

基于双边拍卖的群智感知数据复用机制设计

江常坤¹, 高林²

(1. 深圳大学计算机与软件学院, 广东 深圳 518060; 2. 哈尔滨工业大学(深圳)电子与信息工程学院, 广东 深圳 518055)

摘要: 群智感知是一种有前景的感知范式, 其主要借助大量移动设备中的多样嵌入式传感器来完成数据感知任务。群智感知中的一个关键问题是如何有效地协调移动设备用户同时执行多项感知任务。通过在感知任务和用户之间引入一个新的数据层, 从而有效地利用感知任务的相似性和用户的异质性, 在数据层之上建立联合任务选择和用户调度问题, 旨在最大化整个系统的社会福利。这个问题因其组合特性以及存在感知任务和用户两方面的私有信息而很难解决。为了处理上述问题, 提出了一种双边随机拍卖机制, 并证明其能够满足计算高效性、个体理性和期望意义下的激励相容性。仿真结果表明, 所提的随机拍卖机制可以接近最优的社会福利, 并且数据复用带来的社会福利收益会随着任务相似性的增强而显著增加。

关键词: 群智感知; 数据复用; 激励机制; 拍卖机制

中图分类号: TN915

文献标识码: A

doi: 10.11959/j.issn.2096-3750.201900109

Mechanism design for crowd sensing with data reuse based on two-sided auction

JIANG Changkun¹, GAO Lin²

1. College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

2. School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China

Abstract: Crowd sensing is a promising sensing paradigm, which mainly uses a variety of embedded sensors in a large number of mobile devices to accomplish data sensing tasks. One of the key issues in crowd sensing is how to effectively coordinate mobile device users to perform multiple sensory tasks simultaneously. By introducing a new data layer between the sensing task and the user, the similarity of the sensing task and the heterogeneity of the user were effectively utilized, and the joint task selection and user scheduling problems were established on the data layer, aiming at maximizing the social welfare of the whole system. This problem was difficult to solve due to its combinatorial nature and the presence of private information on both the sensing tasks and the users. In order to deal with these problems, a two-sided randomized auction mechanism was proposed, and it was proved that it can satisfy the desirable properties of the computational efficiency, the individual rationality, and the incentive compatibility in expectation. The simulation results show that the proposed stochastic auction mechanism can achieve nearly optimal social welfare, and the social welfare benefits brought by data reuse will increase significantly with the enhancement of task similarity.

Key words: crowd sensing, data reuse, incentive mechanism, auction mechanism

1 引言

带有多种嵌入式传感器的手持移动设备数量的激增催生了一种新的、快速增长的感知范式, 被称为群智感知^[1]。在群智感知中, 可以通过处理大

量的由个人移动设备收集的感知数据来获取关于某些事件的有用信息。在一个通用的多任务群智感知系统中^[2-3], 感知任务所有者可以通过感知平台发起感知任务, 然后推送到移动设备前端, 一部分移动设备将感知所需数据和收集的数据发送回系统。

用户参与感知任务时，会消耗如电池能量和 CPU 时间等资源，需要付出一定的成本。因此，用户可能不会参与到感知任务中，除非他们获得令人满意的奖励用以补偿感知成本。许多前期的研究^[4-13]专注于激励用户参与到群智感知系统中。

这些群智感知的前期工作^[4-13]侧重于感知任务和用户之间的直接交互，如通过恰当的匹配在用户之间分配感知任务，而没有考虑多个不同感知任务之间的共同数据需求或不同用户异构的感知能力。然而，在实际系统中，多个感知任务可能需要一些共同的数据^[1]，如特定时间和地点的道路交通数据可能对 Waze、Uber 和 Google Traffic 同时有用。因此，如果多个感知任务由相同的用户分别完成，则可能在多感知任务场景中导致重复的感知和处理。此外，由于位置和设备类型不同，用户可能具有不同的感知能力，如用户更容易感知靠近其当前位置的数据。因此，在数据层调度用户比在任务层调度用户具有更强的灵活性和高效性。

为了更有效地完成多个感知任务，确定这些任务共同的数据需求并在不同任务中复用感知数据至关重要。具体而言，群智感知平台可以允许任务开发人员以某种高级语言指定他们的数据需求，然后可以跨多个感知任务识别并复用共同的数据，以减少或避免重复的感知与处理。允许数据复用有以下优点：1) 感知数据是一种数字商品，重复使用无须额外成本；2) 多个感知任务可以通过平台共享大量移动用户；3) 通过不同任务复用数据，可以极大地提高系统效率。

为此，本文提出了一种新的 3 层以数据为中心的群智感知模型，包括数据层、任务层和用户层，这与文献[4-13]中传统的两层以任务为中心的模型不同。具体来说，在此模型中，任务和用户通过数据层连接：每个任务都被转换为它需要的一组数据项，每个用户都与一组可以感知的数据项相关联。此外，不同的任务可能需要共同的数据项（可以复用这些数据项）；不同的用户可以感知相同的数据项（竞争感知机会）。因此，此模型能够利用任务相似性（数据需求）和用户异质性（感知能力）。3 层以数据为中心的群智感知模型如图 1 所示，图 1 显示了具有 2 个任务、3 个用户和 6 个数据项的模型，其中，任务 1 和任务 2 都需要数据项 {5}，左侧和中间用户能够同时感知数据项 {2}，中间和右侧用

户能够同时感知数据项 {5}。

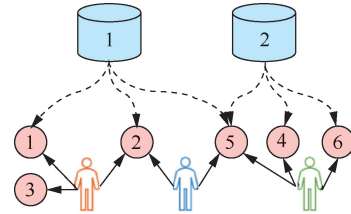


图 1 3 层以数据为中心的群智感知模型

在这个模型中，群智感知平台收集感知任务的数据需求和用户的感知能力，然后决定是否完成以及如何通过一组正确的用户选择完成这些任务。本文专注于解决两个问题：哪些感知任务可以完成？哪些用户将被调度用来感知哪些数据？

同时，专注于整个系统社会福利最大化，社会福利定义为已完成任务的总价值与被调度用户的总成本之间的差异。然而，解决这个问题具有巨大的挑战性。首先，由于本质上是组合问题，这个问题是 NP-hard 问题；其次，求解这个问题需要任务价值和用户感知成本的完整信息，但这些信息通常是私有信息。因此，有必要设计一个真实的激励机制，让任务所有者和用户真实上报这些私有信息。然而，一些经典的真实拍卖机制如 VCG 拍卖^[14-15]，因其具有较高的计算复杂性而不适合求解此问题。

为了解决上述问题，本文采用一种随机拍卖的方法^[16]，将感知平台作为拍卖师，将感知任务和用户作为投标人。提出的真实随机拍卖机制由两个规则组成：一是随机分配规则，根据某个概率分布，从一组目标解决方案中随机选择一个可行的“分配”；二是支付规则，为每个竞标者在任何可能的分配中指定一个相应的支付。与广泛应用于网络^[17]和云计算^[18]中的随机拍卖不同，本文提出的随机拍卖结构是双边的，需要在互相信息不对称的情况下决定感知任务选择（价值）和用户调度（成本）；而文献[17]和文献[18]中的随机拍卖结构是单边的（价值或成本）。本文的主要成果和贡献如下。

1) 提出了一种以数据为中心的新型群智感知模型。目前利用多个感知任务的数据复用来优化多任务群智感知系统的研究还较少。本文提出了一种新型的 3 层以数据为中心的模型，能够利用任务相似性和用户感知异质性。

2) 随机拍卖机制的设计和分析。为了解决感知

任务和用户选择的复杂性问题并获得感知任务和用户双边的私有信息,本文提出了一种双边的真实随机拍卖机制,这种机制具有计算高效性、个人理性和期望上的真实性。

3) 仿真结果和见解。仿真结果显示,本文提出的随机拍卖机制可以达到最大社会福利基准的90%。此外,允许数据复用时,所达到的社会福利随着任务相似性的增强而增加,因为高度相似的任务可以由较小的一组调度用户完成;而在没有数据复用时,由于感知任务之间的竞争加剧,所达到的社会福利随着任务相似性的增强而降低。

2 系统模型与问题描述

2.1 网络模型

考虑一个通用的多任务群智感知平台,包含一组表示为 $[J]=\{1,2,\dots,J\}$ 的 J 个任务,一组表示为 $[I]=\{1,2,\dots,I\}$ 的 I 个移动用户以及一组表示为 $[K]=\{1,2,\dots,K\}$ 的 K 个目标数据项。每个数据项 $k\in[K]$ 由一组参数,如数据类型、位置和时间来表征。每个任务 $j\in[J]$ 与一组数据需求 $K_j\subset[K]$ 相关联,并且每个用户 $i\in[I]$ 能够感知一个特定集合 $[S_i]\subset[K]$ 的数据项。由于不同任务可以复用相同的数据项,因此,可能存在具有重叠数据需求的多个任务 j_1 和 j_2 ,即 $K_{j_1}\cap K_{j_2}\neq\emptyset$ 。

图1以时隙方式运行,其中,整个时间段被划分为多个时隙,每个时隙可以是一小时或一天,取决于感知任务或用户的变化。为了不失一般性,本文考虑一个特定时隙的运行。在该时隙开始时,每个任务发起者在平台上注册其任务,指示任务的数据需求和任务完成时可以获得的潜在价值;每个用户在平台上报告其感知信息,指示其感知能力(即其可以感知的数据项集合)和在其能力范围内感知任何数据项子集的潜在成本。在收集了所有任务发起者和用户报告的信息后,平台决定具体的任务选择(即选择要完成的一组任务)和用户调度(即调度一组用户来感知相关的数据项)。

要注意的是,任务发起者或用户可能会误报信息以寻求改善其个人收益。本文考虑以下简化的场景:1) 用户可能误报其感知成本,但不会误报其感知能力;2) 任务发起者可能误报其任务价值,但不会误报其数据需求(因为平台可以检查这些信息的

正确性)。

2.2 感知任务模型

每个任务 $j\in[J]$ 与一组数据需求 $K_j\subset[K]$ 以及完成后的任务价值 $v_j>0$ 相关联。任务价值 v_j 是任务 j 的私有信息,不能被平台、用户或其他任务发起者观察到。假设任务 j 被完成,当且仅当 K_j 中的每一个数据项被至少一个用户感知。令 $z_j\in\{0,1\}$ 表示任务 $j\in[J]$ 是否完成, $y_k\in\{0,1\}$ 表示数据项 $k\subset[K]$ 是否被至少一个用户感知,对于每个任务 $j\in[J]$ 有以下约束

$$z_j \leq y_k, \forall k \in K_j \quad (1)$$

给定一组可行的任务选择 $z=(z_j, j\in[J])$,总共达到的任务价值(所有已完成的任务)为

$$V(z) = \sum_{j\in[J]} v_j z_j \quad (2)$$

2.3 用户模型

每个用户 $i\in[I]$ 与其感知能力相关联,即他可以感知数据项的集合 $[S_i]$ 。平台可以调度用户 i 去感知其感知能力内的数据项子集 $S\subset[S_i]$,相应的感知成本为 $c_i(S)$ 。令 $x_i(S)\in\{0,1\}$ 表示用户 i 是否被调度去感知数据集 $S\subset[S_i]$ 。当 $S=\emptyset$ 时,则 $x_i(\emptyset)=1$ 表示用户 i 未被调度去感知任何数据集,因此,感知成本为0,即 $c_i(\emptyset)=0$ 。

假设用户只能被调度感知其能力范围内的一个数据集,也就是说,对于每个用户 $i\in[I]$,有以下用户调度约束

$$\sum_{S\subset[S_i]} x_i(S) = 1 \quad (3)$$

如果用户被调度感知多个数据集,如 S_{i_1} 和 S_{i_2} ,总能等效地认为他被调度来感知数据集 $S_{i_1}\cup S_{i_2}$ 。令 $\mathbf{x}_i=(x_i(S), S\subset[S_i])$ 表示用户 i 的调度向量,给定一个可行的用户调度 $\mathbf{x}=(\mathbf{x}_i, i\in[I])$,所有调度用户总共产生的感知成本为

$$C(\mathbf{x}) = \sum_{i\in[I]} \sum_{S\subset[S_i]} c_i(S) x_i(S) \quad (4)$$

令 $y_{ki}\in\{0,1\}$ 表示用户 i 是否感知到数据项 k ,即 $y_{ki} = \sum_{S\subset[S_i]: k\in S} x_i(S)$ 。同时,已定义 $y_k\in\{0,1\}$ 表示数据项 $k\subset[K]$ 是否已被至少一个用户感知。那么,对于每个数据项 $k\subset[K]$,有以下约束

$$y_k \leq \sum_{i \in [I]} y_{ki} \quad (5)$$

此外，用 $\mathbf{c}_i = (c_i(S), S \subset S_i)$ 表示用户 i 的感知成本向量，包含其可以感知的所有可能数据项子集的成本。在实际中，感知成本向量 \mathbf{c}_i 是用户 i 的私有信息，不能被平台、任务发起者或其他用户观察到，这是本文面临的关键挑战之一。

2.4 问题描述

整个系统的社会福利 $W(x, z)$ 被定义为所有已完成任务的总价值 $V(z)$ 与所有被调度用户的总感知成本 $C(x)$ 之差，即

$$W(x, z) = V(z) - C(x) \quad (6)$$

该平台的目标是确定最好的任务选择 z 和用户调度 x ，以最大化社会福利 $W(x, z)$ 。可以定义如下的联合任务选择和用户调度问题 (P1)。

P1:

$$\begin{aligned} \max_{x, y, z} & V(z) - C(x) \\ \text{s.t.} & (1) - (5), \forall i \in [I], j \in [J], k \in [K] \end{aligned} \quad (7)$$

其中， $\mathbf{y} = (y_k, k \in [K])$ 是一个中间变量，表示每个数据项是否被至少一个用户感知。不难看出，问题 P1 是二进制整数线性规划问题。用 $\{x^o, y^o, z^o\}$ 表示 P1 的最优解，为了更清晰，还可以将 $\{x^o, y^o, z^o\}$ 写为 $\{x^o(c, v), y^o(c, v), z^o(c, v)\}$ ，因为它们为用户感知成本 $\mathbf{c} = (c_i, i \in [I])$ 和任务值 $\mathbf{v} = (v_j, j \in [J])$ 的函数。

但是，解决问题 P1 仍具有不少挑战。首先，问题 P1 是 NP-hard 问题 (P1 的特殊情况可以简化成经典的集合覆盖问题)，因此，有必要设计一种低复杂度的近似算法来找到近似解。其次，解决问题 P1 需要完整的信息，包括所有任务的数据需求、任务价值以及所有用户的感知能力和感知成本。但是，正如前文所述，用户的感知成本和任务价值是他们的私有信息，平台无法观察到。因此，需要设计一个真实的激励机制来得到这样的私有信息。

3 双边 VCG 拍卖机制设计

本节首先提出一个双边 VCG 拍卖机制来解决问题 P1，而不考虑计算复杂性问题。在第 4 节基于本节的 VCG 拍卖机制，进一步提出一种易于

处理的低复杂度随机拍卖机制。

3.1 基于双边拍卖的激励结构

本节提出一种基于双边拍卖的激励结构，其中，平台（拍卖师）向移动用户（一方投标人）购买数据，并将数据出售给感知任务所有者（另一方投标人）。在该拍卖中，首先，平台宣布分配规则（任务选择和用户调度）和支付规则（对被调度用户的支付价格和对所选取任务收取的价格）。然后，每个任务提交一个价值（出价），并且每个用户向平台提交感知成本向量（出价），其可以与真实任务值或成本向量不同。最后，平台根据所有任务和用户的出价以及其他公有信息计算分配和收付款。本文主要考虑设计真实的拍卖机制，其中，任务和用户将如实提交其私有信息。

令 $\mathbf{b}_i = (b_i(S), S \subset S_i)$ 表示用户 i 报告的感知成本向量，其中， $b_i(S)$ 表示用户报告的数据集 $S \subset S_i$ 的感知成本。设 $\mathbf{b} = (\mathbf{b}_i, i \in [I])$ 表示所有用户的出价，并设 $\mathbf{u} = (u_j, j \in [J])$ 表示所有任务的出价，其中， u_j 是任务 j 的报告值。如果拍卖是真实的，均衡时有 $\mathbf{b} = \mathbf{c}$ 和 $\mathbf{u} = \mathbf{v}$ 。用 $\{\mathbf{x}(\cdot), \mathbf{z}(\cdot)\}$ 表示分配规则，其中， $\mathbf{x}(\cdot) = (x_i(\cdot), i \in [I])$ 是用户调度向量， $\mathbf{z}(\cdot) = (z_j(\cdot), j \in [J])$ 是任务选择向量；用 $\{\mathbf{p}(\cdot), \mathbf{q}(\cdot)\}$ 表示支付规则，其中， $\mathbf{p}(\cdot) = (p_i(\cdot), i \in [I])$ 是给用户付费向量， $\mathbf{q}(\cdot) = (q_j(\cdot), j \in [J])$ 是向任务收费向量。注意 $\mathbf{x}(\cdot), \mathbf{z}(\cdot)$ 和 $\mathbf{p}(\cdot), \mathbf{q}(\cdot)$ 可写成 $\mathbf{x}(b, u), \mathbf{z}(b, u)$ 和 $\mathbf{p}(b, u), \mathbf{q}(b, u)$ 。另外，本文将机制 \mathcal{M} 写为 $\mathcal{M} = \{\mathbf{x}(\cdot), \mathbf{z}(\cdot); \mathbf{p}(\cdot), \mathbf{q}(\cdot)\}$ 或 $\mathcal{M} = \{\mathbf{x}(b, u), \mathbf{z}(b, u); \mathbf{p}(b, u), \mathbf{q}(b, u)\}$ 。

3.2 双边 VCG 拍卖机制的意义

本节提出的双边 VCG 拍卖机制是经典 VCG 拍卖的扩展^[14-15]。在拍卖中，分配规则旨在最大化在上报的感知成本和任务价值上定义的社会福利；支付规则旨在向每个调度用户支付该用户的存在对整个系统产生的社会收益，并且对每个选中的任务收取该任务的存在对整个系统产生的社会损失。

为方便起见，可将上述双边 VCG 拍卖机制记为 $\mathcal{M}^o = \{\mathbf{x}^o(\cdot), \mathbf{z}^o(\cdot); \mathbf{p}^o(\cdot), \mathbf{q}^o(\cdot)\}$ 。通过将标准的 VCG^[14-15] 拍卖机制扩展到双边场景中，可以证明双边拍卖机制 \mathcal{M}^o 是真实的，因此，每个用户和每个

任务发起者将如实报告信息，即 $b=c$ 和 $u=v$ 。这意味着 \mathcal{M}^o 是高效的，因为它的分配规则最大化了式(6)中定义的社会福利。

然而，计算双边 VCG 拍卖结果需要解决 NP-hard 问题 P1，这是计算上难以解决的问题。为此，将进一步研究低复杂性的拍卖机制。

4 随机拍卖机制设计

受文献[17]和文献[18]中随机拍卖的启发，本文提出一种低复杂度的真实随机拍卖机制，它在多项式时间内运行。值得注意的是，本文所提的随机拍卖机制由于其双边结构而不同于传统的单边随机拍卖机制^[17-18]。

下面从问题 P1 的线性规划松弛开始，在分数域中获得相应的线性规划问题 P2，进一步推导出分数 VCG 拍卖（在实际中可能无法实现）。然后，通过适当的分解，可将分数 VCG 拍卖转换为可实现的双边随机拍卖。

4.1 线性规划松弛

首先将问题 P1 的可行域放宽到分数域（即每个二进制变量集合 $\{0,1\}$ 都放宽到连续区间 $[0,1]$ ），并将相应的线性规划问题称为问题 P2。问题 P2 可以在多项式时间内求解，因为它是标准的线性规划问题^[19]。将问题 P2 的最优解称为分数最优解，即 $\{x', y', z'\}$ 或 $\{x'(c, v), y'(c, v), z'(c, v)\}$ 。值得注意的是，问题 P2 的最优值是问题 P1 的最大值的上限，它们之间的差距称为整性间隙^[19]。直观上讲，可以将分数解视为用户被调度或任务被选中的时间比例。

下面可以设计相应的分数 VCG 拍卖 \mathcal{M}' ，其分配规则旨在最大化分数域中的社会福利（定义在用户出价 b 和任务所有者出价 u 上）；其支付规则旨在向每个被调度的用户支付该用户的存在对系统产生的社会收益，并对每个被选中的任务收取该任务的存在对系统产生的社会损失。详细的 \mathcal{M}' 机制与 \mathcal{M}^o 类似，除了用分数最优解 $x'(\cdot), z'(\cdot)$ 替换整数最优解 $x^o(\cdot), z^o(\cdot)$ ，还在决定支付时求解问题 P2 而不是问题 P1。

命题 1 分数 VCG 拍卖 \mathcal{M}' 满足个体理性和激励相容（真实）性，并最大化分数域中的社会福利。

值得注意的是，问题 P2 的最优解（或机制 \mathcal{M}' 的结果）可能对问题 P1 来说不可行，这意味着，分数 VCG 拍卖 \mathcal{M}' 可能无法直接实现。本文将分数

VCG 拍卖 \mathcal{M}' 转换为（整数）随机拍卖，可以根据一定的概率随机生成问题 P1 的可行解。

4.2 随机拍卖机制设计基础和标准

4.2.1 随机机制设计基础

首先介绍随机拍卖的基本知识和期望真实性的概念^[16]。

双边确定性机制 $\mathcal{M} = \{x(\cdot), z(\cdot); p(\cdot), q(\cdot)\}$ 由确定性的分配规则 $\{x(\cdot), z(\cdot)\}$ 和支付规则 $\{p(\cdot), q(\cdot)\}$ 组成，并在给定任意的出价 b 和 u 时，返回确定性结果 $\{x(b, u), z(b, u); p(b, u), q(b, u)\}$ 。注意，上面提到的 \mathcal{M}^o 和 \mathcal{M}' 都是确定性的机制。

一个机制 $\mathcal{M}^r = \{x^r(\cdot), z^r(\cdot); p^r(\cdot), q^r(\cdot)\}$ 可以随机化，也就是说，可以通过抛硬币来决定具体的分配和支付，而不是给出确定性的分配规则和支付规则，换句话说，给定任何的出价 b 和 u ，结果 $x_i^r(\cdot), z_j^r(\cdot), p_i^r(\cdot), q_j^r(\cdot)$ 都是随机变量。此外，每个任务的效用（即价值减去付费）和每个用户的效用（即收到的费用减去感知成本）也是随机变量。直观上讲，这种随机机制可以被视为确定性机制的一组随机化机制。对于随机机制，真实性的概念是在期望意义上定义的，也就是说，如果随机机制 \mathcal{M}^r 是期望真实的，上报真实感知成本和价值会最大化每个用户和每个任务的期望效用。

4.2.2 随机机制设计标准

真实随机机制设计的核心思想是找到与一个确定性机制相等效的随机机制。

受文献[9]和文献[10]中定义的 α -缩放的分数机制的启发，首先为一个确定性的机制 $\mathcal{M} = \{x(\cdot), z(\cdot); p(\cdot), q(\cdot)\}$ 引入 (α, β) -缩放的分数机制。文献[9]和文献[10]中的 α -缩放的分数机制与这里的 (α, β) -缩放的分数机制之间的关键区别在于：前者只考虑单边缩放，而后者考虑的是双边缩放。

定义 1（缩放的分数机制） 机制 $\mathcal{M} = \{x(\cdot), z(\cdot); p(\cdot), q(\cdot)\}$ 的一个 (α, β) -缩放的分数机制 $\mathcal{M}_{(\alpha, \beta)} = \{x_\alpha(\cdot), z_\beta(\cdot); p_\alpha(\cdot), q_\beta(\cdot)\}$ 定义为

$$x_\alpha(\cdot) = \alpha \cdot x(\cdot), p_\alpha(\cdot) = \alpha \cdot p(\cdot) \quad (8)$$

$$z_\beta(\cdot) = \beta \cdot z(\cdot), q_\beta(\cdot) = \beta \cdot q(\cdot) \quad (9)$$

其中， α, β 是缩放因子，并且满足 $0 \leq \alpha \cdot x(\cdot) \leq 1$ 和 $0 \leq \beta \cdot z(\cdot) \leq 1$ 。

与原始机制 \mathcal{M} 相比，在 (α, β) -缩放的分数机

制中，每个用户产生的成本和支付用 α 缩放，每个任务的价值和付费用 β 缩放。这意味着，这两种机制中的用户和任务所有者的最优出价策略是等价的，因此，其真实性相互等价。

命题 2 如果机制 \mathcal{M} 是真实的，那么它的 (α, β) -缩放的分数机制 $\mathcal{M}_{(\alpha, \beta)}$ 也是真实的。

基于上述内容，本文提出以下真实的随机机制设计标准：找到一个随机机制 \mathcal{M}' ，它产生与 $\mathcal{M}_{(\alpha, \beta)}$ 等效的结果，其中， $\mathcal{M}_{(\alpha, \beta)}$ 是分数 VCG 拍卖机制 \mathcal{M}' 的 (α, β) -缩放的分数机制。由于 \mathcal{M}' 是真实的，根据命题 2 可以获得其 (α, β) -缩放的分数机制的真实性。此外，由于随机机制 \mathcal{M}' 产生与 $\mathcal{M}_{(\alpha, \beta)}$ 相同的任务选择、用户调度和支付，可以进一步获得 \mathcal{M}' 的真实性（在期望意义下）。

4.2.3 真实的随机机制设计

本节给出真实随机拍卖设计的详细内容。

为方便起见，可把一个随机机制 $\mathcal{M}' = \{x^r(\cdot), z^r(\cdot); p^r(\cdot), q^r(\cdot)\}$ 表达成在所有可能的分配下的一组分配概率 $\lambda = (\lambda^i)_{i \in \mathcal{A}}$ 和一组支付规则 $\{p^i(\cdot), q^i(\cdot)\}_{i \in \mathcal{A}}$ ，其中， \mathcal{A} 是所有可行的整数分配的集合（关于 x 和 z ）， $\lambda^i \geq 0$ 是选取一个特定的分配 $\{x^i, z^i\}$ 和相应的付款 $\{p^i, q^i\}$ 的概率。这样，设计随机机制 \mathcal{M}' 等价于找到一组分配概率 $\lambda = (\lambda^i)_{i \in \mathcal{A}}$ 和一组支付规则 $\{p^i(\cdot), q^i(\cdot)\}_{i \in \mathcal{A}}$ 。

下面提出随机拍卖机制 \mathcal{M}' ，其目标是在将分数最优解精确分解为整数解的加权和的情况下，最大化双边缩放的社会福利。由于考虑的是双边社会福利的最大化， \mathcal{M}' 非平凡地拓展了文献[17]和文献[18]中的单边效用最大化和单边成本最小化。

随机拍卖机制 \mathcal{M}' ：从分数 VCG 拍卖机制 $\mathcal{M}' = \{x^r(\cdot), z^r(\cdot); p^r(\cdot), q^r(\cdot)\}$ 出发，定义如下由分配规则和支付规则组成的随机拍卖机制 \mathcal{M}' 。

1) 分配规则 $\lambda^r = (\lambda^i)_{i \in \mathcal{A}}$

$$\begin{aligned} \lambda^r &= \operatorname{argmax}_{\lambda, 0 < \alpha, \beta \leq 1} \beta V^r - \alpha C^r \\ \text{s.t. } &\sum_{i \in \mathcal{A}} \lambda^i x_i^i = \alpha x_i^i(b, u), \forall i \in [I] \\ &\sum_{i \in \mathcal{A}} \lambda^i z_j^i = \beta z_j^i(b, u), \forall j \in [J] \end{aligned} \quad (10)$$

其中， V^r 和 C^r 是关于分数最优解 $z_j^r(b, u)$ 和 $x_i^r(b, u)$ 的最优总任务价值和总用户成本。

2) 支付规则 $\{p^i(b, u), q^i(b, u)\}_{i \in \mathcal{A}}$

$$p_i^i(b, u) = \alpha p_i^i(b, u) \frac{C_i(x_i^i)}{\sum_{i \in \mathcal{A}} \lambda_i C_i(x_i^i)}, \forall i \in [I] \quad (11)$$

$$q_j^i(b, u) = \beta q_j^i(b, u) \frac{V_j(z_j^i)}{\sum_{i \in \mathcal{A}} \lambda_i V_j(z_j^i)}, \forall j \in [J] \quad (12)$$

其中， $C_i(x_i^i)$ 是用户 i 在分配 x_i^i 下的总成本， $V_j(z_j^i)$ 是任务 j 在分配 z_j^i 下的总价值。

可以看到，在 \mathcal{M}' 中每个用户的期望收入和感知成本以及每个任务的期望付费和价值都与分数 VCG 拍卖 \mathcal{M}' 中的相同，这意味着， \mathcal{M}' 在期望意义下是真实的。

命题 3（期望激励相容性） 随机拍卖机制 \mathcal{M}' 在期望意义下是激励相容的，因为每个用户和任务报告真实感知成本和价值时可最大化其期望效用。

可以进一步验证，在 \mathcal{M}' 中，每个用户和任务可以在任何可能的分配下实现非负效用，这意味着， \mathcal{M}' 满足严格意义上的个体理性。

命题 4（个体理性） 随机拍卖机制 \mathcal{M}' 在严格意义上是个体理性的，因为每个用户和任务都可以实现非负效用。

此外，可以看出，在 \mathcal{M}' 中每个用户的感知成本等于 \mathcal{M}' 中感知成本的 α' 倍，而每个任务的价值等于 \mathcal{M}' 中任务价值的 β' 倍（其中， α' 、 β' 是 \mathcal{M}' 中的最优解）。 \mathcal{M}' 的效率在这个意义上是有保证的。

命题 5（机制 \mathcal{M}' 的效率） 随机拍卖机制 \mathcal{M}' 保证能够达到分数 VCG 机制 \mathcal{M}' 中总任务价值的 β' 倍和总感知成本的 α' 倍。

5 仿真结果与分析

本节通过仿真结果来评估本文提出的随机拍卖机制的性能，首先验证随机拍卖的性能，然后评估数据复用带来的性能增益。

5.1 仿真设置

在仿真中，将任务数量固定为 $J = 50$ ，将数据项数量固定为 $K = 30$ ，同时，将用户数量从 $I = 10$ 以 10 为增量扩大到 100。每个数据项都是基于位置的（如特定位置的温度），并随机均匀分布在 $1000 \text{ m} \times$

1 000 m 的区域内。每个用户随机移动到时隙中的特定位置，并且可以感知距离其位置 100 m 内的所有数据项。每个用户感知一个数据项的单位成本 β_c 从 [1,5] 中随机选择，用户感知数据项集合 S 的成本是 $\beta_c |S|$ ；每个任务对一个数据项的单位价值 β_v 也从 [1,5] 中随机选择，因此，数据项集合 S 的任务价值是 $\beta_v |S|$ 。

任务相似性（在数据需求方面）的表征如下。将数据项的受欢迎程度定义为任务需要此特定数据项的概率，并将 p_k 表示为所有数据项中第 k 高的受欢迎程度。前期研究表明，互联网上的内容受欢迎程度遵循参数为 $\mu \geq 0$ 的 Zipf 分布，其中， μ 越大，表示任务相似度越高^[20]。

在每次仿真中，选择特定的用户数量 I 和参数 μ ，并进行 1 000 次随机平均来统计结果。

5.2 社会福利差距

首先比较双边 VCG、分数 VCG 和随机拍卖的社会福利，可以帮助理解随机拍卖与最大社会福利（在双边 VCG 拍卖中实现）或分数最大社会福利（在分数 VCG 拍卖中实现）的性能差距。

社会福利随用户数量的变化如图 2 所示，图 2 显示了在不同用户数量下不同拍卖的社会福利，其中，Zipf 分布的 $\mu = 1$ 。从图 2 中可以看出，整数最大社会福利与分数最大社会福利之间的差异可以忽略不计。此外，3 种拍卖机制的社会福利都随着用户数量的增加而增加。随机拍卖机制的社会福利与最大社会福利（基准）的性能差距随着用户数量的增加而增加，在仿真中（当有 100 个用户时）最大差距小于 10%。

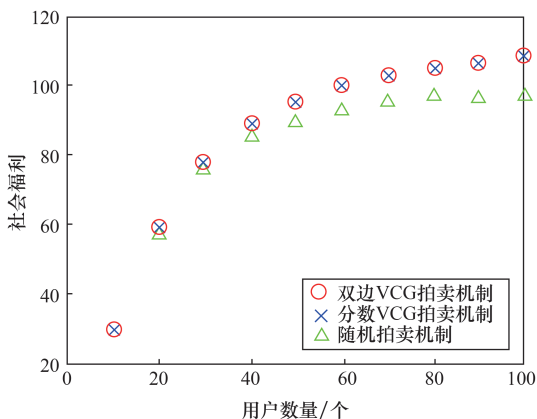


图 2 社会福利随用户数量的变化

5.3 数据复用的性能增益

通过比较系统中有数据复用实现的社会福利

与没有数据复用实现的社会福利来评估数据重用带来的性能提升。在没有数据重用的系统中，用户提供的-一个数据项只能由一个任务使用。因此，需要相同数据的多个任务需要由一个或多个用户提供的同一数据项的多个副本。也就是说，对于每个数据项 $k \in [K]$ ，如果有 N 个任务需要它，那么它应该被至少 N 个用户感知。

社会福利随任务相似度 μ 的变化如图 3 所示，图 3 显示了具有数据复用和不具有数据复用的情形下社会福利在不同 μ 值下的结果，其中，用户数量固定为 $I=60$ 。从图 3 可以看出，在有数据复用的情况下，社会福利随着任务相似度参数 μ 的增加而增加；在没有数据复用的情况下，社会福利随着任务相似度参数 μ 的增加而减小。这是因为在高任务相似度的情况下，大多数任务的数据需求将集中在-小组高流行度的数据上。这意味着，有数据复用的情况下，这些任务可以通过较小的用户集（仅覆盖高流行度数据）来完成，从而潜在地增加了社会福利；在没有数据复用的情况下，这些任务必须由比低任务相似性更小的用户集完成。然而，由于没有数据复用，高任务相似性中的有效用户的数量（即仅能感知高流行度数据的用户）少于低任务相似性中的有效用户的数量（即能够感知任何数据项的用户），因此，系统的社会福利变得更小。

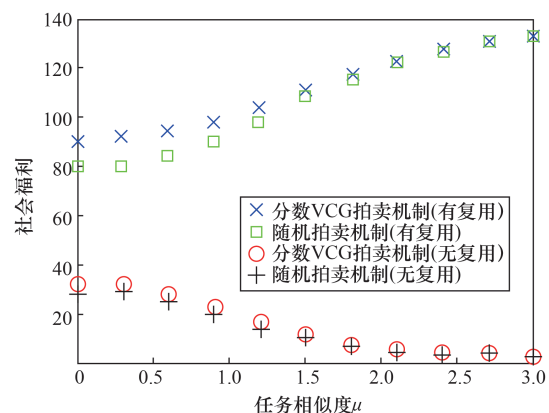


图 3 社会福利随任务相似度 μ 的变化

6 结束语

本文提出了一种新颖的 3 层以数据为中心的群智感知模型，可以实现数据复用，并且可以利用任务的相似性和用户的异构性。本文专注于联合任务选择和用户调度问题，旨在最大化社会福利，这是

一个带有私有信息的 NP-hard 问题。为了解决计算复杂性和私有信息问题，提出了一种双边随机拍卖机制，它具有高计算效率、个体理性和期望激励相容（真实）性。仿真结果证明了本文所提解决方案的有效性。

参考文献：

- [1] GANTI R K, YE F, LEI H. Mobile crowd sensing: current state and future challenges[J]. IEEE Communications Magazine, 2011, 49(11): 32-39.
- [2] DAS T, MOHAN P, PADMANABHAN V N, et al. PRISM: platform for remote sensing using smart phones[C]//The 8th International Conference on Mobile Systems, Applications, and Services, 2010.
- [3] RA M R, LIU B, LA PORTA T F, et al. Medusa: a programming framework for crowd-sensing applications[C]//ACM MobiSys. ACM, 2012.
- [4] DUAN L J, KUBO T, SUGIYAMA K, et al. Incentive mechanisms for smart phone collaboration in data acquisition and distributed computing[C]//IEEE INFOCOM. IEEE, 2012.
- [5] YANG D J, XUE G L, FANG X, et al. Incentive mechanisms for crowdsensing: crowd sourcing with smart phones[J]. IEEE/ACM Transactions on Networking, 2016, 24(3): 1732-1744.
- [6] LUO T, THAM C K. Fairness and social welfare in incentivizing participatory sensing[C]//2012 9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON). IEEE, 2012.
- [7] FENG Z N, ZHU Y M, ZHANG Q, et al. TRAC: truthful auction for location-aware collaborative sensing in mobile crowdsourcing[C]//IEEE INFOCOM. IEEE, 2014.
- [8] ZHANG X, XUE G L, YU R Z, et al. Truthful incentive mechanisms for crowd sourcing[C]//IEEE Conference on Computer Communications. IEEE, 2015: 2830-2838.
- [9] GAO L, HOU F, HUANG J W. Providing long-term participation incentive in participatory sensing[C]//IEEE INFOCOM. IEEE, 2015.
- [10] CHEUNG M H, SOUTHWELL R, HOU F, et al. Distributed time-sensitive task selection in mobile crowdsensing[C]//The 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing. ACM, 2015: 157-166.
- [11] JIN H M, SU L, CHEN D Y, et al. Quality of information aware incentive mechanisms for mobile crowd sensing systems[C]//The 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing. ACM, 2015: 167-176.
- [12] LI Y P, COURCOUBETIS C, DUAN L J. Dynamic routing for social information sharing[J]. IEEE Journal on Selected Areas in Communications, 2017, 35(3): 571-585.
- [13] JIN H M, SU L, NAHRSTEDT K. Centurion: incentivizing multi-requester mobile crowd sensing[C]//IEEE Conference on Computer Communications. IEEE, 2017: 1-9.
- [14] NISAN N, RONEN A. Computationally feasible VCG mechanisms[J]. Journal of Artificial Intelligence Research, 2007(29): 19-47.
- [15] YANG S, HAJEK B. VCG-Kelly mechanisms for allocation of divisible goods: adapting VCG mechanisms to one-dimensional signals[J]. IEEE Journal on Selected Areas in Communications, 2007, 25(6): 1237-1243.
- [16] DOBZINSKI S, DUGHMI S. On the power of randomization in algorithmic mechanism design[C]//2009 50th Annual IEEE Symposium on Foundations of Computer Science. IEEE, 2009.
- [17] LAVI R, SWAMY C. Truthful and near-optimal mechanism design via linear programming[J]. Journal of the ACM, 2011, 58(6).
- [18] ZHANG L Q, LI Z P, WU C. Dynamic resource provisioning in cloud computing: a randomized auction approach[C]//IEEE INFOCOM. IEEE, 2014.
- [19] SMITH D K. Theory of linear and integer programming[J]. Journal of the Operational Research Society, 1987, 38(6).
- [20] BRESLAU L, CAO P, FAN L, et al. Web caching and Zipf-like distributions: evidence and implications[C]//IEEE INFOCOM. IEEE, 1999.

[作者简介]

江常坤（1988—），男，河南信阳人，深圳大学计算机与软件学院助理教授，主要研究方向为移动计算、智能物联网和网络经济学。

高林（1980—），男，浙江金华人，哈尔滨工业大学（深圳）电子与信息工程学院副教授，主要研究方向为无线通信与网络、博弈论和网络经济学。