市场设计中的定价数据引擎

陈 愉[1] 邓小铁[2] 王凯南[2] 阎 翔[1] 张云聪[1]

摘 要 本文中，我们针对互联网市场环境下分配和定价的问题，设计了一种基于数据引擎进行理论分析和实际应用的一般框架。其核心为在网格分布下建立一个简化型（reduced form）的数据库。基于以下事实，在适当的条件下，买方估值分布的适当偏差不会明显地改变简化型，进一步的，对于估值分布在一个很大范围内的买方用户群体，这样的数据库可以为市场定价者提供近似最优拍卖的可行解。与此同时，我们提出了对于这种模式的可行性研究所需要的理论原则、实现结构和实验验证方案。

关键词 简化型, 近似计算, 最优拍卖

引用格式 陈愉，邓小铁，王凯南，阎翔，张云聪：市场设计中的定价数据引擎. 第三届CCF大数据学术会议

DOI 10.16383/j.aas.20xx.cxxxxxx

Price Data Engine for Market Design

CHEN Yu[1] DENG Xiao-Tie[2] WANG Kai-Nan[2] YAN Xiang[1] ZANG Yun-Cong[1]

Abstract In this work, we develop a thorough investigation of a general framework for theoretical analysis and practical implementation of a data engine for allocation and pricing in Internet market design. We build the reduced form database for a grid of distributions. Based on the fact that, under suitable conditions, a fair deviations of buyer valuation distributions would not change the reduced form significantly, such that a database provides answers for a significantly broad class of value distributions of agents in a market setting.

We present the theoretical principles and implementation structures as well as experimental studies for a feasibility study of this paradigm.

Key words Reduced form, approximation, optimal auction

Citation Chen Yu, Deng Xiaotie, Wang Kainan, Yanxiang, Zhang Yuncong. Price Data Engine for Market Design *CCF BigData*, 2015

市场设计已经成为了由互联网技术引导的物质与服务传递方式革命中的一个重要课题。以搜索引擎为背景，人们已经完成了第一波针对付费搜索市场定价机制的研究, 具体来说就是设计了广义第二价格拍卖机制（GSP）. 但是在更多更新的市场中，GSP 机制并不适用。

基于Meyerson的工作，我们从其应用到不同互联网广告网站这一角度考虑，探讨了最优拍卖设计的一般方法。在拥有大量的显示设备的条件下，我们的定价机制可以实现为每一个互联网广告网站单独设计。这一一般方法将对任一用户有所帮助：他们可以通过在一个我们预先建立且持续更新的数据库中查询，得到最优的分配和定价方案。

这种方法的主要思想基于这样一个事实，即最优的机制设计可以被降维。在一个具有诚实性机制的市场中，分配和价格的可行解是所有可能的解的一个有限子集，这一子集被成为简化型（reduced form）。尽管枚举所有可行解仍然是一个艰巨的任务，但通过查询的方式得到其中一个解则变得简单很多。随着互联网广告应用，最优拍卖设计这一过程可以通过建立关于价格和分配的数据库完成，对应的获取其参数的问题就可以转化为对数据库的一个查询。这种方法虽然繁琐，而且需要一个巨大的预处理时间，但它得到的是一个一劳永逸的解决方案，将市场设计问题，简化成为数据工程问题。它为每一位互联网市场的设计者提供了计算一次即可永久使用的（电子）字典，用于查询适合自己拍卖网站的最佳拍卖协议。

粗略地说，在一次拍卖中，所有n位买家所给出的报价类型，决定了他们各自赢得该物品的概率。与之对应的是n个，以n位买家的报价类型，(和他们报价类型的概率分布规则，我们假设其固定不变）为自变量的函数。所谓的简化型，就是那些代表赢得物品概率的函数，它评估了所有其他买家对被拍卖物品估价的联合概率分布。

给定一个拍卖，它的简化型原则上可以通过其定义得到。相反，从一个给定的简化形式，即，一组n个单一变量函数，是否存在一个拍卖与之对应，

收稿日期 XXXX-XX-XX 录用日期 XXXX-XX-XX
Manuscript received Month Date, Year; accepted Month Date, Year

1. 致远学院上海交通大学 上海 200240 2. 计算机科学与工程系上海交通大学 上海 200240
1. Zhiyuan College, Shanghai Jiaotong University, Shanghai 200240 2. Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200240

则是实现过程中的一个重要课题。这一反演过程大部分由Border利用Farkas得到的引理 [1]彻底解决，得到一个性质刻画定理和一个求解的算法。与此同时，对于变量在离散域中取值的问题，相关的解的结构在 [4]一文也已经得到。另外，在 [3]提到，利用VCG求解器，在一般限制条件下的近似得具有诚实性机制的市场，它得到的近似最优收益可以被进步一加强。它的主要思路是将一个任意的机制分解为凸多面体边界上端点的凸组合。

由此推测，人们可以将预处理的可行解集存储在一个数据库当中，而在线市场通过查询操作来得到它的市场分配和定价方案。自然，会有人质疑我们提出的这个方案可用性，谁需要通过这么庞大的数据库，来从一个精确设计的拍卖中获益？

至少我们可以考虑这样一个例子：在线广告的互联网市场，需要考虑市场处于各种不同状态下的情况。每个拥有广告位的网站，可能会吸引到一组完全不同类型的访问者，他们中有的人只关注网站的其他部分，而有的人却是广告商潜在的买家。另一方面，在每一时刻，这样待售的网络广告位可能就有数十亿。这时，并非对于所有广告位的买家和卖家来说，计算最优的匹配方案以及获益最大的定价方式都是一件很容易的工作。

本文中，我们对这一目标进行了初步的分析。我们在一节中针对单一物品最优拍卖为例，介绍了必要的概念、符号和分析框架。接下来，在第二节中，通过一系列离散拍卖将多物品连续拍卖进行离散化的方法，给出了这一离散拍卖列逼近及收敛到连续拍卖的证明。从而建立了离散分析对问题解决的有效性。证明了两个接近的概率分布对应的简化型也是接近的，这使得那些已经预处理好的简化型可以用来处理其他具有相关条件的情况。在第三节中我们进一步讨论了一些相关的应用，给我们的方法一个清晰的应用前景。由于问题本身的复杂性，我们在第四节分析离线实现这一过程的方法论。这样，与传统经济学解决方案的不同，就完全体现了出来。古典解决方法注重在线的运作来发现价格，而定价引擎的方法，在离线状态下利用简化型的优点，将解决方案离线求解，存于数据库，用于在线查询。在第五节，我们利用简化型能够被离散逼近的性质，探讨降低离散数据库的数量，进一步简化实施方案的难度。最后，我们探讨目前的困难和将来的扩展空间。

1 单物品拍卖中的符号与概念框架

在这里，我们用单物品拍卖作为一个例子来说明我们的模型以及符号、概念和基本结果。

1.1 诚实性拍卖设计

考虑在一个拍卖中，n位买家，$1, 2, \ldots, n$对一件物品竞价。卖方的目标为在诚实性和个体理性的限制下，最大化它的期望收益。这里称一个机制具有诚实性是指，在这一机制下，没有任何买家可以通过谎报他/她对物品的私人估价而增加其期望效用。而一个机制具有个体理性则要求保证每一位买家都不能获得负的效用（即对于风险中性的买家来说负的期望效用）。

这时一个拍卖可以被描述为一组函数(Q, p)，其中Q将竞价向量$\vec{v}$映射到$[0, 1]^n$空间中的向量，表示在这一竞价方式下，各位买家获得物品的概率；而p将竞价向量$\vec{v}$映射到$\mathbb{R}^n$，表示各位买家需支付的钱。

同时，我们用$\vec{t}$表示非模糊性的出价向量。这里买家i对于拍卖物品的私人估价记作t_i，这是只有买家i知道的信息。令向量$\vec{v} = (v_1 \ldots v_n)$表示所有买家的出价。当所有买家以其真是估价出价时，$v = t$。

1.1.1 简化型

对于连续值拍卖，我们考虑买家i对物品的估价t_i由一组有界区间$[a_i, b_i]$上的连续值概率分布表示。令$f_i : [a_i, b_i] \to \mathbb{R}_+$相应的概率密度函数。这里我们假设$-\infty < a_i < b_i < +\infty$以及$\forall i \forall t_i : f_i(t_i) > 0$。令$F_i$表示$f_i$对应的概率分布函数：

$$F_i(t_i) = \int_{a_i}^{t_i} f_i(s_i) ds_i \tag{1}$$

令T表示所有买家可能的估价组合构成的集合：

$$T = [a_1, b_1] \times \cdots \times [a_n, b_n] \tag{2}$$

这里我们定义$T_{-i} = \prod_{j \neq i} [a_j, b_j]$，表示除去第$i$位买家外所有买家可能的估价组合

类似的，T和T_{-i}的联合概率密度函数分别由

$$f(t) = \prod_{j=1}^{n} f_j(t_j) \tag{3}$$

和

$$f_{-i}(t_{-i}) = \prod_{j \neq i} f_j(t_j) \tag{4}$$

表示。

给定一个拍卖机制(Q,p)，我们定义的简化型$q_i(t_i)$，表示买家i在估价为t_i时获得物品的概率。这里的概率来自于其他买家估价的不确定性，以及机制的随机性。在连续的情况下，我们有

$$q_i(t_i) = \int_{T_{-i}} Q_i(t) f_{-i}(t_{-i}) dt_{-i} \tag{5}$$

1.1.2 诚实性和个体理性

δ-贝叶斯激励相容(δ-BIC) 一个拍卖协议是贝叶斯激励相容（BIC）当且仅当每个参与者的效用在他/她对每个对象的出价为真实估价时达到最优。在以下的意义上，我们放松了一些$\delta > 0$的限制

$$\vec{q}_i(t_i)\cdot\vec{v}_i(t_i)-p_i(t_i) \geq \vec{q}_i(s_i)\cdot\vec{v}_i(t_i)-p_i(s_i)-\delta \ \ \forall t_i, s_i$$

如果上式中$\delta = 0$，对应的拍卖满足IC。

个体理性(IR) 一个机制满足个人理性（IR）当且仅当每一个以真实估价出价的买家效用是非负的，即

$$\vec{q}_i(t_i) \cdot \vec{v}_i(t_i) - p_i(t_i) \geq 0 \ \forall t_i$$

如果在一个机制下，每一个以真实估价出价的买家基于其他买家出价类型的期望效用是非负的，则这一机制满足临时个体理想（interim IR）。

1.1.3 收益最大化

z卖家的收益R被定义为所有买家付款之和的期望，即

$$R = \int_T \sum_{i\in N}(p_i(t_i)f_i(t_i))dt$$

这里收益最大化是指卖家在满足IC和IR条件下，最大化它的收益R。

我们将分别在第一两小节列出对于估价在连续和离散的区域上取值时前人已有的结果。然后，我们继续讨论我们关于近似方面的研究。

1.2 连续条件下的最优机制

对于每一个买家i和他的出价t_i，定义虚出价：

$$c_i(t_i) = t_i - \frac{1-F_i(t_i)}{f_i(t_i)} \tag{6}$$

令

$$P_i(q) = \int_{a_i}^{b_i} c_i(t_i)q_i(t_i)f_i(t_i)dt_i \tag{7}$$

以及机制Q的期望收益为：

$$R(Q) = \int_T \left(\sum_{i=1}^{n} c_i(t_i)Q_i(t)\right) f(t)dt = \sum_{i=1}^{n} P_i \tag{8}$$

最优机制是可以通过这样的方式得到：首先选择使$R(Q)$最大同时对于任意i都有$Q_i(t_i)$非递减的Q，然后选择适当的p使得这一机制具有诚实性和个体理性。

在最大化$R(Q)$时，可以令$h_i(s) = c_i(F_i^{-1}(s))$, $H_i(s) = \int_0^s h_i(r)dr$, $L_i(s) = \texttt{conv}\ H_i(s)$, $\bar{c}_i(t_i) = L_i'(F_i(t_i))$，其中$\texttt{conv}\ H_i(s)$为$[0,1]$上满足$L_i(s) \leq H_i(s), \forall s \in [0,1]$的最高的凸函数。可以证明，通过下面的方式选择$Q$可以使得$R(Q)$达到最大：对任意$t = (t_1, \ldots, t_n)$，计算每个$t_i$对应的$\bar{c}_i(t_i)$，同时令$M$表示具有最大$\bar{c}_i$的买家集合。对于$i \in M$，令$Q_i(t) = 1/|M|$，对于$j \neq M$，令$Q_j(t) = 0$ for $j \neq M$。

1.3 离散模型

类似的方法也适用于在离散条件下 [4]构建最优拍卖，其中拍卖$A = (n, W, F)$定义如下：

买家i的估价是从一个有限集$W_i = \{w_i^k \mid k = 1, \ldots, K\}$中得到，而非连续模型中的区间。为了简便我们可以假设$w_i^1 < \cdots < w_i^K$，这一点可以通过置换指标实现。记投买家i估价的概率分布为$f_i^1, \ldots, f_i^K$。令$F_i^k = \sum_{j=1}^{k} f_i^j$.

同时，令$W = W_1 \times \cdots \times W_n, W_{-i} = \prod_{j\neq i} W_j$，而$Q_i(w_1^{k_1}, \ldots, w_n^{k_n})$表示买家$i$在出价情况为$w = (w_1^{k_1}, \ldots, w_n^{k_n})$时获得物品的概率。则此时可定义简化型：

$$q_i^k = \sum_{k_1=1}^{K} \cdot \sum_{k_{i-1}=1}^{K} \sum_{k_{i+1}=1}^{K} \cdot \sum_{k_n=1}^{K} Q_i(w_1^{k_1}, \cdot, w_i^k, \cdot, w_n^{k_n}) \\ f_1^{k_1} \cdot f_{i-1}^{k_{i-1}} f_{i+1}^{k_{i+1}} \cdot f_n^{k_n} \tag{9}$$

表示出价w_i^k获得物品的概率。

另外，这时买家i的虚出价可被定义为：

$$c_i^k = w_i^k - (w_i^{k+1} - w_i^k)\frac{1-F_i^k}{f_i^k} \tag{10}$$

再令

$$P_i = \sum_{k=1}^{K} c_i^k q_i^k f_i^k \tag{11}$$

以及机制Q的期望收益为：

$$\begin{aligned} R_Q &= \sum_{(k_1,\ldots,k_n)\in W} \left(\sum_{i=1}^{n} c_i^{k_i} Q_i(w_1^{k_1}, \ldots, w_n^{k_n})\right) \prod_{i=1}^{n} f_i^{k_i} \\ &= \sum_{i=1}^{n} P_i \end{aligned} \tag{12}$$

这里我们需要选择在满足条件$q_i^k \leq q_i^{k+1}, \forall i$下使得$R_Q$最大的$Q$。

这里我们注意到，$F_i^k = f_i^1 + \cdots + f_i^k$。则令$H_i^k = c_i^1 f_i^1 + \cdots + c_i^k f_i^k$，以及$F_i^0 = H_i^0 = 0$，我们可以用$H_i(z) : [0,1] \to \mathbb{R}$表示将$\{(F_i^0, h_i^0), \ldots, (F_i^K, H_i^K)\}$逐个连接时产生的函数，同时令$L_i(z) : [0,1] \to \mathbb{R}$表示集合$\{(F_i^0, H_i^0), \ldots, (F_i^K, H_i^K)\}$的下包络。

将$L_i(F_i^k)$简记为L_i^k.令$\bar{c}_i^k = (L_i^k - L_i^{k-1})/f_i^k$。则通过下面的方式选择$Q$可以使得$R(Q)$达到最大：令$M$为具有最大$\bar{c}_i$的买家集合,对于$i \in M$，令$Q_i(t) = 1/|M|$，对于$j \neq M$，令$Q_j(t) = 0$ for $j \neq M$。

1.4 简化型形式下的最优收益

在连续条件下，一组函数$q_i : T_i \to R$被成为可实现的，当且仅当存在一个与之对应的的有效分配函数。也就是说，存在一组函数Q_i满足：

$$q_i(t_i) = \int_{T_{-i}} Q_i(t_i, t_{-i}) dF_{-i} \tag{13}$$

类似的，在离散条件下，一组函数$q_i : T_i \to R$被成为可实现的，当且仅当存在一个有效拍卖的分配函数与之对应。即存在一组函数Q_i满足：

$$q_i^k = \sum_{k_1=1}^{K} \cdot \sum_{k_{i-1}=1}^{K} \sum_{k_{i+1}=1}^{K} \cdot \sum_{k_n=1}^{K} Q_i(w_1^{k_1}, \cdot, w_i^k, \cdot, w_n^{k_n})$$
$$f_1^{k_1} \cdot f_{i-1}^{k_{i-1}} f_{i+1}^{k_{i+1}} \cdot f_n^{k_n} \tag{14}$$

如果我们从简化型的角度来考虑上面讨论的两种情况，则我们的目标变为找到一个可实现的非递减的简化型q使得

$$R(q) = \sum_{i=1}^{n} P_i(q) \tag{15}$$

最大。

下面我们考虑一个可实现的简化型应满足的条件。在[1]文中，Border证明了，在离散条件下，简化型q是可实现的当且仅当

$$\begin{aligned}
&\forall S_1 \subseteq W_1, S_2 \subseteq W_2, \ldots, S_n \subseteq W_n,\\
&\sum_i \sum_{A_i \in S_i} q_i(A_i) Pr(t_i = A_i)\\
&\leq 1 - \prod_i (1 - Pr(t_i \in S_i))
\end{aligned} \tag{16}$$

这一公式描述的是赢得物品的买家在S_i中（如果买家i赢得物品）的概率，不超过确实存在一个买家i的类型在S_i中的概率。注意到，这个公式的必要性显然成立。即使是在连续条件下也是如此，即:

$$\begin{aligned}
&\forall S_1 \subseteq T_1, S_2 \subseteq T_2, \ldots, S_n \subseteq T_n,\\
&\sum_i \int_{A \in S_i} q_i(A_i) Pr(t_i = A_i)\\
&\leq 1 - \prod_i (1 - Pr(t_i \in S_i))
\end{aligned} \tag{17}$$

1.5 连续机制的离散化

接着，我们讨论连续情况和离散情况之间的关系。我们讨论一套系统化的方法将一个给定的连续拍卖问题离散化。同时应用上述的简化型的性质，得到一个连续拍卖以及其对应的离散拍卖的最大期望收益用于比较。

给定一个连续拍卖$A = (n, T, F)$和一个整数K，我们定义两个离散拍卖：$\check{A}_K = (n, \check{W}, \check{F})$和$\hat{A}_K = (n, \hat{W}, \hat{F})$如下:
对于每个买家i，我们把区间$[a_i, b_i]$ 划分为K个子区间。有两种方法可以得到离散化的拍卖。对于连续拍卖中的每个估价t_i，假设它落在第k个子区间，则我们令买家i在离散拍卖中的估价等于这一子区间的左边界，记作$\check{w}_i^k$。于是买家i估价为$\check{w}_i^k$的概率$\check{f}_i^k$等于f在第k个子区间的积分。如果我们采取右边界的话，可以得到另一个离散结果。这两种离散结果我们都将在这里讨论。令

$$\begin{aligned}
\check{W}_i &= \{\check{w}_i^k \mid \check{w}_i^k\text{是第}k\text{个子区间的左边界}\}\\
\hat{W}_i &= \{\hat{w}_i^k \mid \hat{w}_i^k\text{是第}k\text{个子区间的右边界}\}
\end{aligned}$$

以及对应的

$$\check{f}_i^k = \int_{\check{w}_i^k}^{\check{w}_i^{k+1}} f(t_i) dt_i \tag{18}$$

$$\hat{f}_i^k = \int_{\hat{w}_i^{k-1}}^{\hat{w}_i^k} f(t_i) dt_i \tag{19}$$

其中$\check{w}^{K+1}$记为b_i,同时$\hat{w}^0$记为a_i。易得对于任意的i和k，我们有$\check{f}_i^k = \hat{f}_i^k$。

另外，我们定义$d_i = (b_i - a_i)/K$以及$d = \max\{d_i\}$。

其它的概念和符号与之前的定义相同。

1.5.1 离散化拍卖的收益与原拍卖收益之间的关系

首先我们给出两个引理，其中引理一我们将在后面的章节中证明：

引理**1.1.** 对于任意连续拍卖A, $R(\check{A}_k) \leq R(A) \leq R(\hat{A}_k)$.

引理**1.2.** 对于任意连续拍卖A, $R(\hat{A}_k) \leq R(\check{A}_k) + d$

证明. 由$\check{c}$以及$\hat{c}$的定义可知：

$$\hat{c}_i{}^k = \check{c}_i{}^k + d_i \tag{20}$$

假设$\check{A}_K$和$\hat{A}_k$的最优简化型分别是$\check{q}$和$\hat{q}$。对于$\hat{P}_i(\hat{q}_i)$和$\check{P}_i(\hat{q}_i)$我们可得

$$\begin{aligned}
\hat{P}_i(\hat{q}) - \check{P}_i(\hat{q}) &= \sum_{k=1}^{K} d_i \hat{q}_i^k \hat{f}_i^k\\
&= d_i P(\text{买家i 得到物品})
\end{aligned} \tag{21}$$

因此

$$R(\hat{A_K}) - R(\check{A_K}) \le \hat{R}(\hat{q}) - \check{R}(\hat{q}) \le \sum_{i=1}^{n} dP(\text{买家i得到物品}) \le d \quad (22)$$

□

由上述两个引理，我们可直接得到：

定理1.1. 对于任意拍卖$A = (n, T, F)$,我们有$\lim\limits_{K\to\infty} R(\check{A_K}) = R(A)$

推论1.1. 对于任意拍卖$A = (n, T, F)$ 以及任意$\epsilon > 0$, 存在K 满足$K = O(\max\limits_i\{b_i - a_i\}, \frac{1}{\epsilon})$及$R(A) - R(\check{A_K}) < \epsilon$

证明. 由引理 1.2可知, 要满足$R(A) - R(\check{A}_k) < \epsilon$,我们只需要$d < \epsilon$, 因此我们需要$\frac{\max_i\{b_i - a_i\}}{K} < \epsilon$，也就是说$K > \frac{1}{\epsilon}\max\limits_i\{b_i - a_i\}$ □

2 连续多物品拍卖的离散近似

目前，在连续值条件下的单物品拍卖已经有了完整的解决方案 [5]，与此同时，在离散值条件下的多物品拍卖可以通过近似方法求解 [3]，而我们感兴趣的是在连续值条件下的多物品拍卖。

2.1 多物品拍卖

一个多物品拍卖$A = (n, m, T, v, F)$定义如下：

买家和物品 有n位买家和m件物品.

类型 买家i有一个类型集T_i (例如，$T_i = [0, 1]$). 买家的类型是T_i中的一个随机变量，它的概率密度函数为f_i，概率分布函数为F_i。令$T = \Pi_{i=1}^n T_i$。T中的每一个向量都是一组所有买家可能的类型。

估价 每位买家有一个估价函数$\vec{v_i} : T_i \to \prod_{1\le j\le m}[a_{ij}, b_{ij}]$，其中$a_{ij}$ 和b_{ij} 分别是买家i对物品j估价可能的最小和最大值。也就是说$v_{ij}(t_i)$是买家i在类型为t_i时对物品j的估价。

拍卖 一个拍卖可以由一组函数(Q, p)表示。其中$Q : T \to R^{mn}$，$p_i : T_i \to R$。$Q_{ij}(t)$表示买家i在类型为t时获得物品j的概率。令$p_i(t_i, t_{-i})$表示买家i类型为t_i而其他买家类型为t_{-i}时的付款。令$p_i(t_i)$表示买家i付款的期望:

$$p_i(t_i) = \int_{T_{-i}} p_i(t_i, t_{-i})dF_{-i}$$

简化型 一个拍卖中买家i的简化型$\vec{q_i} : T_i \to R^m$是买家i在类型为t_i时获得每件物品的期望概率向量，即：

$$q_{ij}(t_i) = \int_{T_{-i}} Q_{ij}(t_i, t_{-i})dF_{-i}$$

实现 一组函数$\vec{q_i} : T_i \to R^m$ $(1 \le i \le n)$是可实现的当且仅当存在一组函数$Q_{ij}(\cdot)(1 \le i \le n$ and $1 \le j \le m)$ 满足$\forall i \forall j$

$$q_{ij}(t_i) = \int_{T_{-i}} Q_{ij}(t_i, t_{-i})dF_{-i}$$

其中

$$0 \le Q_{ij}(t) \le 1 \quad \forall i, \forall j \forall t$$

$$\sum_{i\in N} Q_{ij}(t) \le 1 \quad \forall j \forall t$$

效用 买家i的效用函数U_i为他/她获得物品的期望估价减去其付款。因此买家i的期望效用为：

$$U_i(p, Q_{ij}(t_i, t_{-i}), t_i) = E_{t_{-i}}(\sum_{j\in M} Q_{ij}(t)v_{ij}(t_i) - p_i(t_i, t_{-i}))$$

当t_{-i}对应于F_{-i}时，简化型的效用可以简化为：

$$U_i(p, \vec{q_i}(t_i), t_i) = \vec{q_i}(t_i) \cdot \vec{v_i}(t_i) - p_i(t_i)$$

2.2 有限支撑集上拍卖的近似能力

首先我们给出三个引理：

引理2.1. 假设有两个拍卖$A_1 = (n, m, T, v, F)$,$A_2 = (n, m, T, u, F)$。

如果一组可实现的简化型和付款(q, p)满足A_1的δ-*BIC*和*IR*条件，令$p'_i(\vec{t_i}) = p_i(\vec{t_i}) + \vec{q_i}(\vec{t_i}) \cdot (\vec{u_i}(\vec{t_i}) - \vec{v_i}(\vec{t_i}))$, 则$(q,p')$满足$A_2$的$(\delta + 2m|u - v|_\infty)$-*BIC*和*IR*条件。

引理2.2. 对于一个可实现的简化型q和一个常数$c \in (0, 1)$, qc仍然是可实现的。

证明. 假设分配函数Q为q的一个实现,由于$c \in (0, 1)$，易得Qc也是一个可行的分配函数，因此Qc是qc的一个实现. □

引理2.3. 对于一个拍卖$A = (n, m, T, v, F)$, 如果每个买家对每个物品的估价增加一个常向量$\vec{d}$, 那么最大收益将会增加，但不超过$|\vec{d}|_1$。

结合这三个引理我们可得以下定理：

定理2.1. 对于一个拍卖$A = (n, m, T, v, F)$和一个常数d,定义$A^d = (n, m, T, v_d, F)$，其中$v^d_{ij}(\vec{t_i}) = d\lfloor v^d ij(\vec{t_i})/d\rfloor$, 则$\lim_{d\to 0} R(A^d) = R(A)$。

证明. 假设(q,p)是一组使A^d收益最大的简化型和付款。由引理 1.1可知，(q,p) 满足A的$(\delta+dm)$-BIC 和IR 条件。易得$(\frac{\delta}{\delta+2dm}q, \frac{\delta}{\delta+2dm}p')$也满足$A$的$\delta$-BIC和IR条件(其中$p'$定义与引理 2.1中类似)。因此我们有$R(A^d) \leq \frac{\delta+2dm}{\delta}R(A)$。令$A'^d = (n,m,T,v'_d,F)$，其中$v'^d_{ij}(\vec{t_i}) = v^d_{ij}(\vec{t_i}) + d$。由上面的证明我们可以发现$R(A) \leq \frac{\delta+2dm}{\delta}R(A'^d)$。又由引理 2.1我们知道$R(A'^d) \leq R(A^d) + dm$。故$\frac{\delta}{\delta+2dm}R(A^d) \leq R(A) \leq \frac{\delta+2dm}{\delta}(R(A^d)+dm)$。证毕。 □

2.3 关于以上引理的讨论

在线广告定价问题中，广告主在每个市场中因为各种因素，其价值会有不同的分布。互联网上每一个网页位置广告是一个不同的在线广告市场。甚至它在每一个ip地址的展示，每一个移动设备上不同信号激发的点击，都可以成为不同市场的定价问题。这样大量市场的最优定价将依赖于以上的计算模型进行定价。

然而，同样的广告在不同市场上价值差异在不同区域类似环境中有时可以在同一广告发布系统中表示为广告商类型加上一个相同的向量的情况。以上引理就可以帮助我们简单地得到这些市场的最优定价之间的关系。

以此类推，当每个买家的类型乘以一个常数时，我们只需要通过改变付款来满足新拍卖的BIC和IR条件。考虑在不同区域中这些广告的潜在客源，不同地区的客户有不同的购买力，因此，广告商对于那些在发达地区里可以通过WiFi访问的页面估价就比较高，而对于那些在欠发展地区可以通过WiFi访问的页面估价就比较低。我们还应该考虑到不同页面的页面视图。广告商对于热门网页上的广告位估价更高。在这种情况下，我们可以定义一个变量来表示在页面视图上不同区域的购买力，这相当于每个买家的类型乘以一个相同的常数的情况。

两种方法的结合使用，可以大大降低我们应用这一问题的计算要求。

3 应用范例:为多物品价值独立型市场建立定价数据库

下面我们考虑为这样的市场建立数据库：市场中买家对所有物品的估价都是相互独立的。假设买家对物品的估价服从截断正态分布$N_t(\mu_{ij},\sigma_{ij}^2,\mu_{ij}-2\sigma_{ij},\mu_{ij}+2\sigma_{ij})$,即一个期望为$\mu_{ij}$方差为$\sigma^2$的正态分布，在$\mu_{ij}-2\sigma$ and $\mu_{ij}+2\sigma$处截断。其概率密度函数为

$$f_{ij} = \frac{\frac{1}{\sigma_{ij}}\varphi\left(\frac{x-\mu_{ij}}{\sigma_{ij}}\right)}{\Phi(2)-\Phi(-2)} \tag{23}$$

对于每一组μ_{ij}和σ_{ij},我们需要高效地计算其对应的简化型和每位买家的付款。当一个拍卖需要进行许多次时，如果我们提前建立好关于简化型和每位买家的付款的数据库，那么每一次拍卖只需通过在数据库中查询即可完成，这将使得整个工作变得非常高效。由于μ_{ij}和σ_{ij}的取值是连续的，因此我们不可能建立数据库存储每一个μ_{ij}和σ_{ij}取值对应的计算结果。故我们考虑将它们的取值区间离散化为网格，然后只计算和存储μ_{ij}和σ_{ij}在网格结点上取值时的情况。然后对于每一个拍卖，我们寻找与它的参数最接近的结点，使用这一结点对应的计算结果作为它的计算结果，这一结果还需要做一些小的修正，我们会在后面讨论。本文中我们证明的结果可以保证这样的近似将是一个好的近似，其误差正比与网格的粒度。

3.1 数据库的构建

这里我们假设每个市场中买家的估价函数v_{ij}是线性的，即$v_{ij}(0)=a_{ij}$，$v_{ij}(1)=b_{ij}$，$v_{ij}(t)=a_{ij}+t(b_{ij}-a_{ij})$。同时买家的类型$t_{ij}$服从$[0,1]$上期望为$\mu=0.5$方差为$\sigma=0.25$的截断正态分布。在这一条件下，我们的模型可以精确的模拟一个确定的拍卖。

在建立好数据库之后，对于每个拍卖，我们需要寻找与其参数最接近的结点。这里我们记每个结点对应参数为μ_{ij}和σ_{ij},而这个拍卖的参数分别是$\tilde{\mu}_{ij}$和$\tilde{\sigma}_{ij}$。假设我们使用的网格尺寸为ε,那么我们能够找到的最近的结点与真实参数对应的结点的距离将小于ε。通过下面的方法，可以利用已经存储的结点得到这个拍卖的近似简化型：令$d=\max_{ij}(|\mu_{ij}-\tilde{\mu}_{ij}|)$,那么新的简化型为$\tilde{q}=\frac{\delta}{\delta+2dm}q$,同时付款为$\tilde{p}=\frac{\delta}{\delta+2dm}p'$,其中$p'$与引理 2.1中定义类似。由引理 2.1可知, $(\tilde{q},\tilde{p})$仍然满足δ-BIC和IR条件。根据定理 2.1,由$(\tilde{p},\tilde{q})$对应的期望收益与最大收益之间的差距上界为ε的常数倍。

3.2 数据库构建和查询

首先，网格中结点的个数N_g为$\frac{1}{\epsilon^{nm}}$，并且对于每个网格，我们要找到最优的简化型和付款函数。如附录所示，这部分工作需要进行一次双层线性规划,使用椭球法可以在多项式时间内完成对简化型的检验。内层检验需要对每个类型$t\in T$进行一次最大匹配。通常，最大匹配的求解需要$O(N^3)$的时间，其中N是网格结点的个数。所以，建立数据库总的时间

复杂度为

$$O(\frac{|T|}{\epsilon^{nm}}(n+m)^3 F(n,m,\{T_i\}))$$

其中，$F(n,m,\{T_i\})$检验的个数。

至于空间复杂度，我们需要临时储存线性规划中的不等式，总共将需要占据$O((m\sum_{i=1}^{n}|T_i|)^2)$的空间，具体取决于实现细节。

另外，对于每个网格结点，数据库中还需要存储它对应的简化型和付款函数，这将占据$O(\frac{m\sum_{i=1}^{n}|T_i|}{\epsilon^{nm}})$的磁盘空间。

而对于每一次查询，我们可以直接找到网格中最近的结点并加载对应的数据，这部分工作需要$O(m\sum_{i=1}^{n}|T_i|)$的时间和空间。

3.3 复杂性

我们在这一小节中计算在线计算一类拍卖的复杂度。详细的求法见附录，在这里我们作复杂性分析。首先假设有M个买家,每位买家平均有T种类型,而拍卖的物品有N个。方便起见，我们假设$X = M*T*N$。

假设我们使用椭球法求解线性规划问题。这个算法中，椭球迭代需要花费$O(D^2)$的时间找到解(D为线性规划问题的维度，在我们的问题中$D = (M*T*N+M*T)$)，而每一步迭代的计算需要$O(D^2)$的时间。因此这个算法的时间复杂度为$O(X^2 * max(X^2, Time(SO)))$。

接下来我们考虑分离预示(SO)，每运行一次SO,我们还需要求解一个维度为$M*N*T+1$的线性规划，因此$Time(SO) = O(X^2 * max(Time(VCG), X^2))$。

如果我们计算无限制的虚VCG,我们可以在$O(M)$时间内计算一个简化型(只是将每个买家类型比其他买家大的概率相乘),但如果考虑一般情况，我们需要枚举所有可能类型并计算社会福利(social welfare),例如，如果我们考虑每个买家最多买一个物品的拍卖，我们需要对每一种情况运行一次匹配算法。因此为了计算每个简化型，我们至少需要$O(T^M)$的时间来枚举每种情况。而总共有$M*T$个简化型。

综上,求解一个无限制的多物品拍卖，整个算法的时间复杂度为$O(X^4 * max(M^2 * T, X^2))$,而对于有限制情况下，则是$O(X^4 * M * T * T^M * Time(max\ social\ welfare))$

4 离线实现

在多物品拍卖的实际应用中，如互联网广告，我们通常要求，每个买家只能获得一个物品（因为两个相同的广告出现在同一页是没有意义的），而且所有物品必须都被卖出（存在没有被卖出的广告位既是浪费空间，也因为在页面留有空白空间而显得不美观）。由于互联网上的广告拍卖非常频繁，因此每次向数据库查询也会有比较大的代价，所以将一些计算离线完成是很自然的想法。我们发现几个实例中，买家类型的变化不会改变我们的问题的简化型，这使得通过计算分离预示来检验简化型实现的工作也可以离线完成。

4.1 线性变化

这一小节我们将说明如果广告商的分布出现线性变化，那么我们可以离线快速求解。

我们先解释一下我们问题中的线性规划：

令$p_i(\vec{v_i})$表示买家i在类型为$\vec{v_i}$时的期望付款。下面的线性规划计算所有满足δ-贝叶斯激励相容和个体理性条件的机制中的最优期望收益。

- 变量:
  - $p_i(\vec{v_i})$，表示买家i在类型为$\vec{v_i}$时的期望付款。
  - $\pi_{ij}(\vec{v_i})$，表示买家i在类型为$\vec{v_i}$时得到物品j的概率。
- 约束:
  - $\vec{\pi}_i(\vec{v_i}) \cdot \vec{v}_i - p_i(\vec{v_i}) \geq \vec{\pi}_i(\vec{w_i}) \cdot \vec{v}_i - p_i(\vec{w_i}) - \delta$.($\delta$-$BIC$ 的定义)
  - $\vec{\pi}_i(\vec{v_i}) \cdot \vec{v}_i - p_i(\vec{v_i}) \geq 0$.($IR$的定义)
  - $SO(\vec{\pi})$ = yes，保证简化型$\vec{\pi}$在$F(\mathcal{F}, \mathcal{D}')$中(可行性)。
- 最大化:
  - $\sum_{i=1}^{m}\sum_{\vec{v_i}\in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i})$，期望收益。

易知，当每个买家的类型乘以同一个常数时，我们只是令付款乘以相同的常数，则对应的简化型和新的付款函数仍然满足的约束条件和最大化期望收益。

接下来我们证明，如果每个买家的类型加上一个相同的向量，我们只需改变付款函数，则对应的简化型和新的付款函数仍然满足的约束条件和最大化期望收益。假设上述线性规划的解为$\vec{\pi}_i(\vec{v_i})$和$p_i(v_i)$。当所有买家的预算都增加$\vec{d}$个单位,也就是说，他们对第j个物品的估价提高d_j,我们希望通过原线性规划的解直接得到新的线性规划的解。记新线性规划的解为$\vec{\pi}'_i(\vec{v_i})$和$p'_i(\vec{v_i})$。令$\vec{\pi}'(\vec{v_i}+\vec{d}) = \vec{\pi}(\vec{v_i})$,即它们的简化型相同。同时$p'_i(\vec{v_i}+\vec{d}) = p_i(\vec{v_i}) + \vec{\pi}_i(\vec{v_i}) \cdot \vec{d}$，也就是说每个买家额外支付$d_j$乘以他获得每个物品$j$的概率。我们希望证明这事实上就是新线性规划

的解。
令$T_i' = \{\vec{t_i} + \vec{d} \mid \vec{t_i} \in T_i\}$。可以验证所有的约束条件仍然满足。

$$\vec{\pi}_i(\vec{v_i}) \cdot \vec{v_i} - p_i(\vec{v_i}) \geq \vec{\pi}_i(\vec{w_i}) \cdot \vec{v_i} - p_i(\vec{w_i}) - \delta$$
$$\forall \vec{v_i}, \vec{w_i} \in T_i$$
$$\vec{\pi}_i(\vec{v_i}) \cdot (\vec{v_i} + \vec{d}) - (p_i(\vec{v_i}) + \vec{\pi}_i(\vec{v_i}) \cdot \vec{d})$$
$$\geq \vec{\pi}_i(\vec{w_i}) \cdot (\vec{v_i} + \vec{d}) - (p_i(\vec{w_i}) + \vec{\pi}_i(\vec{w_i}) \cdot \vec{d}) - \delta$$
$$\vec{\pi}_i'(\vec{v_i}) \cdot \vec{v_i} - p_i'(\vec{v_i}) \geq \vec{\pi}_i'(\vec{w_i}) \cdot \vec{v_i} - p_i'(\vec{w_i}) - \delta$$
$$\forall v_i, w_i \in T_i'$$

第二个约束条件可以通过同样的方式证明，同时$SO(\vec{\pi})$的结果应为“yes”，因为$\vec{\pi}$类型的分布并未改变。
现在我们希望证明这个解是最优解。注意到如果我们将p_i改为p_i',函数只需要加上一个因子

$$\begin{aligned}
& \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i'(\vec{v_i}) \\
= & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}](p_i(\vec{v_i}) + \vec{\pi}_i(\vec{v_i}) \cdot \vec{d}) \\
= & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i}) + \\
& \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \vec{\pi}_i(\vec{v_i}) \cdot \vec{d} \\
= & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i}) + \\
& \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \sum_{j=1}^{n} \vec{\pi}_{ij}(\vec{v_i}) d_j \\
= & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i}) + \\
& \sum_{j=1}^{n} d_j \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \vec{\pi}_{ij}(\vec{v_i}) \\
= & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i}) + \sum_{j=1}^{n} d_j
\end{aligned}$$

所以，如果新的线性规划存在另一个解$(\tilde{\pi}_i', \tilde{p}_i')$，满足它的收益大于$(\pi_i', p')$,那么$(\tilde{\pi}_i, \tilde{p}_i)$的收益将大于$(\pi_i, p_i)$,这与「$(\pi_i, p_i)$是原线性规划问题的最大解矛盾。

4.2 类型轻微改变的情况

求解上述线性规划问题需要花费大量的时间，其时间复杂度是指数级的。我们希望确保如果一位买家改变了一点点他/她的类型，我们可以只对相应的数据做出更新，尽管更新过后的的收益不是最大值，但距离最大值非常靠近。

假设买家k的预算改变了$\vec{d}$个单位,也就是说他对第j个物品的估价改变d_j，而其他买家的出价不变。同时，我们假设买家k的改变上界为ϵ,即$|d_j| < \epsilon$，$j = 1, 2, \cdots, n$。

令$\vec{\pi}'(\vec{v_k} + \vec{d}) = \vec{\pi}(\vec{v_k})$，即简化型不变，同时$p_k'(\vec{v_k} + \vec{d}) = p_k(\vec{v_k}) + \vec{\pi}_k(\vec{v_k}) \cdot \vec{d}$，也就是说买家$i$额外支付$d_j$乘以他/她获得每个物品$j$的概率。我们希望证明：(i)它仍然满足约束条件(ii)对应的收益与真实的最大收益相差不大。

令$T_k' = \{\vec{t_k} + \vec{d} \mid \vec{t_k} \in T_k\}$。可以验证所有约束条件仍然满足。

$$\vec{\pi}_k(\vec{v_k}) \cdot \vec{v_k} - p_k(\vec{v_k}) \geq \vec{\pi}_k(\vec{w_k}) \cdot \vec{v_k} - p_k(\vec{w_k}) - \delta$$
$$\vec{\pi}_k(\vec{v_k}) \cdot (\vec{v_k} + \vec{d}) - (p_k(\vec{v_k}) + \vec{\pi}_k(\vec{v_k}) \cdot \vec{d})$$
$$\geq \vec{\pi}_k(\vec{w_k}) \cdot (\vec{v_k} + \vec{d}) - (p_k(\vec{w_k}) + \vec{\pi}_k(\vec{w_k}) \cdot \vec{d}) - \delta$$
$$\vec{\pi}_k'(\vec{v_k}) \cdot \vec{v_k} - p_k'(\vec{v_k}) \geq \vec{\pi}_k'(\vec{w_k}) \cdot \vec{v_k} - p_k'(\vec{w_k}) - \delta$$

第二个约束条件可以通过同样的方式证明，同时$SO(\vec{\pi})$的结果应为“yes”，因为$\vec{\pi}$类型的分布并未改变。

现在我们希望证明这个解与最优解很接近。

$$\begin{aligned}
& \max \Bigg\{ \sum_{i \neq k} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] p_i(\vec{v_i}) + \\
& \sum_{\vec{v_k} \in T_k'} Pr[t_k = \vec{v_k}] p_k'(\vec{v_k}) \Bigg\} \\
\leq & \max \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] p_i(\vec{v_i}) + \\
& \max \sum_{\vec{v_k} \in T_k} Pr[t_k = \vec{v_k}] \vec{\pi}_k(\vec{v_k}) \cdot \vec{d} \\
\leq & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i}) + \epsilon \sum_{\vec{v_k} \in T_k} Pr[t_k = \vec{v_k}] \\
= & \sum_{i=1}^{m} \sum_{\vec{v_i} \in T_i} Pr[t_i = \vec{v_i}] \cdot p_i(\vec{v_i}) + \epsilon
\end{aligned}$$

其中第二个不等式是基于假设没有人会获得超过一件物品。

由上述我们可以发现，新线性规划问题的最大值与原线性规划问题的最大值相差不超过ϵ。当然，新的最大值与更新原最大值之后的期望收益相差也不超过ϵ。尽管不能保证真实的最大值很接近我们得到的简化型和定价。

另一方面，当类型改变很小时，简化型可能有很大改变。例如，考虑只有一个物品和一个买家，买家有两种类型：10-ϵ和20,它们的概率都是0.5,那么对应的简化型为(0,1)。当他的类型变为10+ϵ和20时,简化型变为(1,1).

5 讨论与未来的工作

本文提出的方法考虑了建立数据库预处理在线定价与分配的市场组成。它为市场设计问题的求解开辟了新的可能性。如何才能进一步提高效率呢？我们把它作为未来研究的课题。特别地，一个关键的难点在于在预示构建中对于线性规划的处理。

5.0.1 椭球法的难点

椭球法是第一个用来求解线性规划问题的多项式时间算法，但它在实际应用中并未得到充分的体现。椭球法要求非常精确的计算，而当线性规划问题的维度为D时,它需要精确度达到D^2比特,这一需要在我们的问题中意味着更大计算量。

References

1 Border, Kim C.: Reduced form auctions revisited. Economic Theory 31.1 (2007): 167-181.

2 Che, Yeon-Koo, Jinwoo Kim, and Konrad Mierendorff.: Generalized Reduced-Form Auctions: A Network-Flow Approach. Econometrica 81.6 (2013): 2487-2520.

3 Cai, Yang, Constantinos Daskalakis, and S. Matthew Weinberg.: Optimal multi-dimensional mechanism design: Reducing revenue to welfare maximization. Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on. IEEE, 2012.

4 Elkind, Edith.: Designing and learning optimal finite support auctions. Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, 2007.

5 Myerson, Roger B.: Optimal auction design. Mathematics of operations research 6.1 (1981): 58-73.

附录: 引理 1.1证明

证明. 我们首先分析H和$\check{H}$之间的关系：

$$\begin{aligned}
\check{H}_i^k &= \sum_{j=1}^{k} \check{c}_i^j \check{f}_i^j \\
&= \sum_{j=1}^{k} \check{w}_i^j \check{f}_i^j - (\check{w}_i^{k+1} - w_i^1) + \sum_{j=1}^{k} (\check{w}_i^{j+1} - \check{w}_i^j) \check{F}_i^j \\
&= \sum_{j=1}^{k} \check{w}_i^j (\check{F}_i^j - \check{F}_i^{j-1}) - (\check{w}_i^{k+1} - \check{w}_i^1) \\
&\quad + \sum_{j=1}^{k} (\check{w}_i^{j+1} - \check{w}_i^j) \check{F}_i^j \\
&= \check{w}_i^{k+1} \check{F}_i^k - \check{w}_i^{k+1} + a \\
&= H_i(F_i(\check{w}_i^{k+1}))
\end{aligned} \tag{24}$$

其中最后一个等式由下式得到：

$$\begin{aligned}
H_i(F_i(x)) &= \int_{a_i}^{x} c_i(t) f_i(t) \mathtt{d}t \\
&= \int_{a_i}^{x} (t f_i(t) - 1 + F_i(t)) \mathtt{d}t \\
&= t F_i(t)|_0^x - (x - a) \\
&= x F_i(x) - x + a
\end{aligned} \tag{25}$$

对于$\check{A}_k$的任意分配规则$\check{Q}$，我们可以为A构建一个对应的规则：

$$q_i(t) = \max\{\check{q}_i^k | \check{w}_i^k \leq t\} \tag{26}$$

注意到q可以被这样的分配函数$Q_i(t)$实现：$Q_i(t)$的值为$\check{Q}_i(\check{w}_1^{k_1}, \ldots, \check{w}_n^{k_n})$，其中$\check{w}_i^{k_i}$是大于$t_i$的最小类

型。接下来我们可以计算买家i付款：

$$\begin{aligned} P_i(q_i) &= \int_{a_i}^{b_i} c_i(t)q_i(t)f_i(t)dt \\ &= \sum_{k=1}^{K}(\int_{\check{w}_i^k}^{\check{w}_i^{k+1}} c_i(t)f_i(t)dt)\check{q}_i^k \\ &= \sum_{k=1}^{K}(\check{H}_i^{k+1}-\check{H}_i^k)\check{q}_i^k \\ &= q_i(b_i)H_i(F_i(b_i)) \\ &\quad -\sum_{k=1}^{K-1} H_i(F_i(\check{w}_i^{k+1}))(q_i^k-q_i^k) \end{aligned} \tag{27}$$

$$\begin{aligned} \check{P}_i(\check{q}_i) &= \sum_{k=1}^{K}\check{c}_i^k\check{q}_i^k\check{f}_i^k \\ &= \sum_{k=1}^{K}(\check{H}_i^k-\check{H}_i^{k-1})\check{q}_i^k \\ &= -\sum_{k=1}^{K-1}\check{H}_i^k(\check{q}_i^{k+1}-\check{q}_i^k)-\check{H}_i^0\check{q}_i^1+\check{H}_i^K\check{q}_i^K \\ &= -\sum_{k=1}^{K-1} H_i(F_i(\check{w}_i^{k+1}))(q_i^{k+1}-q_i^k) \\ &\quad +q_i(\check{w}_i^{K+1})H_i(F_i(\check{w}_i^{K+1})) \\ &= q_i(b_i)H_i(F_i(b_i)) \\ &\quad -\sum_{k=1}^{K-1} H_i(F_i(\check{w}_i^{k+1}))(q_i^k-q_i^k) \end{aligned} \tag{28}$$

因此，对于$\check{A}_k$最优的简化型，我们有对应的简化型A，并且它们的最大收益相同。

故$R(\check{A}_k) \leq R(A)$。接下来我们证明$R(A) \leq R(\hat{A}_k)$。

由(20)以及$\check{H}$和$\hat{H}$的定义我们有：

$$\hat{H}(z)=\check{H}(z)+d_i z \tag{29}$$

对于A的任意分配规则Q,我们可以为$\hat{A}_k$构建一个相应的规则：

$$\hat{q}_i(\hat{w}_i^k)=\left(\int_{\hat{w}_i^{k-1}}^{\hat{w}_i^k} q_i(t_i)f_i(t_i)dt_i\right)/\hat{f}_i^k \tag{30}$$

由(16)和(17)可以直接证明$\hat{q}$是可实现的。[2] 提供了一个算法实现了$\hat{q}$。

下一步我们计算$\hat{A}_k$和A中买家i的付款：

$$\begin{aligned} \hat{P}_i(\hat{q}_i) &= \sum_{k=1}^{K}\hat{c}_i^k\hat{q}_i^k\hat{f}_i^k \\ &= \sum_{k=1}^{K}\hat{c}_i^k\int_{\hat{w}_i^{k-1}}^{\hat{w}_i^k} q_i(t_i)f_i(t_i)dt_i \\ &= \int_{a_i}^{b_i} q_i(t)d\hat{H}_i(F_i(t_i)) \\ &= q_i(t)\hat{H}_i(F_i(t))|_{a_i}^{b_i}-\int_{a_i}^{b_i}\hat{H}_i(F_i(t))dq_i(t) \\ &= q_i(b_i)\hat{H}_i(F_i(b_i))- \\ &\quad \int_{a_i}^{b_i}\hat{H}_i(F_i(t))dq_i(t) \end{aligned} \tag{31}$$

$$\begin{aligned} P_i(q_i) &= \int_{a_i}^{b_i} c_i(t)q_i(t)f_i(t)dt \\ &= \int_{a_i}^{b_i} q_i(t)dH(F_i(t)) \\ &= q_i(t)H(F_i(t))|_{a_i}^{b_i}-\int_{a_i}^{b_i} H_i(F_i(t))dq_i(t) \\ &= q_i(b_i)H(F_i(b_i))- \\ &\quad \int_{a_i}^{b_i} H_i(F_i(t))dq_i(t) \end{aligned} \tag{32}$$

注 意 到$F_i(b_i) = 1$以 及$\hat{H}(1) = H(1)+d_i$,故 比 较(31)与(32),只 需 证 明$\hat{H}(z) \leq H(z)+d\forall z \in [0,1]$，即 可 证 明$R(q) \leq R(\hat{q})$。对 于任 意 的$\hat{w}_i^k$,由(24)及(29)，可 以 证 明$\hat{H}(F(\hat{w}_i^k)) = H(F(\hat{w}_i^k))+d_iF(\hat{w}_i^k)$。而对于$t_i \in [a_i,b_i]$,假设$\hat{w}_i^k \leq t_i \leq \hat{w}_i^{k+1}$,且$\lambda=\frac{F(t_i)-F(\hat{w}_i^k)}{F(\hat{w}_i^{k+1})-F(\hat{w}_i^k)}$ $(\lambda \leq 1)$。我们有

$$\check{H}(F(\hat{w}_i^k)) = a-\hat{w}_i^k+\hat{w}_i^kF(\hat{w}_i^k) \tag{33}$$

$$\begin{aligned} \check{H}(F(\hat{w}_i^{k+1})) &= a-\hat{w}_i^k-d_i \\ &\quad +(\hat{w}_i^k+d_i)(F(\hat{w}_i^k)+\hat{f}(\hat{w}_i^{k+1})) \\ &= \check{H}(F(\hat{w}_i^k))-d_i+d_iF(\hat{w}_i^k)+ \\ &\quad d_if(\hat{w}_i^{k+1})+\hat{w}_i^kf(\hat{w}_i^{k+1}) \end{aligned} \tag{34}$$

因为$\check{H}$是分段线性函数，所以

$$\begin{aligned}\check{H}(F(t_i)) &= \check{H}(F(\hat{w}_i^k)) + \lambda(-d_i + d_iF(\hat{w}_i^k) + d_if(\hat{w}_i^{k+1}) + \hat{w}_i^kf(\hat{w}_i^k)) \\ &\le \check{H}(F(\hat{w}_i^k)) + \lambda\hat{w}_i^kf(\hat{w}_i^k) \\ &= a - \hat{w}_i^k + \hat{w}_i^kF(t_i) \\ H(F(t_i)) &= a - t_i + t_iF(t_i) \\ \hat{H}(F(t_i)) - H(F(t_i)) &\le d_i(F(t_i)) + (t_i - \hat{w}_i^k)(1 - F(t_i)) \\ &\le d_i\end{aligned}$$

□

附录:引理 2.1证明

证明. 因为(q,p)满足A_1的δ-BIC和IR条件, 我们有

$$\vec{q}_i(\vec{t}_i)\cdot\vec{v}_i(\vec{t}_i) - p_i(\vec{t}_i) \ge \vec{q}_i(\vec{s}_i)\cdot\vec{v}_i(\vec{t}_i) - p_i(\vec{s}_i) - \delta \ \ \forall\vec{t}_i, \vec{s}_i$$

$$\vec{q}_i(\vec{t}_i)\cdot\vec{v}_i(\vec{t}_i) - p_i(\vec{t}_i) \ge 0 \ \forall\vec{t}_i$$

由p'的定义知

$$\vec{q}_i(\vec{t}_i)\cdot\vec{v}_i(\vec{t}_i) - p_i(\vec{t}_i) = \vec{q}_i(\vec{t}_i)\cdot\vec{u}_i(\vec{t}_i) - p'_i(\vec{t}_i)$$

$\forall\vec{t}_i$ 故(q,p')满足A_2的IR条件。

另一方面，

$$\begin{aligned}&\vec{q}_i(\vec{s}_i)\cdot\vec{u}_i(\vec{t}_i) - p'_i(\vec{s}_i) \\ =\ &\vec{q}_i(\vec{s}_i)\cdot\vec{v}_i(\vec{t}_i) + \vec{q}_i(\vec{s}_i)(\vec{u}_i(\vec{t}_i) - \vec{v}_i(\vec{t}_i)) - p_i(\vec{s}_i) - \vec{q}_i(\vec{s}_i)(\vec{u}_i(\vec{s}_i) - \vec{v}_i(\vec{s}_i)) \\ \le\ &\vec{q}_i(\vec{s}_i)\cdot\vec{v}_i(\vec{t}_i) + |\vec{u}_i(\vec{t}_i) - \vec{v}_i(\vec{t}_i)|_1 - |\vec{u}_i(\vec{s}_i) - \vec{v}_i(\vec{s}_i)|_1 \\ \le\ &\vec{q}_i(\vec{s}_i)\cdot\vec{v}_i(\vec{t}_i) + 2m|u-v|_\infty \qquad (35)\end{aligned}$$

$\forall\vec{s}_i,\vec{t}_i$. 所以$(q,p')$也满足$A_2$的$(\delta + 2m|u-v|_\infty)$-BIC条件。 □

附录:引理 2.3证明

证明. 假设估价增加之后，拍卖变为了$A' = (n,m,T,u,F)$，其中$\forall i,\forall j,\forall t, u_{ij}(\vec{t}_i) = v_{ij}(\vec{t}_i) + d$。假设一组简化型和付款$(q,p)$最大化$A$的收益。令$p'$表示另一个付款函数，满足$p'_i(\vec{t}_i) = p_i(\vec{t}_i) + \vec{d}\cdot\vec{q}_i(\vec{t}_i)$。

$$\begin{aligned}&\vec{q}_i(\vec{s}_i)\cdot\vec{u}_i(\vec{t}_i) - p'_i(\vec{s}_i) \\ =\ &\vec{q}_i(\vec{s}_i)\cdot(\vec{v}(\vec{t}_i) + \vec{d}) - p'_i(\vec{s}_i) \\ =\ &\vec{q}_i(\vec{s}_i)\cdot\vec{v}_i(\vec{t}_i) + \vec{d}\cdot\vec{q}_i(\vec{s}_i) - p'_i(\vec{s}_i) \\ =\ &\vec{q}_i(\vec{s}_i)\cdot\vec{v}_i(\vec{t}_i) - p_i(\vec{s}_i) \qquad (36)\end{aligned}$$

$\forall\vec{t}_i$, $\vec{s}_i$。所以(q,p')满足A'的δ-BIC和IR条件。

当使用(q,p')时, A'收益等于

$$\begin{aligned}&\int_T\sum_{i\in N}(p'_i(\vec{t}_i)f_i(\vec{t}_i))dt \\ =\ &R(A) + \int_T\sum_{i\in N}(\vec{d}\cdot\vec{q}_i(\vec{t}_i)f_i(\vec{t}_i))dt \\ =\ &R(A) + \sum_{j\in M}\int_T\sum_{i\in N}(d_jq_{ij}(\vec{t}_i)f_i(\vec{t}_i))dt \\ =\ &R(A) + \sum_{j\in M}d_j\sum_{i\in N}\int_{T_i}(q_{ij}(\vec{t}_i)f_i(\vec{t}_i))d\vec{t}_i \\ =\ &R(A) + \sum_{j\in M}d_j\sum_{i\in N}Pr[\text{买家}i\text{获得物品}j] \\ =\ &R(A) + \sum_{j\in M}d_jPr[\text{物品}j\text{被卖出}] \qquad (37)\end{aligned}$$

所以$R(A') \ge R(A) + \sum_{j\in M}d_jPr[\text{物品}j\text{在最优拍卖}A\text{中被卖出}]$ 类似的$R(A) \ge R(A') - \sum_{j\in M}d_jPr[\text{物品}j\text{在最优拍卖}A'\text{中被卖出}]$ 所以$R(A) \le R(A') \le R(A) + |\vec{d}|_1$. □

附录：简化型的虚拟VCG求解方案

通过简化型解决拍卖问题，也是多物品拍卖中的一个核心问题。Cai等人设计了一个FPRAS算法，将最大化拍卖收益的问题转化成为社会福利最大化问题 [3]。首先证明了一组函数的实现可以通过一个分离预示来检验，然后设计出一个线性规划来求解最大收益满足IC和IR条件时对应的简化型。最后通过几个虚拟VCG实现简化型。这一过程可分解为以下步骤。

5.1 虚拟VCG

定义**5.1.** 一个虚VCG分配规则是一组加权函数，$f_i : T_i \rightarrow R^n$. f_i将买家i的一个类型映射到买家i 的一个虚类型。对于任意类型组合$\vec{v}$，虚VCG分配规则以及函数组$\{f_i\}_{i\in[m]}$在输入为$(f_1(\vec{v_1}),\ldots,f_m(\vec{v_m}))$时得到$VCG_{\mathcal{F}}$。$VVCG_{\mathcal{F}}(\{f_i\}_{i\in[m]})$表示在可行性约束$\mathcal{F}$和加权函数$\{f_i\}_{i\in[m]}$的条件下的虚$VCG$分配规则。

定义**5.2.** 令$\vec{w}\in R^{n\sum_i=1^m|T_i|}$。定义$f_i$使得$f_{ij}(A) = \frac{w_{ij}(A)}{Pr[t_i=A]}$。那么$VVCG_{\mathcal{F}}(\vec{w})$即为虚$VCG$分配规则$VVCG_{\mathcal{F}}(\{f_i\}_{i\in[m]})$。

5.2 简化型

定义**5.3.** 一个分配规则的简化型是一个向量值函数$\pi(\cdot)$，对应在任意买家i、物品j 和类型$A\in T_i$ 上

取值为$\pi_{ij}(A)$。$\pi_{ij}(A)$表示买家i在类型为A时得到物品j 的概率，其中概率涉及所有其他买家的类型的随机性以及分配规则的内部随机性。这里我们假设其他买家关于他们的类型也都是诚实的。

对于任意加权向量$\vec{w} \in R^{n\sum_{i}=1^{m}|T_i|}$，令$R_{\mathcal{F}}(\vec{w})$表示$VVCG_{\mathcal{F}}(\vec{w})$的简化型，$W_{\mathcal{F}}(\vec{w})$表示所有物品在$VVCG_{\mathcal{F}}(\vec{w})$下总的期望权重，即$W_{\mathcal{F}}(\vec{w}) = R_{\mathcal{F}}(\vec{w}) \cdot \vec{w}$。

给定任意可行性约束条件$\mathcal{F} \subseteq 2^{\mathcal{A}}$以及任意买家类型的分布$\mathcal{D}$，并不是所有满足$\forall i, j, A, 0 \le \pi_{ij}(A) \le 1$的简化型都是可行的，即可以通过特定的分配规则实现。但是，所有可行的简化型组成一个凸多面体。令$F(\mathcal{F}, \mathcal{D})$表示在可行性约束条件为$\mathcal{F}$而买家类型服从分布$\mathcal{D}$时所有可行的简化型构成的集合。则我们有下面的定理：

定理**5.1.** 如果$|\mathcal{D}|$是有限的，则$F(\mathcal{F}, \mathcal{D})$是一个凸多面体。

5.3 分离预示

我们用一个分离预示$SO(\vec{\pi})$来检验$\vec{\pi}$是否属于$F(\mathcal{F}, \mathcal{D})$。如果$\vec{\pi} \in F(\mathcal{F}, \mathcal{D})$,则$SO(\vec{\pi})$的输出结果为"yes";否则输出一个超平面（凸多面体$F(\mathcal{F}, \mathcal{D})$的一个边界），这一超平面将$\vec{\pi}$和$F(\mathcal{F}, \mathcal{D})$分离。$SO$可以通过下面的定理实现：

定理**5.2.** 一个简化型$\vec{\pi}$在$\mathcal{F}$和$\mathcal{D}$意义下是可行的当且仅当对于任意$\vec{w} \in [-1, 1]^{n\sum_{i}=1^{m}|T_i|}$，都有$\vec{\pi} \cdot \vec{w} \le W_{\mathcal{F}(\vec{w})}$。

这样一个分离预示检验$\vec{\pi}$是否属于$F(\mathcal{F}, \mathcal{D})$的步骤如下：运行下面的线性规划$SOLP$，最小化$g_{\vec{\pi}} = W_{\mathcal{F}}(\vec{w}) - \vec{\pi} \cdot \vec{w}$,记输出的最优解为$t^*, \vec{\pi}^*$。如果线性规划的结果为负，那么我们知道$\vec{w}^* \cdot \vec{\pi} > t^* = W_{\mathcal{F}}(\vec{w})$，同时得到所需的超平面。否则，输入的简化型是可行的，输出为"yes"。$SOLP$：

- 变量：
  - t，表示$W_{\mathcal{F}}$的值。
  - $w_i j A$。
- 约束：
  - $-1 \le w_{ij}(A) \le 1$，对于任意买家i、物品j和类型$A \in T_i$。
  - $t \ge R_{\mathcal{F}}(\vec{w}) \cdot \vec{w}$。
- 最小化：
  - $t - \vec{\pi} \cdot \vec{w}$。

5.4 将简化型分解为一组虚VCG

定理**5.3.** n维多面体P的每一个内点$\vec{x}$可以被写成至多$n+1$个P的端点的凸组合。

算法：将$\vec{x}$写成至多$n+1$个端点的凸组合

1:Initialize: $i = 1, \vec{y} = \vec{0}, \vec{z} = \vec{x}, E = \emptyset, c_i = 0, \vec{a_i} = \vec{0}, \forall i \in [n+1], c = \sum_i c_i, \vec{y}\frac{1}{c}\sum_i c_i \vec{a_i}, c\vec{y} + (1-c)\vec{z} = \vec{x}$。
2:WHILE $c < 1$ DO
3: SET$\vec{a_i} = CO(E)$.
4: IF $\vec{a_i} = \vec{z}$ THEN
5: SET $c_i = 1 - c$.
6: OUTPUT $c_1, \ldots, c_{n+1}, \vec{a_1}, \ldots, \vec{a_{n+1}}$.
7: ELSE
8: SET $D = \max\{d | (1+d)\vec{z} - d\vec{a_i} \in P\}$.
9: SET $E_i = SO((1 + D + \epsilon)\vec{z} - (D + \epsilon)\vec{a_i})$ for $\epsilon > 0$.
10: UPDATE $c_i = (1 - \frac{1}{1+D})(1 - c), E = E \Cup E_i, c, \vec{y}, \vec{z}$.
11: END IF 12:END WHILE

一个角预示(CO)是一个输入为一些超平面输出为超平面的公共点（或任意一点，如果这些超平面没有公共点）的预示。

定理5.4. 上述算法的角预示可以通过输出虚VCG $VVCG_{\mathcal{F}}(\vec{w})$的简化型来实现，其中$\vec{w}$是输入超平面（用向量表示）的平均向量。