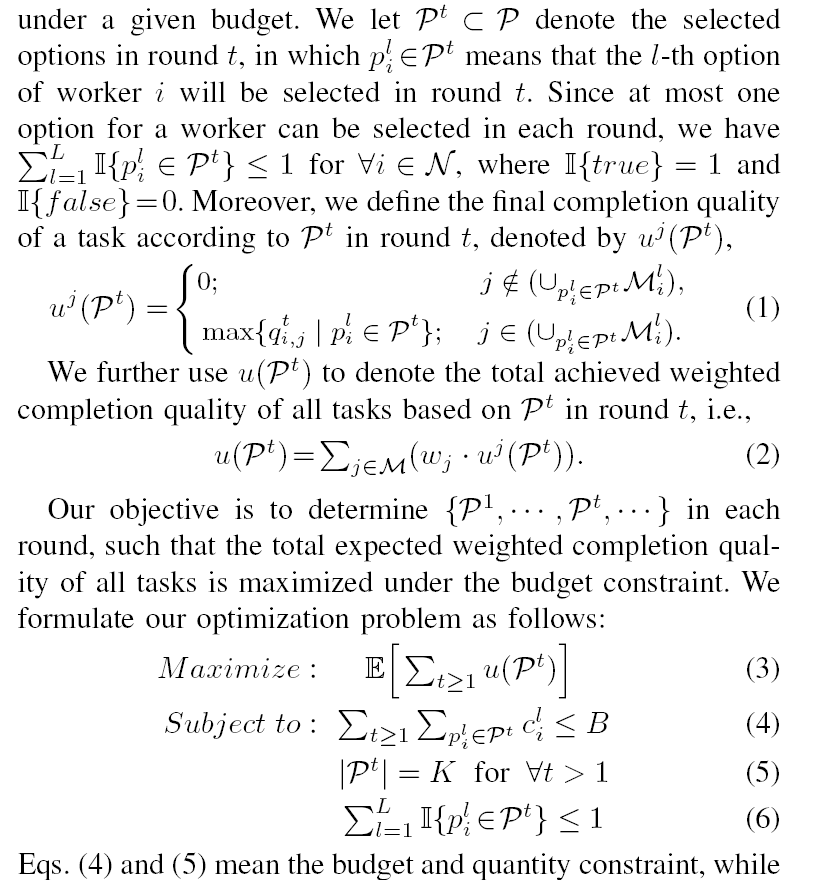
As we know, Upper Confidence Bound (UCB) is a widely used arm-pulling strategy, designed for the traditional multi armed bandit problem [18], [20]. It always selects the arm that has the largest value on the estimated reward and the upper bound of confidence to be pulled. To solve our CMAB problem, we extend the UCB strategy by adding two extra designs. First, when estimating the reward and computing the confidence for each arm, we consider that workers’ sensing qualities might be learned multiple times in one round, due to the reason that each worker has multiple options and covers multiple tasks. Second, we adopt the greedy strategy to solve the budget-limited maximum weighted coverage problem, when determining which arms should be pulled. Next, according to the extended UCB arm-pulling strategy, we design an unknown worker recruitment algorithm. In addition, we extend our problem to the scenario where workers’ costs are also unknown and devise another algorithm.

First, the requester publicizes these tasks to all workers via the platform. Then, each worker will tell the platform which sensing tasks it is willing to perform. Moreover, the worker can provide multiple options, each of which includes a subset of tasks that it can deal with and also attaches a cost that it wants to charge. Next, the platform will recruit some workers to perform the tasks round by round according to some strategy, until the budget is exhausted. Fig. 2 illustrates the main procedures.

For generality, we assume that the MCS system is heterogeneous, where each task can be completed by multiple workers and each worker can also cover multiple tasks in each round. Moreover, each worker has a sensing quality when performing tasks. The quality value can be evaluated only by the platform after the worker completes some tasks and submits the sensed results. If a task is completed by more than one workers, we will only select the best sensing data and let the completion quality of this task be the maximum sensing quality of these workers(他这论文假设选择前不知道, 选择后知道工人的数据质量, 然后选取质量最高的数据质量为此任务的完成质量). It should be pointed out that we mainly focus on the unknown worker recruitment problem in this paper, and thus we assume that the whole system is secure and truthful by leaving the privacy-preserving and incentive issues to be solved in future works. We assume that workers’ sensing qualities follow some unknown distributions. The platform can learn and estimate these distributions after the workers complete some tasks. A profile is used to record the learned quality for each worker (假设工人提交数据的质量服从一定的分布, 能够通过学习得到).

在一个网络中有个workers, 个tasks. 每个workers可以提交一个完成任务的申请would submit candidate options to the platform. We use to denote the th () option submitted by the worker . 是worker 第个选择项, 他愿意做的任务集合, 以及这个任务集合对应的价格集合. Platform选取worker 只能选取中的某1种option. , 是任务的数量, 是一个与任务数量相关的正函数, 任务数量多, 其函数的值就大. 用 denote the set of options submitted by worker for simplicity, and further use to denote the set of all options.

to denote the sensing quality of the worker completing the task in the -th round. 他们认为 follow an unknown independent and identically distribution with an unknown (unique) expectation . 如果选中了workers 的第 个option (i.e., ) is selected in round , must perform all tasks in . 因而, 就会学习次. 问题如下



公式1是说任务的质量是所有完成这个任务中质量最大的工作的那个质量,如果没有工作执行这个任务质量就为0. 公式2是所有任务的加权和. 公式3是一段时间内的总质量最大化, 公式4是预算不能超,这样如果每次花费少,则做的round会多, 反之则少. 公式5是每round选取k个. 每个工人每round只能选取其中的一个option.



上面计算被选取的次数



公式9计算学习得到的质量

公式10是Upper Confidence Bound (UCB) 函数它是Multi-Armed Bandit关键的一个概率计算公式



我们怎么做呢?

用我们的真象发现方法来做

假设我们的模型与他们的一样. 只是我们在workers提交数据后, 仍然不能确定workers的quality. 因而, 我们要用真象发现的方法来做.

同样, 有下面的模型

在一个网络中有个workers, 个tasks. 每个workers可以提交一个完成任务的申请would submit candidate options to the platform. We use to denote the th () option submitted by the worker . 是worker 第个选择项, 他愿意做的任务集合, 以及这个任务集合对应的价格集合. Platform选取worker 只能选取中的某1种option. , 是任务的数量, 是一个与任务数量相关的正函数, 任务数量多, 其函数的值就大. 用 denote the set of options submitted by worker for simplicity, and further use to denote the set of all options.

to denote the sensing quality of the worker completing the task in the -th round. 他们认为 follow an unknown independent and identically distribution with an unknown (unique) expectation . 如果选中了workers 的第 个option (i.e., ) is selected in round , must perform all tasks in . 因而, 就会学习次. 问题是, 我们如何选取workers使得系统的收益最大化呢?

我们把公式(11)加一个检验的收益.