隐私保护移动人群感知的动态定价：一种强化学习方法

张梦媛、陈继明、雷洋、张军山

Abstract

监控监是一种新兴技术，它利用广泛使用的移动设备的巨大传感能力，以成本效益高的方式完成传感任务。在当前监控监系统的所有未决问题中，对参与者的传感数据缺乏隐私保护的关切最近引起了越来越多的关注。针对用户隐私保护要求不变的静态场景，提出了各种保护隐私的MCS机制。然而，在实践中，用户对隐私保护的要求可能是时变的，这进一步使隐私保护MCS的设计复杂化。在本文中，我们首先概述了保护隐私的MCS的多种有前途的方法，在此基础上，我们第一次尝试在动态场景中探索保护隐私的MCS，这是一个马尔可夫决策过程。具体来说，我们开发了一种基于强化学习的方法，通过该方法，平台可以动态地调整其定价政策，以适应参与用户不同的隐私保护水平。我们进一步使用一个案例研究来评估我们提出的方法的性能。

Introduction

最近，移动人群传感（MC S）已成为各种传感任务（如空气质量监测、民用噪声水平测量和交通拥堵监测[1,2，3]）的一种流行和有效的方法。本着众包的精神，监控监将一系列传感任务外包给配备各种传感器（例如摄像机、麦克风、全球定位系统、陀螺仪和加速度计）的人携带的智能设备。与传统的传感器网络相比，MCS可以更灵活、更经济的方式收集细粒度的信息..

移动用户提供的传感数据可能包含用户的敏感（私人）信息，从而在向不受信任的一方发布时引起隐私关注。这成为设计MCS系统的一个关键挑战，即如何设计有效的激励机制来刺激用户的参与，同时保护用户隐私。传统上，广泛的研究采用经济方法来设计激励计划，以吸引移动用户的参与，如拍卖、博弈论，以及合同理论[4-6]。在这些工作中所做的一个共同假设是，系统参数，如感知实用程序和移动用户的配置文件不会随着时间的推移而变化。

近日，统一人群感应平台设计引起了人们的广泛关注。在这样一个平台中，参与的用户可以同时为不同的应用程序完成多个复杂的感知任务，可能跨越很长一段时间。在完成不同的感知任务时，用户可能有不同的隐私保护要求。例如，用户在为医疗应用程序执行感知任务时可能具有较高的隐私保护级别，而基于位置的应用程序的隐私保护级别较低。此外，他们的隐私保护水平可能会随着时间的推移而变化（例如，人们可能在生病时提高医疗应用的隐私保护水平）。这需要一种新的方法来处理这种动态场景，其中系统参数是时变的。

在本文中，我们首先介绍了一个概述现有的MCS激励机制，旨在满足静态隐私保护要求。然后，我们将重点放在动态场景上，每个参与用户都被授权根据一个自指定和时变的隐私保护级别在局部扰动他们的感知数据。在不知道用户的回报功能和他们的隐私保护水平的变化动态的情况下，平台通过离线方法来确定定价策略是很有挑战性的。为此，我们提出了一种基于学习的方法来自适应地调整感知任务的定价，同时保护移动用户的隐私。具体来说，MCS系统被建模为马尔可夫决策过程，其中每个状态对应于所有用户的当前隐私保护级别。基于观察到的系统状态和感知到的奖励，平台使用基于Q学习的算法动态调整其定价策略以优化其效用[7]。直观地说，将向保护隐私水平低的用户提供更多的付款，以鼓励高质量数据的贡献。值得注意的是，与现有的解决方案相比，我们提出的基于学习的定价机制不需要任何关于移动用户支付功能的知识，并且使用户能够在不同的情况下自主地指定他们的隐私保护要求。

本文的其余部分按

数字对象标识符：
10.1109/MNET.2018.1700468

以下：在下一节中，我们回顾了数据隐私攻击模型以及两种有效的数据扰动方法。然后，我们概述了现有的设计，这些设计将数据隐私保护与激励机制结合起来。接下来，我们研究了用户隐私保护需求时变的MCS场景，并提出了一种基于Q学习的动态定价机制。最后，我们总结了这篇文章。

Data Privacy in MCS

在本文中，我们考虑了一个半可信赖的平台，收集移动用户提供的传感数据。在下面，我们首先介绍贝叶斯攻击模型，旨在推断用户的敏感传感数据。然后，我们介绍了能够有效对抗MCS中贝叶斯隐私攻击的差分隐私机制..

Bayesian Privacy Attack Model

我们假设移动用户具有表示数据值的域的c的私有感知数据xCorinc。我们还假设对手对参与用户的感知结果x偶数c的概率分布p（X）以及感知数据x偶数c被混淆到x偶数c的概率p（x\\\x）具有先验知识。然后，通过对x*的观察，攻击者可以根据Bayes规则[8]导出用户真实感知结果p（x|x*）的后验分布：

$$p(x \mid x^*) = \frac{p(x^* \mid x) \cdot \pi(x)}{\sum_{x' \in \chi} p(x^* \mid x') \cdot \pi(x')}. \quad (1)$$

为了打击这种贝叶斯隐私攻击，我们设计了一种保护隐私的机制，它可以约束攻击者的后验知识比他们先前的知识更好。换句话说，希望p（x\\\x）和p（x\\\x‘）的值足够接近，以便给定x*的观测值，攻击者几乎无法区分x和x‘。为此，我们采用了著名的差别隐私概念。

dA tat dIfferentialAl PrIvAcy

在[6]之后，我们定义了用户的差异隐私保护级别如下。

定义1（差异-价格-优惠级别）：用户的不同隐私保护级别（DPL），以e表示，定义为：

$$\varepsilon = \max\left\{ \frac{p(x^* \mid x)}{p(x^* \mid x')}, \frac{p(x^* \mid x')}{p(x^* \mid x)} \right\}, \forall x, x' \in \chi. \quad (2)$$

用户的DPL量化了通过测量报告数据的条件概率之间的不可区分性，在攻击者的先验知识下的隐私。可以很容易地证明，在DPL=e的情况下，攻击者的知识增益g=p（x|x*）/p（x），有界为1/e e≤g≤e，具有先验知识p（x）。显然，e越小，攻击者就越难正确推断出真实数据，因此隐私保护性能越好。接下来，我们将回顾两种局部数据扰动方法，通过这些方法可以为用户的数据达到一定的差异隐私保护水平。

拉普拉斯机制：拉普拉斯机制是广泛使用的提供差异性的技术

反向拍卖机制是以平台为中心的MCS激励机制设计中最流行的方法之一。具体而言，平台作为拍卖商和移动用户

向平台报告其投标，反映其参与成本。

隐私保障[9]。在本文中，我们特别感兴趣的是Laplace机制的一个变体版本，其中Laplace机制的Laplacian噪声h-Lap（D/e）被本地添加到每个用户的数据中，其中D’sx被定义为局部灵敏度，e是该用户的DPL。

随机响应：随机响应是另一个可以提供本地差异隐私保证的工具[9]。在一定概率下，单个用户将其真实数据的随机实例发送给不可信的数据收集器。随机响应机制的参数被仔细选择，以限制平台对数据的真实价值充满信心地学习的能力。

NoIs Inject IonbAsed

Privacy-Preserving Approaches in MCS

在本节中，我们概述了几种最先进的保护隐私的MCS设计，将差分-私人噪声注入与激励机制设计结合起来。我们根据他们使用的激励机制的类型对这些作品进行分类，包括基于拍卖的方法、基于博弈论的方法和基于契约理论的方法。

Auction Based Approach

反向拍卖机制是以平台为中心的MCS激励机制设计中最流行的方法之一。具体来说，平台作为拍卖商，移动用户向平台报告他们的出价，反映他们的参与成本。平台选择参与者并确定相应的支付，目的是在预算约束下最小化总支付或最大化平台效用，或最大化参与者的社会福利[10]。

在Ghosh和Roth[11]的研讨会工作中，提出了一种拍卖机制，其中私人数据被视为由运行调查的数据分析器采购的货物。设计的拍卖机制满足了激励机制设计的两个基本要求，即真实性和个人理性：

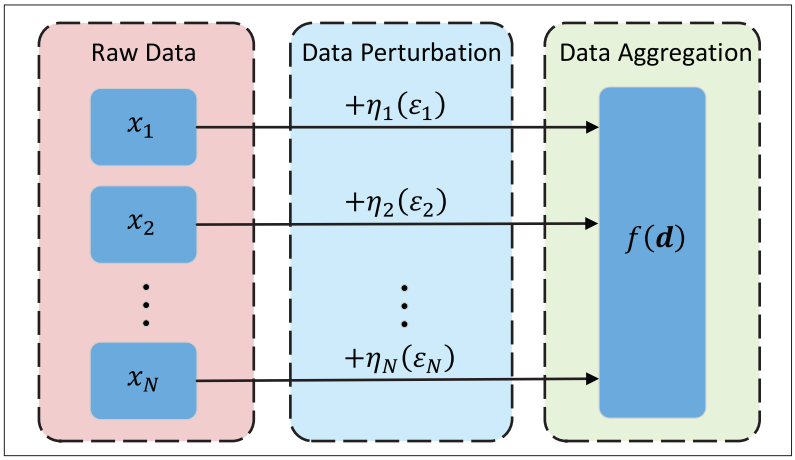
- 真实性：参与的用户不能从不真实的投标中获益。
- 个人合理性：每个参与用户都收到大于或等于其隐私成本的付款。

这项工作[11]是权衡的特点在总付款和结果准确性之间，从以下两个角度：·在给定准确性要求的情况下，尽量减少总付款。
最大限度地提高付款预算下的结果准确性。

沿着这条线，金等人。开发了一个组合-

参考资料	机制类型	平台可信	数据摄动	DPL	特色
[4]	拍卖	是	拉普拉斯机制	静态	考虑用户的可靠性
[5]	拍卖	是	拉普拉斯机制	静态	考虑用户之间的社会关系
[6]	Game	否	随机响应	静态	分析用户的均衡行为
[14]	合同	否	拉普拉斯机制	静态	克服信息不对称
本文	–	否	拉普拉斯机制	Dynamic	应用基于学习的在线定价

表1.现有的移动人群感知数据隐私保护方法综述。



基于拍卖的MCS数据隐私保护框架[4]。具体来说，隐私成本被建模为每个用户感知成本的一部分。此外，在用户选择过程中还进一步考虑了移动用户的可靠性，这影响了聚合感知结果的准确性。相关作品[5]被认为是移动用户嵌入社交网络的设置。值得注意的是，两者在模型中都假设了一个可信的数据收集器。

Game Theory Based Approach

博弈论模型是激励机制设计的另一种流行方法[10，12，13]。与基于拍卖的机制设计类似，游戏理论方法需要指定一种激励用户参与的支付政策。然而，在基于游戏的方法中，参与的用户不是由平台选择的；相反，鉴于平台的支付政策，用户做出战略参与决策。

王等人在[6]设计了一个游戏理论机制中的设置，其中数据收集器购买用户的私有数据，这表示他们对底层系统状态的知识。与[4，5，11]信任平台对聚合数据进行集中数据扰动不同，[6]中的一个基本假设是数据收集器不可信，在这种情况下，允许每个参与用户在向数据收集器发布之前对其原始数据进行战略性的局部扰动。在他们的博弈论模型中，个体用户是游戏的参与者，其行为与他们的数据扰动策略相对应。定价策略由数据采集器精心设计，使参与用户的隐私成本得到补偿，并满足所采集数据的准确性要求，从而达到博弈的纳什均衡。

Contract Based Approach

在基于合同的机制中，合同设计者将设计一个复杂的努力支付对菜单，以激励不同类型的个人参与，同时优化其自身的收益。基于合同的方法不需要实时的投标信息，而是利用个人成本的统计信息来确定支付合同，从而克服了信息不对称，减少了通信和计算开销。在[14]中，提出了一种基于合同的隐私保护MCS框架..具体来说，平台设计和广播一个合同菜单，每个合同菜单指定一种不同的隐私保护级别，以及用户如果同意该合同将收到的相应付款。然后，每个用户选择一个使其效用最大化的合同。作者用适当的度量方法导出了个体隐私与聚合精度之间的定量关系。

以上审查的三种方法是：表1概述。虽然他们应用了不同的机制设计方法，但它们是一个共同的线程：在整个人群感知过程中，每个用户的DPL保持不变。在下一节中，我们将当前场景扩展到更一般的场景，在这种情况下，移动用户的DPL不是静态的，平台需要动态调整定价策略以优化MCS系统的操作。

Learning Based Dynamic Pricing for Privacy-Preserving MCS

在这一部分中，我们开发了一种用于MCS系统中动态定价问题的强化学习方法。在基于博弈的隐私保护MCS方法中，纳什均衡的计算需要了解用户的收益函数和DPL，这是很有挑战性的。对于基于合同的方法，平台至少需要用户隐私成本的统计信息来进行合同设计，这可能是不可用的。在以平台为中心的拍卖方法中，用户几乎没有权力决定他们的参与状态。这里提出的基于学习的方法可以用来处理信息

平台与移动用户之间的不对称问题。此外，它还可以应用于用户的DPL是时变的。

Problem Statement

我们考虑一个由一个半可信任的平台和一个N个移动用户的N个庄{1，2，...，N}组成的MCS。我们假设用户面临着先前引入的贝叶斯数据推理攻击，并且具有不同的差分隐私级别（DPL）。我们让eRe表示不参与感知任务的选择，并让离散集E={e（1），.，e（L）}是参与用户DPL的可能值集，e（L）>.>e（1）>0。我们使用eit来表示移动用户i在t个子{0，1，2，...}阶段所做的决定。在实践中，每个用户我首先决定是否参与。每个参与用户将在一个可用的选项列表中指定一个语义隐私偏好级别，例如低、中、高或非常高，这些选项中的每一个对应于集合E中提供的DPL值。根据定义1，EI越小，隐私保护级别越高。

我们假设我拥有的每个用户-ING数据，表示为xiErinc，但只显示一个局部扰动数据，表示为~xiErinc，到半信任平台。向量d=[~...xS]表示平台收集的所有数据，r=f（D）是通过实值函数f得到的聚合结果，局部数据扰动和数据聚合过程如图所示。1.

我们定义了nIt=SN $\sum_{i=1}^{N}\mathbb{1}\{\varepsilon_i^t=\varepsilon_{(l)}\},\ l\in\{1,\cdots,L\}$ as保护隐私的用户数量水平等于e在t阶段，并定义向量NT=（n1t，..!），NL作为移动人群的整体隐私保护级别配置文件。我们进一步表示pt=（p1t，...，pL t）作为平台指定的定价配置文件，其中每个元素pI t≥0表示用DPL=e（L）支付给用户。我们让R（NT）表示平台在t阶段感知到的奖励，这是整体隐私保护级别配置文件NT的函数。平台在t阶段的效用被定义为奖励和总支付之间的差额，即U0=R（N T）-ptnt。最后，我们将每个用户的个人收益定义为UI=pIt-e（L）ci是用户i在t阶段的单位隐私成本。

Q-Learning Based Dynamic Pricing

在我们的MCS的每个阶段，如果货币支付不能补偿他们的隐私损失，即最大的e，pt）<0，用户I将退出参与。否则，用户i参与并选择使其收益最大化的DPL，即EI**********接下来，用户I向平台报告他们的DPL以及噪声传感数据，这些数据以差分私有的方式被他们的DPL参数化。值得注意的是，虽然平台知道每个用户的DPL，但他们无法以高度的信心恢复每个用户的原始数据，从而实现了对用户数据隐私的有效保护。

观察移动用户单位隐私成本{ci}fCorton是时变的，不为平台和其他移动用户所知，我们采用了一种基于学习的方法，该方法是设计为

应对参与用户的隐私保护要求可能面临的挑战

我们提出了一种基于Q学习的定价机制，通过该机制，MCS平台可以动态地调整其定价政策以优化其效用。

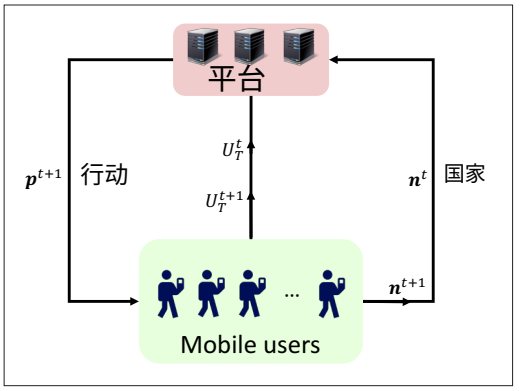


图2.《公约》原则的说明

基于Q学习的动态定价机制。

使平台能够通过试错了解最优定价策略。最近，肖等人将MCS平台与自私移动用户之间的交互作为一种动态游戏，并将平台的定价过程建模为马尔可夫决策过程[15]。在[15]的启发下，我们将MCS平台的定价过程建模为马尔可夫决策过程（MD P）。与[15]不同，我们没有从博弈论的角度来解决定价问题。在我们的模型中，MDP的状态空间对应于整体的隐私保护级别配置文件，即{n}偶核RL。我们设计了一种基于Q学习的求解MD P的算法[7]..具体来说，我们让Q（nt，pt）表示国家和定价政策在t阶段的Q功能。平台按照以下程序更新Q功能：

$$\begin{cases} Q(\mathbf{n}^t,\mathbf{p}^t) \\ \quad =(1-\alpha)Q(\mathbf{n}^t,\mathbf{p}^t)+\alpha(u_0^t(\mathbf{n}^t,\mathbf{p}^t)+\eta V(\mathbf{n}^{t+1})), \\ V(\mathbf{n}^t)=\max_{\mathbf{p}} Q(\mathbf{n}^t,\mathbf{p}\ ). \end{cases} \quad (3)$$

其中\64257；（0，1）是学习速率，h偶[0，1]是表示服务器近视性质的折扣因子，V（·）是状态的最高值。学习过程的原理如图所示。2.假设平台在每个阶段应用s-greedy策略来确定定价策略p..具体来说，最优价格p=argmaxp Q（nt1，p）将以概率1-s选择，其他支付策略则以相对较小的概率均匀地选择。算法1总结了基于Q学习的平台动态定价机制..

Case Studies

在这一部分中，我们通过仿真演示了基于Q学习的动态定价机制的性能..我们将系统中的默认代理数设置为N=200，四种可能的DPLs为e（1）=0.1（非常高的级别），e（2）=0.2（高级别），e（3）=0.3（中等级别），e（4）=0.4（低级别）。

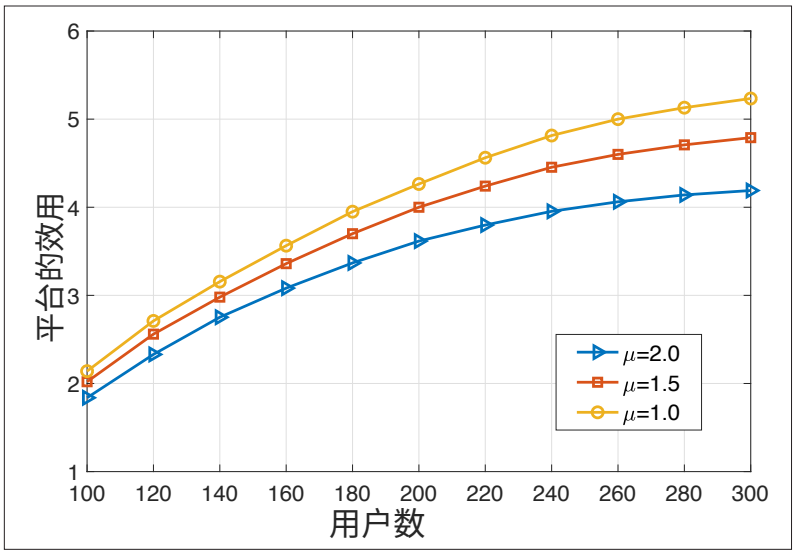
1: 初始化：
2: Set $\alpha \in (0, 1]$, $\eta \in [0, 1]$, $Q(\mathbf{n}, \mathbf{p}) = 0$, $V(\mathbf{n}) = 0$, $\forall \mathbf{n}, \mathbf{p}$.
3: 观察初始系统状态n0；
4: 循环为t = 1, 2, 3, ...
5: 选择定价策略pt=argmaxp Q（nt+1， p）使用概率1-s；否则，随机选择一个具有概率s/|P-1的定价策略pt BriendP。
6: 通过计算nt=（n）观察系统状态 t], ..., n_t^t).
7: 计算即时效用u0 t;
8: 根据Eq更新Q（NT，pt）。
9: 根据Eq更新V（NT）。3；
10: 结束循环

ALGO RITHM1.基于Q学习的动态定价监控监
。

根据[10]中的方程（2.2），我们定义了平台的奖励函数为R（N T）=Mlog（a · nt1）。日志函数描述了平台对用户参与的回报递减，并且缩放因子M是由平台定制的系统参数。权向量a=（a1，a2，a3，a4）量化了不同隐私保护级别的用户感知数据的贡献，并在我们的案例研究中设置为a=（2，1.5,1，0.5）。为了便于阐述，我们根据id正态分布，即ci-N（m，0.1），在单个阶段生成每个用户的单位隐私成本。在经验上，我们将学习算法的参数设置为

$\alpha = 0.2$, $\eta = 0.8$, and $\sigma = 0.2$.

我们对不同水平的用户的平均单位隐私成本（即m=1.0、1.5、2.0）。图3显示了在用户平均单位隐私成本m的不同值下，随着迭代次数的增加，平台效用的变化趋势。可以看出，我们的算法在300次迭代后很快收敛。结果还表明，对于较小的单位隐私成本，平台的可实现效用较大。在图中我们说明了用户总数对平台效用的影响。具体来说，随着用户数量从100增加到300，对于所有三种不同平均单位隐私成本的情况，可实现的效用单调增加。In addi-



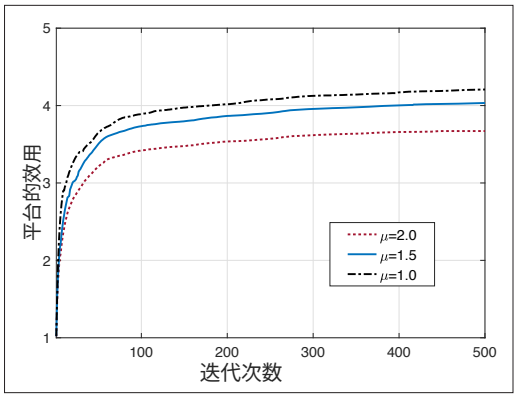


图3.说明学习算法的收敛性能。

可以观察到，效用的边际增加随着用户总数的增加而减小。这是由于我们在案例研究中使用的奖励函数的凹性，这也同意了在收集了足够数量的传感数据后，新收集的数据带来的值将收缩的事实..

Conclusion

在本文中，我们概述了几种最先进的移动人群感知隐私保护方法，这些方法刺激了具有隐私关注的移动用户的参与，以释放他们的敏感数据。为了解决参与用户的隐私保护需求可能是时变的挑战，我们提出了一种基于Q学习的定价机制，通过该机制，MCS平台可以动态地调整其定价策略，以优化其效用。进一步提供了一个案例研究，以评估我们提出的方法的性能。

承认门特

这项工作得到了CNS-1559696和IIA-1301726赠款项下美国国家科学基金会和自然科学基金会61429301赠款项下的部分支持。

参考文献

[1]“ Waze：全球定位系统的自由导航与轮流，”http://www.。waze.com/.
[2]S.Hu等人，“ Smartroad：Smartphone-based crowd Sensing ”用于交通调节器检测和识别，“ ACMTrans。传感器网络，第二卷。11，没有。4，2015年，第55页。
[3]Y.Cheng等人，“ Aircloud：一个面向每个人的基于云的空气质量监测系统”，Proc。ACM SenSys2014，pp.[4]H.Jin等人，“ Incention：激励隐私保护 ”移动人群感知系统的数据聚合，“ Proc。ACM MobiHoc2016，vol.16，pp.341 – 50.
[5]M.Zhang等人，“ 隐私保护人群感应：隐私 ”估值、网络效应和利润最大化，“ Proc。IEEEGlobecom2016，pp.1 – 6.
[6]W.Wang等人，“ 隐私的价值：战略数据主体、激励机制和基本界限 ”，Proc。ACMSigmetrics2016，pp.249 – 60.
[7]R.S.Sutton和A.G.Barto，强化学习：An导言，麻省理工学院出版社，1998年。
[8]L.Wang等人，《稀疏的差分位置隐私》移动Crowdsensing，“ Proc。IEEE ICDM2016，2016年12月，pp. 1257–62.
[9]C.Dwork等人，”微分的算法基础隐私，“ 理论计算机科学的基础和趋势，第二卷。9，没有。第3-4页，2014年。211 – 407.
[10]D.Yang等人，”Crowdsensing的激励机制：众包与智能手机，“ IEEE/ACM Trans。联网，第二卷。24，没有。3，2016年，pp.1732 – 44.
[11]A.Ghosh和A.Roth，《在拍卖中出售隐私》，Games

和经济行为，第二卷。91，2015年，pp.334 – 46.
[12]S.He等人，“对移动的交易所市场办法”
众包：定价、任务分配和Walrasian均衡，“IEEEJSAC，2017年。
[13]G.Yang等人，“通过移动人群感知的社会激励机制促进合作”，IEEE Commun。Mag.，vol.55，没有。3，2017年，pp.86 – 92.
[14]Z.Zhang等人，“REAP：一种有效的激励机制”
对于在Crowdsensing中协调聚合精度和个人隐私，“Corr，第二卷。abs/1711.00757，2017年；提供：http://arxiv.org/abs/1711.00757
[15]L.Xiao等人，“Secure Mobile Crowdsensing Game”，Proc。
IEEE GLOBECOM 2015, 2015, pp. 7157–62.

Biographies

张梦媛2011年获得中国杭州浙江大学光学科学与工程专业学士学位，2012年获得澳大利亚悉尼新南威尔士大学光电与可再生能源工程学院硕士学位。他目前正在攻读博士学位。浙江大学控制科学与工程系D学位。他目前的研究兴趣包括移动社交网络和众包网络中的安全和隐私。

陈继明[M’08，SM’11]收到了B.Sc。和博士2000年和2005年分别获得浙江大学控制科学和工程学位。他是滑铁卢大学2008年至2010年的访问研究员。现任控制科学与工程学院正教授，中国浙江大学工业控制技术国家重点实验室和工业过程控制研究所副所长。他目前的研究兴趣包括网络安全、传感器网络、智能电网和网络控制。他曾担任多种国际期刊的副编辑，包括并行和分布式系统上的IEEE事务、IEEE网络和IEEE网络系统控制事务。他是IEEE自动控制交易的客座编辑。雷洋[M13]分别于2005年和2008年获得中国南京东南大学电气工程学士和硕士学位，并获得博士学位。2012年获得美国亚利桑那州坦佩市亚利桑那州立大学电气计算机和能源工程学院D学位。他是美国新泽西州普林斯顿普林斯顿大学的博士后学者，也是亚利桑那州立大学电气、计算机和能源工程学院的助理研究教授。他目前是美国内华达州雷诺大学计算机科学与工程系的助理教授。他是IEEE IN FOCOM最佳论文奖亚军

2014.

张俊山[M‘00，SM’06，F’12]获得博士学位。D.2000年普渡大学欧洲经委会学院学位。他于2000年8月加入亚利桑那州立大学欧洲经委会学院，自2015年以来一直担任富尔顿教授。他的研究方向是信息网络和数据科学的一般领域，包括通信网络、物联网（物联网）、雾计算、社交网络和智能电网。他是IEEE的研究员，2005年获得ON R青年调查员奖，2003年获得NSF职业奖。他在2016年获得IEEE无线通信技术委员会表彰奖。他的论文获得了几个奖项，包括KennethC.Sevcik优秀学生论文奖，ACMS IGME TRICS/IFIP表现2016，IEEE IN FOCOM2009和IEEE IN FOCOM2014最佳论文亚军奖，IEEEICC2008和ICC2017最佳论文奖。他是TPC在通信网络中的一些主要会议的共同主席，包括IEEE IN FOCOM2012和ACMMOBIHOC2015。他是ACM/IEEESEC2017、Wiopt2016和IEEE通信理论研讨会2007的总主席。他是IEEE通信学会的杰出讲师。他是IEEE无线通信交易的副编辑，计算机网络杂志的编辑，IEEE无线通信杂志的编辑。他目前担任IEEE/ACM网络交易的总编辑和IEEE网络的编辑。