基于区块链的群智感知中任务预算约束的位置隐私保护参与者选择方法

高 胜[1] 陈秀华[1] 朱建明[1] 袁丽萍[1] 马鑫迪[2] 林 晖[3]

[1](中央财经大学信息学院 北京 100081)

[2](西安电子科技大学网络与信息安全学院 西安 710071)

[3](福建师范大学计算机与网络空间安全学院 福州 350117)

摘 要 群智感知中,任务发布者基于有限的任务预算招募合适的参与者来执行感知任务.但是,现有的相关工作依赖于可信第三方来执行参与者选择或者忽视了参与者位置隐私泄露问题.为了解决这些问题,本文提出一种基于区块链的群智感知中任务预算约束的位置隐私保护参与者选择方法LPWS.通过保序加密和Merkle树来为参与者提供个性化的位置隐私保护,确保参与者将精确位置隐藏于隐匿区域.在有限的任务预算下,LPWS将参与者选择问题建模为目标优化问题,并基于动态规划来确定一组合适的参与者以增加高质量感知数据获取的可能性.此外,在保证数据隐私和奖惩公平性下,LPWS基于数据质量评估结果完成报酬支付和信誉更新,从而激励参与者尽可能地提供高质量数据.仿真实验表明,LPWS在参与者选择方面具有可行性与有效性,保证了安全公平的参与者选择以及数据质量评估.与相关工作对比,在有限的任务预算下,LPWS不仅取得了更好的质价比,而且在确保任务完成质量的同时提供了位置隐私保护和数据隐私保护.

关键词 群智感知;区块链;位置隐私;参与者选择;质量评估

中图法分类号 TP309 **DOI号** 10.11897/SP.J.1016.2022.01052

A Location Privacy-Preserving Worker Selection Scheme Under Limited Budget for Blockchain-Based Crowdsensing

GAO Sheng[1] CHEN Xiu-Hua[1] ZHU Jian-Ming[1] YUAN Li-Ping[1] MA Xin-Di[2] LIN Hui[3]

[1](*School of Information, Central University of Finance and Economics, Beijing* 100081)

[2](*School of Cyber Engineering, Xidian University, Xi'an* 710071)

[3](*College of Computer and Cyber Security, Fujian Normal University, Fuzhou* 350117)

Abstract In crowdsensing, task publishers recruit a set of suitable workers based on a limited task budget to perform sensing tasks. However, some work relies on a trusted third party to perform worker selection or ignores the problem of location privacy leakage. To tackle these issues, a Location Privacy-preserving Worker Selection scheme under a limited budget for blockchain-based crowdsensing (LPWS) was proposed. Based on order-preserving encryption and Merkle tree, workers can mask precise locations in hidden areas to achieve personalized location privacy protection. With a limited task budget, LPWS models worker selection as a target optimization problem, and then select a set of suitable workers based on the dynamic programming to increase the possibility of obtaining high-quality sensing data. In addition, for guaranteeing the data privacy and the

收稿日期:2020-12-31;在线发布日期:2021-07-13.本课题得到国家自然科学基金(62072487,61902290)、北京市自然科学基金面上项目(M21036)、全国统计科学研究重大项目(2020LD01)资助.**高 胜**,博士,副教授,中国计算机学会(CCF)会员,主要研究方向为区块链技术与应用、数据安全与隐私计算.E-mail:sgao@cufe.edu.cn.**陈秀华**(通信作者),硕士研究生,主要研究方向为群智感知、区块链.E-mail:xiuhuachen1003@163.com.**朱建明**(通信作者),博士,教授,主要研究领域为数字经济、金融信息安全、区块链.E-mail:zjm@cufe.edu.cn.**袁丽萍**,硕士研究生,主要研究方向为联邦学习、区块链.**马鑫迪**,博士,讲师,主要研究方向为数据安全、隐私保护.**林 晖**,博士,教授,主要研究领域为区块链、机器学习、移动边缘计算.

fairness of rewards and punishments, LPWS completes remuneration payment and reputation update based on the results of data quality evaluation, thereby encouraging workers to provide high-quality sensing data as much as possible. Simulation results indicate that LPWS is feasible and effective in worker selection, ensuring safe and fair worker selection and data quality evaluation. Compared with related work, under the limited task budget, LPWS not only achieves a better quality-price ratio, but also provides location privacy protection and data privacy protection while ensuring the task completion quality.

Keywords　crowdsensing; blockchain; location privacy; worker selection; quality evaluation

1 引　言

群智感知作为一种物联网的新兴应用，其旨在利用群体的智慧来完成任务，有效解决了传感器网络所存在的高维护费、移动性差等问题[1-2]. 传统的群智感知包含三个实体，分别为任务发布者、参与者以及中心化感知平台[3]. 任务发布者发布感知任务到感知平台，然后感知平台进行参与者选择以招募合适的参与者来收集感知数据. 通过贡献有效的感知数据，参与者获得来自任务发布者的报酬奖励. 然而，传统的群智感知存在单点故障、操作不透明等威胁风险，从而降低了参与者的参与积极性[4]. 最近，区块链技术的发展促进了群智感知的可持续发展，基于区块链的群智感知系统[5-7]能够有效克服中心化感知平台的弱点.

群智感知中，任务发布者通常基于有限的任务预算发布感知任务，而参与者往往需要通过位置移动来收集感知数据. 这种类型的数据收集模式已在许多大规模的实际应用中使用，例如环境质量监测[8]、交通流量监测[9]等. 然而，由于群智感知的开放性，任何持有感知设备的参与者都能参与感知任务，但其所提交的感知数据并非都满足任务需求[10-11]. 在参与者选择方面，通常一些高信誉或者近距离的参与者被视为具备较高的可能性来提供高质量数据. 信誉值反映了参与者过去执行任务的表现，而位置距离则体现了参与者当前执行任务的条件. 在基于位置的感知任务中，距离任务位置越远意味着参与者需要消耗更多的时间与精力到达指定任务位置，这可能导致参与者无法拥有充足的时间来收集感知数据，进而导致感知数据的质量降低. 因此，位置因素是参与者选择所需考虑的影响因素之一. 然而，提交位置信息的行为可能导致位置隐私泄露，而位置隐私的泄露将导致个人敏感隐私的泄露[12-13]，如家庭地址、个人偏好、生活习惯等. 相关工作[14-17]虽然综合考虑了信誉与位置距离来执行参与者选择以最大化感知数据质量，但是忽视了参与者的位置隐私保护需求.

在参与者选择阶段，主要目标是综合考虑信誉水平和位置距离来选择最有可能提供高质量数据的参与者. 为了激励参与者在实际数据收集过程中能够积极地完成任务，任务发布者需要支付报酬奖励来激励参与者[18-19]. 基于数据质量评估，符合要求的参与者将获得报酬奖励. 然而，数据质量评估需要考虑感知数据的安全性以及质量评估的公平性. 为了防止恶意参与者通过复制其他诚实参与者的感知数据以进行数据提交，需要保证感知数据的安全性. 此外，感知数据的评估结果将影响参与者的应获报酬和信誉更新，进而影响参与者的参与积极性，因此需要保证质量评估的公平性. 但是，现有的相关工作[17,20-22]很少考虑数据质量评估或者忽视数据质量评估中的感知数据安全性.

为了解决上述问题，本文提出一种基于区块链的群智感知中任务预算约束的位置隐私保护参与者选择方法 LPWS. 在有限的任务预算下，选择一组合适的参与者来最大化高质量数据的获取可能性，并且保证参与者的位置隐私安全. 进一步，在保证数据隐私性与评估公平性下，对符合质量要求的参与者进行报酬奖励和信誉奖励，进而激励参与者积极提供高质量的感知数据. 本文的主要贡献包括 3 个方面：

(1) 在基于区块链的群智感知模型下，提出一种具有位置隐私保护的参与者选择方法 LPWS，以缓解传统的群智感知所面临的中心化、透明性差等威胁风险，同时采用保序加密和 Merkle 树实现参与者个性化位置隐私保护. 考虑有限的任务预算，LPWS

将位置距离和信誉水平作为参与者选择指标，并基于动态规划，然后将参与者选择问题建模为目标优化问题，最后给出了基于动态规划的求解方法.

(2)为了提高任务完成质量，LPWS 基于数据质量评估结果完成参与者的报酬支付和信誉更新，从而激励参与者尽可能地提供高质量数据. 通过采用随机扰动以及 Pedersen 承诺，实现具有隐私性和公平性的数据质量评估.

(3)通过仿真实验表明，LPWS 在参与者选择方面具有可行性与有效性. 对比于相关工作，LPWS 在保证任务完成质量下，不仅实现了安全公平的参与者选择，而且保证了质量评估的隐私性与公平性以及取得了更好的质价比.

2 相关工作

为了保证感知数据质量，参与者选择是群智感知中不可忽视的重要环节. 参与者选择旨在借助参与者的相关属性信息来预估参与者提供高质量数据的可能性，进而选择合适的参与者来执行感知任务.

在群智感知中，信誉和位置距离通常被用作参与者选择的筛选指标. 现有一些工作借助中心化感知平台来完成参与者选择. 例如，文献[23]提出基于统计和投票的两种信誉评分方法，并基于信誉完成参与者选择，忽略了任务预算有限的约束条件. 文献[14]提出一种基于启发式的贪婪方法以实现最大化的任务分配，但其考虑参与者以自愿无偿的形式参与任务. 考虑到有限的任务预算，文献[15]则提出一种基于预算的感知任务分配方法，其考虑参与者的信誉以及位置距离，以在不超过预算下最大限度地提高感知结果的预期质量. 然而，中心化感知平台拥有对参与者选择、信誉更新和报酬支付等方面的绝对控制权，故其容易遭受攻击与破坏. 为了解决中心化带来的安全问题，相关工作则研究基于区块链的群智感知中的参与者选择. 例如，文献[16]基于信誉和位置计算参与者与任务之间的匹配度，进而选择高匹配度的参与者来执行任务. 链上节点负责选择参与者，并基于任务发布者上传的数据质量评估结果完成报酬支付. 文献[17]综合考虑位置、信誉以及时间来计算参与者的任务质量信息，以在人数限制下选择一组具有最大任务质量信息的参与者，并通过公开的数据质量评估以计算参与者的报酬奖励. 文献[24]提出一种声誉机制和仲裁机制，并将其分别应用于参与者选择和数据评估者选择. 但是，其需要提前选择一组可靠的数据评估者来执行数据质量评估. 上述方案并未考虑位置隐私泄露问题或者数据质量评估中的数据隐私安全.

考虑到位置隐私泄露问题，一些工作结合位置混淆、差分隐私来设计具有位置隐私保护的群智感知方案. 例如，文献[21]允许参与者提交感兴趣的任务区域以代替精确的位置，从而基于任务发布者分配的报酬奖励以及位置距离来选择任务. 文献[25]在保证任务覆盖质量与位置隐私下，将参与者选择问题定义为任务覆盖率最优化问题，其同样允许参与者提交包含真实位置的混淆区域信息. 然而，上述方案没有针对被选中参与者提交数据后的情况，描述关于数据质量评估的细节，从而无法实际反馈任务完成情况. 文献[26]提出一种适用于群智感知的具有位置差分隐私的参与者选择方案. 参与者采用地理位置不可区分来生成混淆感知区域，进而服务提供者在有限的预算下求解任务覆盖范围最大化的参与者选择问题. 该方案虽然提供了位置隐私保护，但是参与者选择结果受控制于服务提供者.

考虑到数据隐私安全问题，相关工作采用 k-匿名、秘密共享等技术来保证数据隐私安全. 例如，文献[27]提出一种基于 k-匿名的位置及数据隐私保护方法，其通过多方安全协助的方式来构造等价类以实现 k-匿名位置隐私保护，而等价类中的参与者负责转发数据并通过抛硬币的方式来决定是否上传自己的感知数据，进而实现数据隐私保护下的数据迭代上传. 文献[28]同样基于数据转发的思想，允许每个参与者将自己的加密感知数据以及参与者链上对应子节点发送的数据一起上传给父节点. 虽然他们考虑了数据隐私问题，但是多次的数据转发将造成一定的时间开销. 文献[29]则采用秘密共享来保护数据隐私安全. 参与者把来自任务发布者发送的密钥份额作为盲因子来扰动感知数据，并上传至区块链供节点计算扰动后的数据聚合结果. 任务发布者则使用私钥来消除盲因子，以获得原始数据聚合结果. 但是，其需要一个可信的密钥分配中心来提供盲因子. 文献[30]设计了一种个性化隐私度量算法来获得参与者感知数据的隐私度，并基于隐私度设计博弈游戏来为参与者提供数据上传的最佳决策，进而实现数据隐私保护. 然而，

其依赖感知平台来为参与者提供动态的个性化隐私度量结果.

在群智感知中，对于参与者而言，其希望在不泄露位置隐私和数据隐私下参与感知任务并在公平的数据质量评估中获得报酬奖励. 对于任务发布者而言，其希望在有限的任务预算下选择合适的参与者以获得高质量的感知数据. 为了满足参与者与任务发布者的需求，本文借助区块链提出一种任务预算约束的位置隐私保护参与者选择方法 LPWS. 在保证任务完成质量下，LPWS 致力于实现安全公平的参与者选择以及数据质量评估，进而提高参与者与任务发布者的参与积极性.

3　问题定义与系统模型

本节将定义参与者选择问题并且概述相应的系统模型. 为了更好地理解本文所提的参与者选择方法，表 1 列出了常用的符号和相应的描述.

表 1　符号和描述

符号	描述
w_i	第 i 个参与者
$lw_i(x_{l_i},y_{l_i})$	参与者 w_i 在 l_i 时刻的位置
$task_j$	第 j 个感知任务
obj_m	感知任务所包含的第 m 个观测对象
L	最大矩形隐匿区域
H	区域划分水平
$xMTree$	隐匿区域边界点的横坐标 Merkle 树
$yMTree$	隐匿区域边界点的纵坐标 Merkle 树
fd	感知位置距离上限
$subL$	L 的子隐匿区域
(x_o,y_o)	$subL$ 的外接圆圆心
r^*	$subL$ 的外接圆半径
QV_{w_i}	参与者 w_i 的任务质量价值
rep_{w_i}	参与者 w_i 的信誉值
dl_{w_i}	参与者 w_i 与任务位置的距离
rew_{λ_u}	不同距离间隔 λ_u 内的任务基础报酬
B_1	参与者选择阶段使用的预算金额
B_2	报酬支付阶段使用的预算金额
N	参与者总人数
agg_m	观测对象 obj_m 的数据聚合结果
$dist_i$	参与者 w_i 的数据失真程度

3.1　问题定义

在群智感知中，基于有限的任务预算，任务发布者希望选择一组合适的参与者来完成数据采集工作以最大化感知数据质量. 针对感知任务 $task_j$，在参与者未执行任务之前，本文考虑位置距离以及信誉值对任务质量价值的影响. 一方面，参与者距离任务位置越近，意味着其在任务截止时间之前拥有充足的时间去执行感知任务. 另一方面，参与者的信誉值越高，则表示参与者过去执行感知任务的表现越好. 因此，综合考虑位置距离与信誉值来计算任务质量价值以作为参与者选择阶段的评估指标，这是合理的. 每个提交任务请求的参与者 w_i 则需要提交相应的位置信息 $lw_i(x_{l_i},y_{l_i})$ 以用来计算任务质量价值. 然而，位置信息属于参与者的敏感隐私，即使参与任务会获得报酬奖励，参与者往往也不愿意以暴露位置隐私的代价来参与任务. 此外，在数据质量评估阶段，由于感知数据是参与者获得报酬奖励的凭证，存在恶意参与者通过复制他人的数据来获取不应得的报酬奖励的行为，因此需要确保感知数据的安全性. 为了保证质量达标的参与者能够获得应有的报酬奖励，还需要确保数据质量评估的公平性. 因此，在有限的任务预算下，本文设计了一种考虑位置隐私与数据隐私的参与者选择方法以满足最大化感知数据质量的需求.

3.2　系统模型

LPWS 的系统模型由 3 个实体组成，包括任务发布者、参与者以及区块链，如图 1 所示.

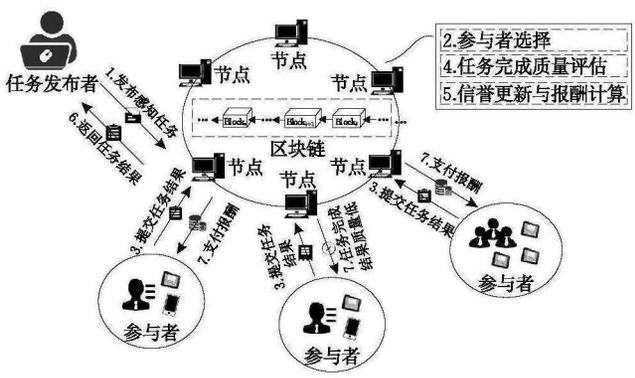


图 1　系统模型

(1) 任务发布者. 任务发布者基于有限的任务预算，发布感知任务到区块链，其希望招募一组合适的参与者来完成感知任务以保证感知数据质量的最大化. 因此，在参与者提交任务参与请求后，任务发布者需要验证参与者的位置信息以确定候选参与者，进而由链上节点完成参与者选择. 在参与者提交感知数据后，任务发布者将协助链上节点完成质量评估，并从区块链上获取符合要求的有效感知数据. 基于事先制定的奖励规则，任务发布者通过区块链向保质完成任务的参与者支付报酬.

(2) 参与者. 有意愿参与感知任务的参与者需要提交位置记录到区块链. 在保护位置隐私下，参与

者确定一个最大矩形隐匿区域并采用保序加密和Merkle树来计算位置密文以及构造关于隐匿区域的Merkle树，进而生成相应的位置记录.在获知感知任务后，参与者自行计算自身位置与任务位置之间的距离，从而基于任务的感知位置距离上限要求以决定是否参与任务响应.确定参与任务响应后，基于不同的位置隐私保护需求，参与者从事先确定的最大矩形隐匿区域中选择一个子矩形隐匿区域，并将精确的位置隐藏于子矩形隐匿区域中，进而采用保序加密来生成位置信息.参与者提交包含位置信息以及验证Merkle根哈希值所需信息的任务请求至区块链.被选中且提交有效数据的参与者可以获得相应的报酬奖励以及信誉奖励.

(3)区块链.区块链是一种分布式账本技术，采用区块和链式结构存储合法有效的交易记录.为了保证任务发布者与参与者之间交易的公平性，需要借助区块链来完成参与者选择、任务完成质量评估、信誉更新与报酬计算.具体而言，基于任务发布者提供的候选参与者信息，链上节点在有限的任务预算下完成参与者选择.在参与者提交感知数据后，链上节点进行数据质量评估，并基于数据评估结果完成参与者的报酬计算与信誉更新.通过合法性验证的交易信息将被记录在分布式账本中.

4 预算受限下具有隐私保护的参与者选择方法LPWS

4.1 注 册

任务发布者与参与者需要在区块链上进行身份注册.参与注册的每个用户拥有一对公私钥对$\{pk, sk\}$.为了方便描述，这里将公钥的哈希值$Hash(pk)$简单当作用户在区块链上的身份标识和交易地址.

4.2 位置记录生成

在位置记录生成阶段，假设参与者提交的位置记录是正确且真实的.目前，已有一些有效的解决方案[31-32]提供了位置验证服务，从而保证了位置记录的真实性.LPWS着重考虑参与者提交任务请求过程中的位置隐私泄露问题，因此对于位置记录的验证细节不展开描述.有意愿参与任务的参与者需要将当前的位置记录上传至区块链.在LPWS中，将位置坐标乘以相同的扩大因子以转换为整数坐标，而扩大因子的取值可根据GPS坐标的精度来确定.假设参与者w_i在t_i时间的位置为$lw_i(x_{t_i}, y_{t_i})$，为了保护位置隐私，参与者借助保序加密[33]对位置进行加密.受启发于文献[34]，LPWS要求参与者事先确定一个最大的矩形隐匿区域$L=\{(bx,by)\mid 0\le bx\le X, 0\le by\le Y\}$.其中，$X$和$Y$分别为$L$的最大横坐标和纵坐标.参与者可自定义区域划分水平$H$，划分$L$为多个等分网格，即$L=\left\{bx_h \mid 1\le h\le 2^H \wedge bx_h = h\cdot\frac{X}{2^H}\right\}\cup\left\{by_h \mid 1\le h\le 2^H \wedge by_h = h\cdot\frac{Y}{2^H}\right\}$.基于边界点$\{bx_1, bx_2, \cdots, bx_{2^H}\}$和$\{by_1, by_2, \cdots, by_{2^H}\}$，参与者分别构造关于隐匿区域边界的Merkle树$xMTree$和$yMTree$.本文使用$hash_{l_i}^{bx_h}$、$hash_{l_i}^{bx_{h+j}\|bx_{h+j+1}\|\cdots}$等形式表示Merkle树中的每个节点.图2显示了$H=3$时的隐匿区域划分结果，此时$L=\{bx_1, bx_2, \cdots, bx_8\}\cup\{by_1, by_2, \cdots, by_8\}$.参与者将包含位置密文哈希值、$xMTree_{root}$以及$yMTree_{root}$的位置记录并附上签名后上传至区块链，算法1描述了位置记录生成的过程.链上节点收到位置记录后，首先验证参与者的注册状态以及签名的有效性，然后将有效的位置记录写入分布式账本.为了消除密文间的关联性，参与者w_i可使用不同的保序加密密钥进行加密.位置记录可表示为$lrecord_{t_i}=pk_{w_i}\|lID_{t_i}\|hash_{t_i}\|xMTree_{root}\|yMTree_{root}\|t_i\|Sig_{w_i}$.其中，$pk_{w_i}$为参与者$w_i$的公钥，$lID_{t_i}$为位置记录的ID编号，$hash_{t_i}=Hash(cipher_{t_i})$为位置密文的哈希值，$t_i$表示时间戳，$Sig_{w_i}$为参与者$w_i$使用私钥$sk_{w_i}$生成对位置记录的数字签名.$xMTree_{root}$和$yMTree_{root}$分别为$xMTree$和$yMTree$的根哈希值.由于人具有移动性，因此有意愿参与任务的参与者若发生了位置变化，则需要在获取任务前将最新的一个位置记录提交至区块链.其中，若隐匿区域L未发生改变，参与者则无需重复上传$xMTree_{root}$和$yMTree_{root}$.

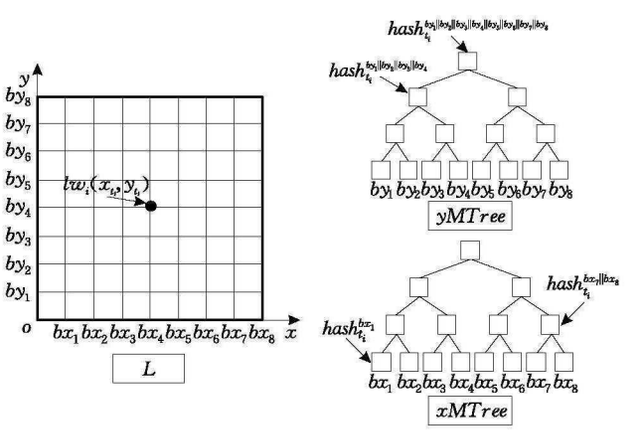


图2 $H=3$时的隐匿区域划分

算法 1.　位置记录生成.

输入：参与者位置 $lw_i(x_{l_i},y_{l_i})$，最大隐匿区域 L，区域划分水平 H

输出：位置记录 $lrecord_{l_i}$

1. 执行 $OPE.Gen(\cdot)$ 以生成密钥 $enc_{l_i,x}$ 和 $enc_{l_i,y}$；
   /* $OPE.Gen(\cdot)$ 为保序加密的密钥生成算法 */
2. $cipher^x_{l_i}=Enc(enc_{l_i,x},x_{l_i})$，$cipher^y_{l_i}=Enc(enc_{l_i,y},y_{l_i})$；
   /* 加密位置的横坐标、纵坐标 */
3. $cipher_{l_i}=cipher^x_{l_i}\|cipher^y_{l_i}$，$hash_{l_i}=Hash(cipher_{l_i})$；
   /* 计算位置密文以及位置密文的哈希值 */
4. FOR $h=1$ to 2^H DO
5. $cipher^{bx_h}_{l_i}=Enc(enc_{l_i,x},bx_h)$；/* 加密隐匿区域 L 各个边界点的横坐标 */
6. $cipher^{by_h}_{l_i}=Enc(enc_{l_i,y},by_h)$；/* 加密隐匿区域 L 各个边界点的纵坐标 */
7. $hash^{bx_h}_{l_i}=Hash(bx_h\|cipher^{bx_h}_{l_i})$；/* 计算各个边界点横坐标的哈希值 */
8. $hash^{by_h}_{l_i}=Hash(by_h\|cipher^{by_h}_{l_i})$；/* 计算各个边界点纵坐标的哈希值 */
9. END FOR
10. $xMTree\leftarrow genTree(hash^{bx_1}_{l_i},hash^{bx_2}_{l_i},\cdots,hash^{bx_{2^H}}_{l_i})$；/* 计算关于隐匿区域边界的横坐标 Merkle 树 $xMTree$ */
11. $yMTree\leftarrow genTree(hash^{by_1}_{l_i},hash^{by_2}_{l_i},\cdots,hash^{by_{2^H}}_{l_i})$；/* 计算关于隐匿区域边界的纵坐标 Merkle 树 $yMTree$ */
12. $lrecord_{l_i}=pk_{w_i}\|lID_{l_i}\|hash_{l_i}\|xMTree_{root}\|yMTree_{root}\|t_i\|Sig_{w_i}$；/* 生成位置记录 */

4.3　感知任务发布

任务发布者 re 发布任务至区块链，任务信息可表示为 $task_j=pk_{re}\|tID_j\|\{obj_1,obj_2,\cdots,obj_M\}\|lt(x_j,y_j)\|B\|fd\|time\|hpk\|Sig_{re}$. 其中，$pk_{re}$ 为任务发布者 re 的公钥，tID_j 表示感知任务的 ID 编号，$\{obj_1,obj_2,\cdots,obj_M\}$ 表示感知任务所包含的 M 个观测对象，$lt(x_j,y_j)$ 表示任务位置，B 为任务总预算，Sig_{re} 为任务发布者 re 使用私钥 sk_{re} 生成对感知任务的数字签名. 任务发布者指定了感知位置距离上限 fd，并且规定了数据提交的截止时间 $time$ 以及参与者执行概率性加密算法所使用的加密公钥 hpk. 链上节点收到任务信息后，首先验证任务发布者的注册状态以及签名的有效性，进而记录有效的任务信息.

4.4　任务请求提交

参与者 w_i 在获知任务信息后，计算自身位置与任务位置的距离 dt_{w_i} 以决定是否参与任务. 假设参与者 w_i 在 t_k 时间提交任务请求，此时参与者的位置表示为 $lw_i(x_{l_k},y_{l_k})$ 以及区块链上对应的位置记录表示为 $lrecord_{l_k}=pk_{w_i}\|lID_{l_k}\|hash_{l_k}\|xMTree_{root}\|yMTree_{root}\|t_k\|Sig_{w_i}$. 基于不同的位置保护需求，参与者 w_i 首先在最大的矩形隐匿区域 L 中选择一个子矩形隐匿区域 $subL$，并确定 $subL$ 的外接圆圆心 (x_o,y_o) 与半径 r^*，以及 $subL$ 的最小横纵坐标 $bx_{\min}$、$by_{\min}$ 和最大横纵坐标 $bx_{\max}$、$by_{\max}$. $subL$ 的大小则代表了位置隐私保护程度，即 $subL$ 越大，位置隐私保护强度越大. 其次，参与者 w_i 执行距离计算函数 $d((x_o,y_o),lt(x_j,y_j))+r^*$ 以得到隐私保护下参与者位置与任务位置之间的距离 dt_{w_i}. 若 $dt_{w_i}\le fd$，参与者 w_i 可选择提交任务请求至区块链. 最后，参与者 w_i 采用位置记录生成阶段的保序加密密钥分别对 $\{bx_{\min},bx_{\max},by_{\min},by_{\max}\}$ 进行加密. 此外，参与者需要提供一些在验证 Merkle 树的根哈希值时所需的辅助节点信息以及对任务请求的数字签名. 图 3 展示了 $H=3$ 时参与者可选择的一个子隐匿区域，其标识了子隐匿区域的边界节点信息以及用以验证 Merkle 根哈希值的辅助节点信息. 算法 2 描述了参与者生成任务请求提交信息的过程.

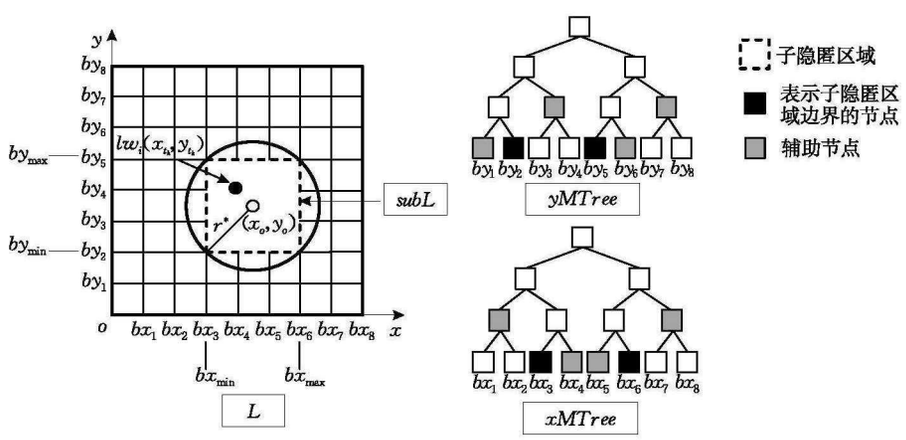


图 3　$H=3$ 时的子隐匿区域示例

算法 2. 任务请求提交.

输入:参与者位置 $lw_i(x_{l_k},y_{l_k})$,任务位置 $lt(x_j,y_j)$,保序加密密钥 $enc_{l_k,x}$、$enc_{l_k,y}$,位置密文 $cipher_{l_k}$,位置记录编号 lID_{l_k}

输出:任务请求 sub_i

1. 确定子矩形隐匿区域 $subL$,并确定 $subL$ 外接圆的圆心(x_o,y_o)和半径 r^*,以及坐标信息$\{bx_{\min},bx_{\max},by_{\min},by_{\max}\}$;
2. $dt_{w_i}=d((x_o,y_o),lt(x_j,y_j))+r^*$; /* 计算参与者与任务之间的位置距离 */
3. IF $dt_{w_i}\leq fd$ THEN
4. $C_1=bx_{\min}\|cipher_{l_k}^{bx_{\min}}$;
5. $C_2=bx_{\max}\|cipher_{l_k}^{bx_{\max}}$;
6. $C_3=by_{\min}\|cipher_{l_k}^{by_{\min}}$;
7. $C_4=by_{\max}\|cipher_{l_k}^{by_{\max}}$;
8. $vpath_{l_k}^{bx}=\{hash_{l_k}^{bx_h},hash_{l_k}^{bx_h+j\|bx_{h+j+1}\|\cdots},\cdots\}$; /* 验证 $xMTree$ 根哈希值所需的辅助节点 */
9. $vpath_{l_k}^{by}=\{hash_{l_k}^{by_h},hash_{l_k}^{by_h+j\|by_{h+j+1}\|\cdots},\cdots\}$; /* 验证 $yMTree$ 根哈希值所需的辅助节点 */
10. $Loc_i=pk_{w_i}\|lID_{l_k}\|(x_o,y_o)\|r^*\|cipher_{l_k}\|C_1\|C_2\|C_3\|C_4\|vpath_{l_k}^{bx}\|vpath_{l_k}^{by}\|Sig_{w_i}^*,sub_i=tID_j\|Loc_i$; /* 生成任务请求 */
11. ELSE
12. $sub_i=$null; /* 任务请求为 null 表示参与者不参与任务 */
13. END IF

4.5 参与者选择

收到参与者的任务请求后,链上节点将验证签名的有效性,而后将有效的任务请求记录到分布式账本中.假设分布式账本中记录了 N 个参与者所提交的任务请求.为了抵御参与者伪造虚假位置的恶意行为,任务发布者可从区块链上获取参与者的位置记录和任务请求以进行位置验证.任务发布者首先验证参与者位置与任务位置之间的距离是否不超过 fd 以及验证圆心(x_o,y_o)到 $subL$ 的任意一个顶点(例如$(bx_{\min},by_{\min})$)的距离是否等于 r^*.本文以圆心(x_o,y_o)到任务位置 $lt(x_j,y_j)$的距离并加上半径 r^* 后作为隐私保护需求下参与者位置与任务位置之间的距离 dt_{w_i}.其次,任务发布者验证参与者所提交的位置密文是否与在位置记录生成阶段所上传的位置密文保持一致,即 $Hash(cipher_{l_k})\stackrel{?}{=}hash_{l_k}$.由于保序加密具备明文与密文保持一致数值顺序的特性,且参与者在任务请求生成中使用位置记录生成阶段的同一保序加密密钥来分别加密子隐匿区域 $subL$ 的顶点坐标,故参与者的位置密文与子隐匿区域的顶点密文之间存在大小比较关系.因此,任务发布者可通过直接比较位置密文与 $subL$ 的顶点密文的大小来验证参与者的位置是否落在子隐匿区域 $subL$ 中,即 $cipher_{l_k}^{bx_{\min}}\stackrel{?}{\leq}cipher_{l_k}^{x}\stackrel{?}{\leq}cipher_{l_k}^{bx_{\max}}$,$cipher_{l_k}^{by_{\min}}\stackrel{?}{\leq}cipher_{l_k}^{y}\stackrel{?}{\leq}cipher_{l_k}^{by_{\max}}$.最后,任务发布者根据辅助节点信息 $vpath_{l_k}^{bx}$ 和 $vpath_{l_k}^{by}$ 来分别计算根哈希值 $xMTree_{root^*}$ 和 $yMTree_{root^*}$,进而验证根哈希值与位置记录生成阶段参与者所上传的根哈希值的一致性,即 $xMTree_{root^*}\stackrel{?}{=}xMTree_{root}$,$yMTree_{root^*}\stackrel{?}{=}yMTree_{root}$.通过位置验证的参与者将成为候选参与者,任务发布者将提交候选参与者信息并附上签名至区块链.基于候选参与者信息,在有限的任务预算下,链上节点将执行参与者选择.LPWS 综合考虑参与者的位置距离 dt_{w_i} 以及信誉 rep_{w_i} 对任务质量价值的影响,定义了任务质量价值 QV 的计算方式,并以 QV 作为参与者选择的筛选指标.参与者位置与任务位置越接近,且参与者的信誉值越高,相应的任务质量价值越大,具体的计算公式如式(1)所示.其中,α、β 为调整因子且 $0\leq\alpha,\beta\leq1$、$\alpha+\beta=1$.

$$QV_{w_i}=\alpha\cdot rep_{w_i}+\beta\cdot\left(\frac{1}{2}\right)^{dt_{w_i}} \tag{1}$$

为了激励参与者积极参与任务并尽可能地提供精确的位置信息,LPWS 将感知任务的有效感知区域进行划分,并为每个子感知区域设置不同的任务基础报酬.具体而言,基于感知位置距离上限 fd 和间隔步长 val,将 fd 划分为 $U=fd/val$ 个距离间隔$\{\lambda_1,\lambda_2,\cdots,\lambda_U\}$,其中每个距离间隔可表示为 $\lambda_u=u\cdot val(u=1,2,\cdots,U)$.假设最大的任务基础报酬为 W,不同距离间隔内的任务基础报酬则可表示为 $rew_{\lambda_u}=W+(u-1)\cdot q$.图 4 显示了 $fd=500$ m、$val=100$ m、$W=10$、$q=-2$ 时的有效感知区域划分结果以及每个子感知区域的任务基础报酬.

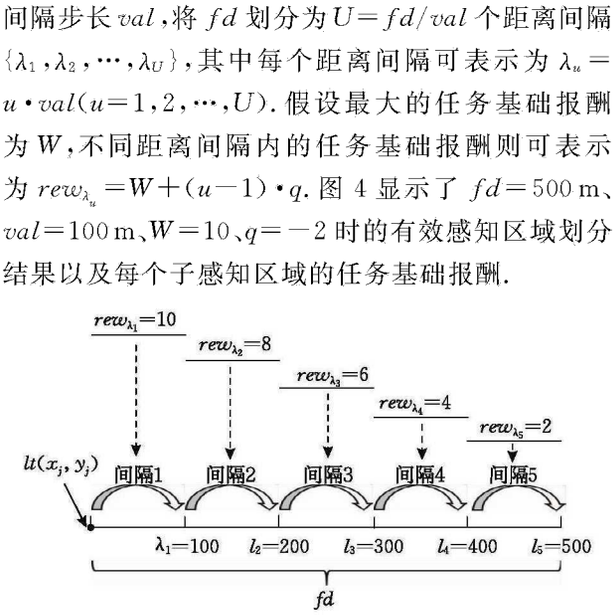


图 4 制定任务基础报酬的示例

在 LPWS 中，任务发布者将任务总预算 B 拆分为两部分，即 $B=B_1+B_2$. B_1 作为在参与者选择阶段使用的预算金额，B_2 则作为在报酬支付阶段使用的预算金额. 在有限的任务预算 B_1 下，任务发布者希望尽可能选择高任务质量价值的参与者来执行任务，因此参与者选择问题可以表示为优化问题：

$$\text{maximize} \sum_{i-1}^{N} y(i)QV_{w_i} \tag{2}$$

$$\text{s.t.} \sum_{i-1}^{N} y(i)rew_{w_i,\lambda_u} \leq B_1 \tag{3}$$

$$y(i)\in\{0,1\},\ \forall i\in\{1,2,\cdots,N\} \tag{4}$$

其中，QV_{w_i} 表示参与者 w_i 的任务质量价值，rew_{w_i,λ_u} 表示参与者 w_i 处于距离间隔 λ_u 范围时所能获得的任务基础报酬. 为了保证参与者选择的公平性，链上节点共同完成参与者选择. 参与者选择问题可以看成是 0-1 背包问题，而 0-1 背包问题是一种组合优化的 NP 完全问题. 本文选择动态规划的方法来求解参与者选择问题，而其主要思想为以自底向上的方式，借助集合存储每选择一个参与者时总报酬和总价值发生改变的跳跃点，进而依次回溯集合以确定选中的参与者编号. 算法 3 描述了参与者选择的过程.

算法 3.　参与者选择.

输入：预算 B_1，参与者人数 N，参与者的任务质量价值 $QV[N]$，参与者的任务基础报酬 $rew[N]$，存放跳跃点的集合 $State$，价值最大的跳跃点 $Best$

输出：参与者选择结果 sel_result

1. 初始化 N 个参与者跳跃点集合 $State$ 为{(0,0)}；/* 跳跃点(0,0)表示未选择参与者时，报酬和价值均为 0 */
2. FOR $i=N-1$ to 0 DO
3. $Temp \leftarrow State_{i+1}+(rew[i],QV[i])$；/* 不超过 B_1 下，更新 $State_{i+1}$ 的跳跃点，并放入集合 $Temp$ */
4. $State_i \leftarrow Merge(State_{i+1},Temp)$；/* 合并 $State_{i+1}$ 和 $Temp$，舍弃 $State_i$ 中报酬大而价值小的跳跃点 */
5. END FOR
6. FOR $i=0$ to $N-1$ DO
7. IF $State_{i+1}$ 中存在跳跃点 (a,b)，使得 $(a+rew[i],b+QV[i])=Best$ THEN
8. $Best=Best-(rew[i],QV[i])$；
9. $sel\_result.add(i+1)$；
10. END IF
11. END FOR

4.6　质量评估

参与者选择阶段结束后，被选中的参与者执行任务并及时上传感知数据至区块链. 此外，为了保护感知数据隐私，参与者通过添加随机数的方式对感知数据进行扰动，并提供基于 Pedersen 承诺[35]的相关承诺值 $commit(v_m^i\cdot QV_{w_i},r_m^i)$ 和随机数密文以用于评估数据质量. LPWS 使用参与者的感知数据与数据聚合结果之间的距离平均值来表示参与者的数据失真程度，并通过比较数据失真程度与阈值的大小以得到参与者的数据质量评估结果. 数据质量评估过程主要包括以下几个步骤：

(1) 假设在数据提交的截止时间之前，共有 K 个参与者上传感知数据，并记参与者 w_i 关于观测对象 obj_m 的感知数据为 s_m^i. 在质量评估阶段，感知数据以及任务质量价值被转换为整数来计算. 在计算最终的数据聚合结果时，可将其除以相应的扩大倍数. 基于保护数据隐私的需求，参与者可随机选择一个随机数来扰动感知数据 s_m^i，即 $\tilde{s}_m^i=s_m^i-v_m^i$. 进一步，参与者 w_i 采用任务发布者生成的加密公钥 hpk 来加密每个随机数 v_m^i 以得到对应密文 $E(v_m^i)$. 此外，参与者还需计算承诺值 $commit(v_m^i\cdot QV_{w_i},r_m^i)$ 并提交至区块链，且需要将用以计算承诺值的随机数的密文 $E(r_m^i)$ 和用来扰动感知数据的随机数的密文 $E(v_m^i)$ 发送给任务发布者. 这里的承诺值是参与者对用以扰动感知数据的随机数 v_m^i 和任务质量价值 QV_{w_i} 的乘积的承诺，r_m^i 则是用以计算承诺值的随机数.

(2) 由于任务发布者拥有相应的私钥，因此任务发布者可在接收到参与者发送的密文后，分别计算 $D_m=v_m^1\cdot QV_{w_1}+v_m^2\cdot QV_{w_2}+\cdots+v_m^K\cdot QV_{w_K}$ 以及 $\sum_{i-1}^{K} r_m^i$，并提交计算结果至区块链.

(3) 链上节点收到相关的数据信息后，首先验证 $commit(v_m^i\cdot QV_{w_i},r_m^i)$ 的乘积之和是否与 $commit\left(D_m,\sum_{i-0}^{K} r_m^i\right)$ 保持一致. 其次，计算 $task_j$ 中观测对象 obj_m 的数据聚合结果，计算公式如式(5)所示.

$$agg_m=\frac{\sum_{i-1}^{K} QV_{w_i}\cdot\tilde{s}_m^i+D_m}{\sum_{i-1}^{K} QV_{w_i}} \tag{5}$$

(4) 基于聚合结果 agg_m，任务发布者或者参与者都可根据事先制定的数据评估规则计算数据失真程度以得到质量评估结果，计算公式如式(6)所示. 链上节点基于任务发布者以及参与者所上传的质量评估结果，验证每个数据失真程度 $dist_i$ 的匹配性. 通过合法性验证的数据失真程度将被用以计算参与者的报酬奖励以及信誉奖励.

$$dist_i=\frac{\sum_{m-1}^{M}|s_m^i-agg_m|}{M} \tag{6}$$

4.7 信誉更新与报酬计算

基于参与者的任务完成情况，节点通过比较数据失真程度 $dist_i$ 与阈值 $thresh$ 的大小，更新参与者的信誉并计算报酬奖励. 如果参与者的数据失真程度小于阈值，参与者除了获得任务基础报酬，还将得到基于数据质量的额外报酬奖励以及信誉奖励. LPWS 采用 Gompertz 函数[36]来计算参与者的贡献分数，进而更新参与者的信誉. 在有限的任务预算 B_2 下，LPWS 根据每个数据失真程度与阈值的绝对差值占总绝对差值的比值分配额外报酬奖励给参与者. 算法 4 描述了信誉更新与报酬计算的过程.

算法 4. 信誉更新与报酬计算.

输入：数据失真程度 $dist_i$，信誉值 rep_{w_i}，预算 B_2，阈值 $thresh$，惩罚因子 δ，调整因子 γ，数据质量达标参与者集合 Z

输出：参与者的信誉值 $rep_{w_i}^*$ 和报酬奖励 $reward_{w_i}$

1. FOR $i=1$ to K DO
2. IF $dist_i \geq thresh$ THEN
3. $rep_{w_i}^*=rep_{w_i}-\delta$; /* 基于惩罚因子 δ，减少参与者的信誉 */
4. $reward_{w_i}=0$; /* 参与者没有得到报酬 */
5. ELSE
6. $Z=Z+\{w_i\}$; /* 将参与者 w_i 放入数据质量达标参与者集合 Z */
7. END IF
8. END FOR
9. FOR $i=1$ to $|Z|$ DO
10. $sum=\sum_{i-1}^{|Z|}|dist_i-thresh|$; /* 计算集合 Z 中的数据失真程度与阈值的总绝对差值 */
11. END FOR
12. FOR $i=1$ to $|Z|$ DO
13. $G(dist_i)=e^{-e^{-dist_i^{-1}}}$; /* 计算参与者的贡献分数 */
14. $rep_{w_i}^*=rep_{w_i}\cdot(1+\gamma\cdot G(dist_i))$; /* 基于贡献分数，增加参与者的信誉 */
15. $reward_{w_i}=rew_{w_{i,\lambda_u}}+\frac{|dist_i-thresh|}{sum}\cdot B_2$; /* 计算参与者所获得的总报酬 */
16. END FOR

4.8 共识过程

在 LPWS 中，为了保证公平可信的参与者选择结果以及奖惩结果，链上节点负责完成参与者选择、质量评估、报酬计算与信誉更新工作. 假设节点 MN_i 为事先选中的主节点. 在参与者选择阶段，基于任务发布者给定的有限任务预算 B_1，MN_i 通过执行算法 3 来确定参与者选择结果，进而将生成的参与者选择结果作为交易广播到链上. 收到参与者选择结果消息后，链上其他节点 SN_j 则执行相同的参与者选择算法来生成参与者选择结果以验证选择结果的匹配性. 具体而言，SN_j 首先验证参与者人数的一致性. 若参与者人数相等，则下一步执行对参与者选择结果的详细验证；反之，该参与者选择结果则被视为无效. 其次，SN_j 将基于本地执行参与者选择算法所得到的参与者选择结果与接收到的参与者选择结果进行匹配性检查. 若参与者选择结果通过匹配性检查，该参与者选择结果则被视为有效；反之，该参与者选择结果则被认定为无效. 最后，对于有效的参与者选择结果，SN_j 将反馈一个同意答复消息. MN_i 基于收到的反馈消息，将有效的参与者选择结果广播给其他节点. 最终通过合法性检查的参与者选择结果则被记录在分布式账本中.

在质量评估阶段，针对感知任务中的每个观测对象，链上节点通过式(5)来计算数据聚合结果，进而验证数据失真程度的匹配性. MN_i 将计算得到的数据聚合结果广播给其他节点，SN_j 则对本地计算得到的数据聚合结果与收到的数据聚合结果进行一致性检查. 基于通过一致性检查的数据聚合结果，进而节点确定每个参与者的数据失真程度. 完成数据质量评估后，基于任务预算 B_2，链上各节点间执行算法 4 更新每个参与者的信誉以及计算相应的报酬奖励. SN_j 基于收到来自 MN_i 广播的信誉更新和报酬奖励结果，验证其与本地计算的信誉更新和报酬奖励结果的一致性. 基于验证成功的信誉更新和报酬奖励结果，节点更新参与者的信誉值并将其记录在分布式账本中，以及向参与者支付相应的报酬奖励.

5 理论分析与性能评估

本节将从理论和性能两方面对本文所提参与者选择方法 LPWS 进行分析与评估. 理论分析上，从隐私性和公平性两方面分析 LPWS 的可行性，并和其他相关工作进行比较. 在性能评估方面，通过仿真实验对比分析 LPWS 的效用.

5.1 理论分析

LPWS 考虑参与者提交任务请求过程中的位置

隐私泄露问题以及质量评估中的数据隐私泄露问题.此外,LPWS还考虑了参与者选择以及质量评估的公平性,从而提高参与者的参与积极性.

5.1.1　隐私性

LPWS为参与者提供了位置隐私保护.在LPWS中,参与者可通过提供隐匿区域的方式来提交位置信息,并采用保序加密得到精确位置密文和隐匿区域边界节点密文.其中,参与者可事先确定一个最大的矩形隐匿区域,进而基于不同的位置隐私保护需求,选择一个子矩形隐匿区域来隐藏精确的位置.通过构造关于隐匿区域边界的横纵坐标Merkle树,参与者提供相应Merkle根哈希值以及验证根哈希值时所需的辅助节点信息,进而帮助任务发布者验证位置信息的一致性.在保序加密下,任务发布者可通过比较位置密文间的大小以及检查Merkle根哈希值的一致性来验证参与者的位置有效性.在有限的任务预算下,LPWS实现了在不泄露位置隐私情况下执行参与者选择.除了位置隐私安全,LPWS还考虑了数据隐私安全.在参与者选择阶段,基于有限的任务预算,我们侧重于选择最有可能提供高质量数据的参与者.进一步,我们通过执行质量评估来实际计算参与者的感知数据质量,从而更新参与者的信誉以作用于下一次的参与者选择.为了防止恶意参与者通过复制其他参与者的感知数据来谋取不应得的报酬,LPWS允许参与者通过添加随机数的方式来扰动感知数据.此外,参与者可通过提供承诺值来辅助节点完成质量评估.在此过程中,参与者的感知数据始终对除了任务发布者和任务参与者之外的无关人员保持机密性.

5.1.2　公平性

在LPWS中,参与者选择以及质量评估由链上节点共同完成,从而确保了选择公平性和奖惩公平性.在有限的任务预算下,LPWS基于参与者的信誉和位置距离计算任务质量价值,并选择一组具有最大任务质量价值的参与者来执行任务.参与者选择只是以最大化高质量数据的提交可能性为目的选择一组合适的参与者.因此LPWS还需通过质量评估来实际反馈参与者的任务完成情况,进而基于任务质量更新参与者的信誉值以及计算报酬奖励.链上节点以公开透明的方式执行参与者选择、报酬计算以及信誉更新,有效抵御了任务发布者可能存在的有偏见参与者选择以及拒绝付款的恶意行为.此外,参与者的信誉值由链上节点进行更新与维护,参与者无法通过篡改信誉值来提高自己的任务质量价值.进一步地,在LPWS中,任意被选中参与者的低质数据提交行为都将被检测,LPWS通过降低信誉值的方式来惩罚参与者.

5.1.3　相关工作对比

本文将LPWS与相关工作AMT[37]、SenseChain[17]、CrowdBC[20]、Yang[21]进行了对比分析.AMT是基于中心化平台来完成参与者选择与质量评估,而其他相关工作则是基于区块链来设计群智感知的相关流程.本文主要围绕位置隐私、数据隐私、选择公平性以及数据质量评估四个方面进行分析,结果如表2所示.从表2来看,对比于现有的相关工作,LPWS在不依赖中心化平台下,能够以不泄露位置隐私的方式完成参与者选择,并且在保证数据隐私下执行数据质量评估来实际反馈参与者的任务完成情况.LPWS提供了一种保证隐私安全且确保公平性的参与者选择方法,并基于数据质量评估结果来完成参与者的报酬支付与信誉更新,进而有利于提高参与者的积极性和任务完成质量.

表2　相关工作对比

特征	AMT[37]	SenseChain[17]	CrowdBC[20]	Yang[21]	LPWS
位置隐私	−	×	−	√	√
数据隐私	×	×	√	×	√
选择公平性	×	√	√	√	√
数据质量评估	√	√	×	×	√

5.2　性能评估

本文从任务完成质量、质价比、时间开销以及平均时延和吞吐量四个方面评估LPWS的性能.仿真实验采用Java实现,在Intel(R) Core(TM) i5-10210U CPU@1.60GHz 2.11GHz处理器、16GB内存上运行.本文以北京空气质量报告中的测量数据为标准模拟生成参与者的感知数据,且每个感知任务包含的观测对象为5个(例如PM2.5、PM10、SO2等空气指标).参与者的位置则根据任务位置随机生成,且经过坐标放大被转换到二维坐标中.在质量评估阶段,本文采用概率公钥加密算法Paillier[38]以供参与者进行数据加密,并设置安全参数为512位.此外,我们在Hyperledger Fabric v1.4.0上部署链码并采用Kafka共识,以及借助Hyperledger Caliper对运行在单个主机上的测试区块链网络进行性能测试.测试区块链网络包含了2个组织,每个组织各包含2个节点.在相同的实验环境下,每个实验重复运行20次以计算平均值,实验的相关参数如表3所示.

表 3 实验参数设置

参数	数值
N	{100,200,300,400,500}
B_1	{50,60,70,80,90}
B_2	100
rep	[0,1]
fd	3 km
val	0.3 km
α	{0.1,0.3,0.5,0.7,0.9}
β	{0.1,0.3,0.5,0.7,0.9}
$thresh$	0.5
δ	0.2
γ	0.5

5.2.1 任务完成质量

为了评估参与者选择结果对任务完成质量的影响，我们采用 MAE 作为数据误差来衡量任务完成质量，并以质量评估阶段的数据聚合结果作为每个观测对象的最终观测值. MAE 可表示为每个观测对象的最终观测数据 agg_m 与真实数据 $groundtruth_m$ 的平均绝对误差，即 $MAE=\frac{1}{M}\sum_{m-1}^{M}|agg_m-groundtruth_m|$. 在 LPWS 中，调整因子 α 和 β 影响了任务质量价值的计算结果，而不同的任务质量价值可能影响参与者选择结果，进而影响任务完成质量. 调整因子 δ 和 γ 则作用于信誉更新与报酬计算，它们的取值可根据任务发布者制定的奖惩规则来确定，取值越大则说明相应的奖惩力度越大. 因此，我们主要是通过实验来直观分析 α 的不同取值对任务完成质量的影响. 在参与者人数固定为 500，预算分别为 50、70、90 下，我们观察了 α 分别取 0.1、0.3、0.5、0.7、0.9，对应的 β 为 $1-\alpha$ 下的任务完成质量的变化情况，如图 5 所示.

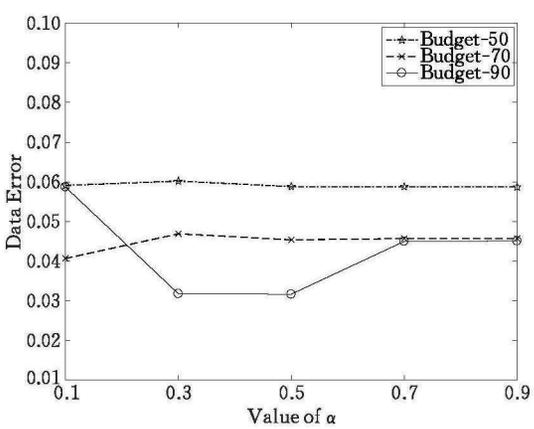
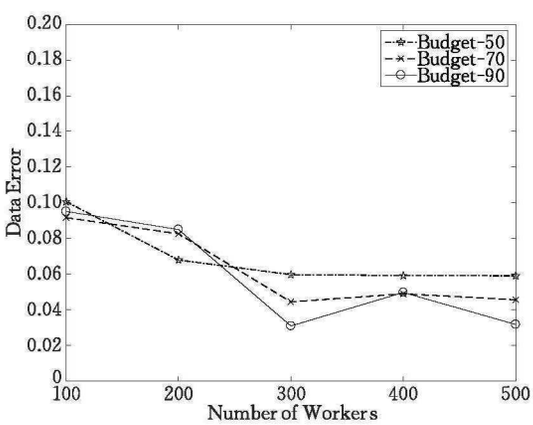


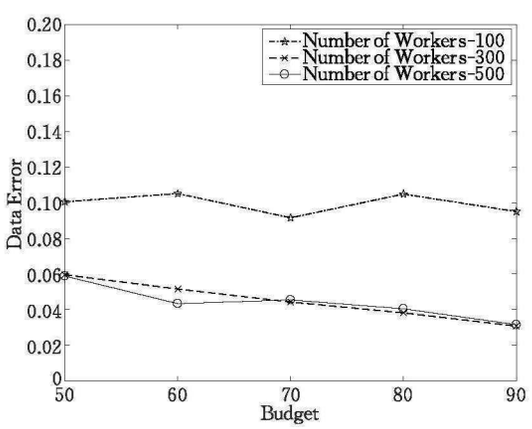
图 5 α 的不同取值对任务完成质量的影响

α 和 β 分别决定了信誉值与位置距离对任务质量价值的影响程度. 由于 $\alpha+\beta=1$，当 α 小于 0.5，β 则大于 0.5，此时位置距离对任务质量价值的影响程度大；反之，则是信誉值对任务质量价值的影响程度大. 观察图 5，当 $\alpha>0.5$ 或者 $\alpha<0.5$ 时，不同预算下的数据误差变化各异. 在三种预算下，$\alpha=0.5$ 时的数据误差均小于 $\alpha>0.5$ 时的数据误差. 相比于 $\alpha<0.5$ 时的数据误差，在不同预算下，虽然 $\alpha=0.5$ 时的数据误差不总是最小，但是总体上 $\alpha=0.5$ 对不同预算下的任务完成质量的影响相对较好. 在实际应用中，α 与 β 的取值大小可根据任务发布者对信誉值与位置距离在任务质量价值上的影响程度的倾向来确定. 结合仿真实验结果，我们建议在任务质量价值上，将信誉值与位置距离视为同等权重的影响因素. 因此，在后续的实验中，α 与 β 均被设置为 0.5.

此外，在 LPWS 中，参与者选择受约束于有限的任务预算 B_1，因此我们以参与者人数与预算为变量来分析参与者选择结果对任务完成质量的影响. 图 6(a)为面向 100、200、300、400 以及 500 个参与者在预算分别为 50、70、90 下执行参与者选择时的任务完成质量. 此外，我们还考虑在预算为 50、60、70、80 以及 90 下分别对 100、300、500 个参与者执行参与者选择的任务完成质量，如图 6(b)所示.



(a) 不同参与者人数下的任务完成质量



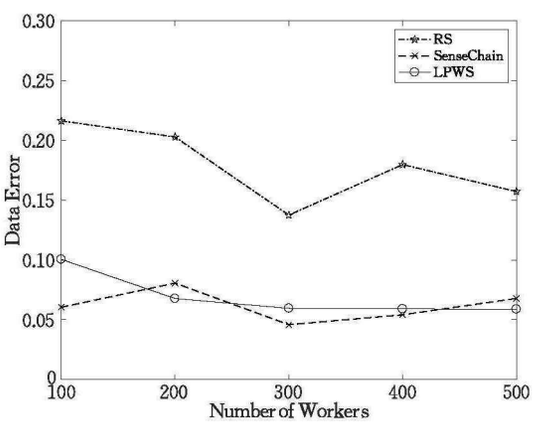
(b) 不同预算下的任务完成质量

图 6 不同变量对任务完成质量的影响

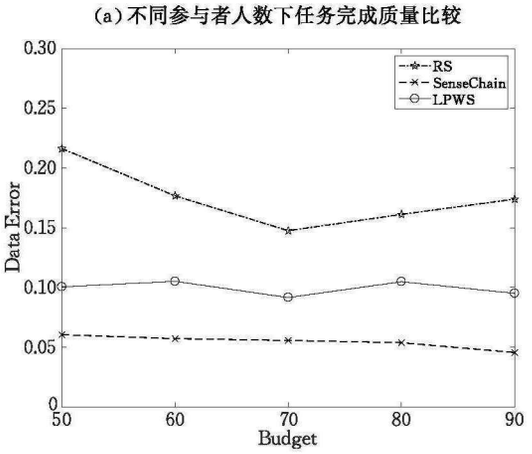
观察图 6(a)，在不同的任务预算下，随着参与者人数的增加，数据误差整体上呈现递减的趋势变

化. 参与者人数增加时，意味着我们可以面向更大的选择范围来选择更加优质的参与者，从而增加获得高质量感知数据的可能性以提高任务完成质量. 从图 6(b)中，我们发现在参与者人数固定时，适当增加预算会减少数据误差. 但是，针对参与者人数较少的情况，持续增加任务预算并不一定会减少数据误差. 这是因为在有限的较少参与者人数下，其包含的优质参与者数量相对较少，因此增加任务预算反而会因为选择了相对低质的参与者而导致任务完成质量降低. 因此，在参与者人数较少时，任务预算则不适宜设置过大，而在参与者人数较多时，可适当增加任务预算以提高任务完成质量.

进一步，为了分析 LPWS 在参与者选择方面的性能，我们将其与随机参与者选择 RS 以及 SenseChain[17] 进行对比分析. RS 即从符合任务要求的候选参与者中随机选择参与者来执行任务，而 SenseChain 则在有限的人数限制下，基于参与者的任务质量信息来选择具有最大任务质量信息的参与者. 由于实验环境不同，我们保留了其核心思想，并对其进行了调整以适合我们的模型. 图 7 显示了三种参与者选择方案在任务完成质量上的对比分析结果. 其中，图 7(a)为任务预算固定为 50，参与者人数分别为 100、200、300、400 以及 500 下的任务完成质量比较结果. 图 7(b)为参与者人数固定为 100，预算分别为 50、60、70、80 以及 90 下的任务完成质量比较结果.



(a) 不同参与者人数下任务完成质量比较



(b) 不同预算下任务完成质量比较

图 7　不同方案间的任务完成质量比较

在图 7(a)中，当预算固定为 50 时，随着参与者人数的增加，RS 的数据误差始终最高，而 LPWS 的数据误差则略高于 SenseChain 的数据误差. 观察图 7(b)中的数据误差，我们可以得到相似的分析结果. 对比于 RS，LPWS 体现了明显的参与者选择效果. 虽然 LPWS 的任务完成质量不如 SenseChain 的任务完成质量，但是两者的差异性不大. 此外，SenseChain 需要提前确定参与者选择人数，且未提供位置隐私保护以及数据隐私保护. 因此，在三种参与者选择方案中，LPWS 在参与者选择方面呈现一定的优势.

5.2.2　质价比

对于任务发布者而言，其希望能够以尽可能少的任务预算来补偿参与者的感知开销，并获得高质量的感知数据. 本文将任务完成质量与实际的任务预算支出的比值定义为质价比，进而分析在固定的任务总预算 $B=B_1+B_2=50+100=150$、不同的参与者人数下，比较 LPWS 与 SenseChain 的质价比，如图 8 所示. 需要注意的是，这里描述的实际的任务预算支出是指实际用以支付数据质量达标者报酬奖励的花费.

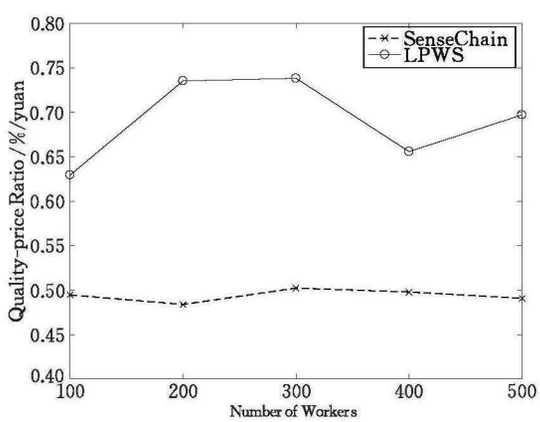


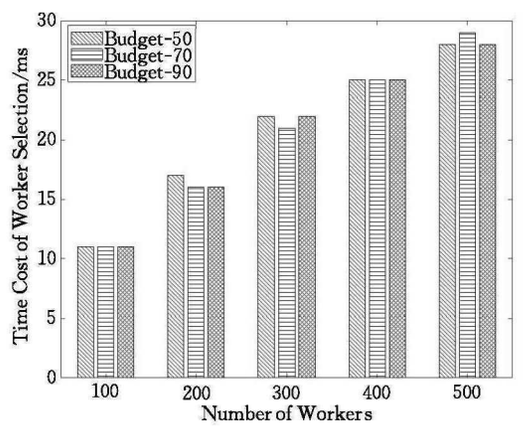
图 8　质价比比较

在图 8 中，当任务总预算固定为 150 时，随着参与者人数的增加，SenseChain 的质价比趋于稳定，LPWS 的质价比则始终优于 SenseChain 的质价比. 在这两种方案中，每个数据质量达标的参与者都可以获得任务基础报酬和基于质量的额外报酬奖励. 但是，在 SenseChain 中，任务总预算即为实际的

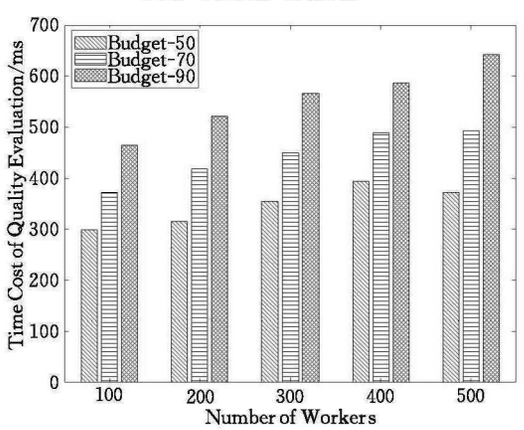
任务预算支出，而 LPWS 则将任务总预算划分为两部分，进而分别使用于参与者选择和报酬支付，因而有利于减少实际的任务预算支出，即实际的任务预算支出可能少于任务总预算. 此外，相比于 SenseChain 中统一固定的任务基础报酬，LPWS 则基于位置距离为不同感知区域内的参与者制定个性化的任务基础报酬. 相比于 SenseChain，LPWS 在质价比方面具有一定的优势，且 LPWS 还进一步考虑参与者的位置隐私保护需求，故 LPWS 更能够吸引参与者与任务发布者的参与兴趣.

5.2.3 时间开销

本文主要从参与者选择时间开销和质量评估时间开销两方面评估 LPWS 的效率性. 在有限的任务预算 B_1 下，本文采用动态规划的方法来执行参与者选择，并在保护数据隐私安全下，完成质量评估以将评估结果用于信誉更新与报酬计算. 因此，在任务预算分别为 50、70、90 下，我们统计了节点对 100、200、300、400 以及 500 个参与者单独执行参与者选择的时间开销，如图 9(a)所示. 此外，面向 100、200、300、400 以及 500 个参与者，我们分别在任务预算为 50、70、90 下统计了节点对被选中参与者单独进行质量评估的时间开销，如图 9(b)所示.



(a) 参与者选择时间开销



(b) 质量评估时间开销

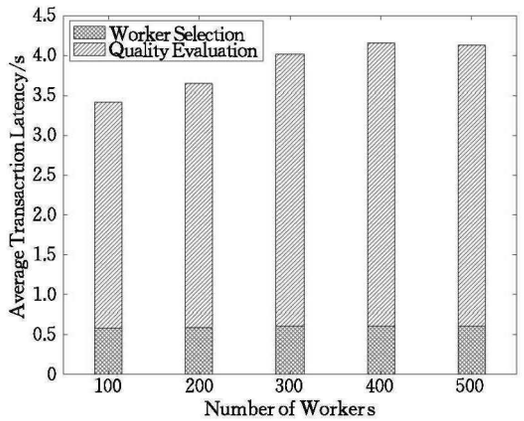
图 9 时间开销

观察图 9(a)，在不同的任务预算下，随着参与者人数的增加，参与者选择的时间开销均呈现增大的趋势. 在相同的参与者人数下，任务预算的增加对参与者选择时间开销的影响较小，这是因为在参与者选择阶段，参与者人数影响了动态规划求解中跳跃点集合的计算时间. 因此，参与者人数的增加导致参与者选择时间开销的增大. 观察图 9(b)，在相同的参与者人数下，随着任务预算的增加，质量评估时间开销保持增大趋势. 在相同的任务预算下，参与者人数的增加导致质量评估时间开销增大. 在质量评估阶段，节点对在参与者选择阶段被选中的参与者执行质量评估，而任务预算的增加使得被选中参与者人数增多，进而导致质量评估时间开销增大. 此外，在相同的任务预算下，随着参与者人数的增加，参与者选择范围增大，因此被选中参与者人数可能会增加，进而增加了质量评估时间开销.

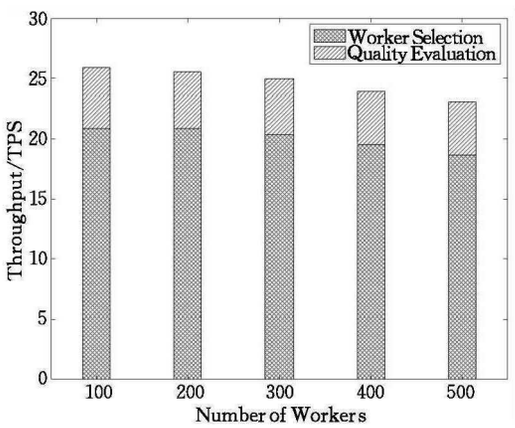
5.2.4 平均时延与吞吐量

在 LPWS 的参与者选择阶段以及质量评估阶段中，参与者选择结果以及数据聚合结果分别以交易的形式广播到区块链上. 为了评估交易从提交到响应的平均时延以及吞吐量，在固定的 30 s 测试时间、固定的 5 个工作进程数量以及预算 B_1 为 50 下，我们分别观察不同参与者人数下参与者选择以及质量评估中交易的平均时延和吞吐量变化，如图 10 所示.

在图 10(a)中，随着参与者人数增加，参与者选择中交易的平均时延保持平稳，而质量评估中交易的平均时延整体呈现上升趋势. 从图 9 可知，参与者人数的增加会导致参与者选择时间开销以及质量评估时间开销的增大，但参与者选择时间开销小且增量较小，而质量评估时间开销大且增量相对较大. 因此质量评估中交易的平均时延相对于参与者选择中交易的平均时延，其增长变化明显且平均时延相对较大. 在图 10(b)中，随着参与者人数的增加，参与者选择与质量评估中交易的吞吐量呈现略微下降的变化趋势. 在有限的时间内，由于参与者选择中交易的平均时延小于质量评估中交易的平均时延，因此参与者选择中交易的吞吐量优于质量评估中交易的吞吐量.



(a) 参与者选择和质量评估中交易的平均时延



(b) 参与者选择与质量评估中交易的吞吐量

图 10　平均时延与吞吐量

6　总　结

本文提出一种基于区块链的群智感知中任务预算约束的位置隐私保护参与者选择方法 LPWS. 在位置隐私保护方面，LPWS 借助保序加密和 Merkle 树为参与者提供个性化的位置隐私保护. 在有限的任务预算下，LPWS 将参与者选择问题建模为目标优化问题，并由链上节点基于动态规划来完成参与者选择以保证参与者选择的公平可信性. 此外，LPWS 不仅仅关注参与者选择结果，还在保护数据隐私下评估参与者的数据质量以提供公平的报酬计算与信誉更新，进而激励参与者尽可能地提供高质量数据. 参与者可通过随机扰动数据的方式来保护数据隐私，并提供承诺值来辅助链上节点完成数据质量评估. 仿真结果表明，在保证任务完成质量下，对比于相关工作，LPWS 不仅实现了安全公平的参与者选择，而且确保了质量评估的公平性以及获得了更好的质价比，进而提高了参与者与任务发布者的参与积极性. 在保证位置隐私安全下，如何进一步提高效率以及更细粒度地分析影响参与者选择的评估指标，这将是下一步的研究重点.

参　考　文　献

[1] Wu Yao, Zeng Ju-Ru, Peng IIui, et al. Survey on incentive mechanisms for crowd sensing. Journal of Software, 2016, 27(8): 2025-2047(in Chinese)
(吴垚，曾菊儒，彭舸等. 群智感知激励机制研究综述. 软件学报，2016，27(8)：2025-2047)

[2] Guo Bin, Wang Zhu, Yu Zhi-Wen, et al. Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm. ACM Computing Surveys, 2015, 48(1): 1-31

[3] Zhao Bo-Wen, Tang Shao-IIua, Liu Xi-Meng, et al. Pace: Privacy-preserving and quality-aware incentive mechanism for mobile crowdsensing. IEEE Transactions on Mobile Computing, 2020, 20(5): 1924-1939

[4] Cai Cheng-Jun, Zheng Yi-Feng, Du Yue-Feng, et al. Towards private, robust, and verifiable crowdsensing systems via public blockchains. IEEE Transactions on Dependable and Secure Computing, 2019, 18(4): 1893-1907

[5] Feng Wei, Yan Zheng. MCS-Chain: Decentralized and trustworthy mobile crowdsourcing based on blockchain. Future Generation Computer Systems, 2019, 95(6): 649-666

[6] Lu Yuan, Tang Qiang, Wang Gui-Ling. ZebraLancer: Private and anonymous crowdsourcing system atop open blockchain// Proceedings of the 38th IEEE International Conference on Distributed Computing Systems. Vienna, Austria, 2018: 853-865

[7] Lin Chao, IIe De-Biao, Zeadally S, et al. SecBCS: A secure and privacy-preserving blockchain-based crowdsourcing system. Science China Information Sciences, 2020, 63(3): 1-14

[8] Liu Liang, Liu Wu, Zheng Yu, et al. Third-Eye: A mobile phone-enabled crowdsensing system for air quality monitoring. Proceedings of the ACM on Interactive Mobile Wearable and Ubiquitous Technologies, 2018, 2(1): 1-26

[9] Liu Zhi-Dan, Jiang Shi-Qi, Zhou Peng-Fei, et al. A participatory urban traffic monitoring system: The power of bus riders. IEEE Transactions on Intelligent Transportation Systems, 2017, 18(10): 2851-2864

[10] Wen Yu-Tian, Shi Jin-Yu, Zhang Qi, et al. Quality-driven auction-based incentive mechanism for mobile crowd sensing. IEEE Transactions on Vehicular Technology, 2015, 64(9): 4203-4214

[11] Peng Dan, Wu Fan, Chen Gui-IIai. Pay as how well you do: A quality-based incentive mechanism for crowdsensing// Proceedings of the 16th ACM International Symposium on Mobile Ad IIoc Networking and Computing. IIangzhou, China, 2015: 177-186

[12] Montjoye Y A D, IIidalgo C A, Verleysen M, et al. Unique in the crowd: The privacy bounds of human mobility. Scientific Reports, 2013, 3(6): 1-5

[13] Montjoye Y A D, Radaelli L, Singh V K, et al. Unique in the shopping mall: On the identifiability of credit card metadata. Science, 2015, 347(6221): 536-539

[14] Kazemi L, Shahabi C, Chen Lei. GeoTruCrowd: Trustworthy query answering with spatial crowdsourcing//Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. Orlando, USA, 2013: 314-323

[15] Miao Chun-Yan, Yu IIan, Shen Zhi-Qi, et al. Balancing quality and budget considerations in mobile crowdsourcing. Decision Support Systems, 2016, 90(10): 56-64

[16] An Jian, Liang Dan-Wei, Gui Xiao-Lin, et al. Crowdsensing quality control and grading evaluation based on a two-consensus blockchain. IEEE Internet of Things Journal, 2019, 6(3): 4711-4718

[17] Kadadha M, Otrok II, Mizouni R, et al. SenseChain: A blockchain-based crowdsensing framework for multiple requesters and multiple workers. Future Generation Computer Systems, 2020, 105(4): 650-664

[18] Li Qing-IIua, Cao Guo-IIong. Providing privacy-aware incentives for mobile sensing//Proceedings of the 34th IEEE International Conference on Distributed Computing Systems. Madrid, Spain, 2014: 1-14

[19] Wang Jing-Zhong, Li Meng-Ru, IIe Yun-IIua, et al. A blockchain based privacy-preserving incentive mechanism in crowdsensing applications. IEEE Access, 2018, 6: 17545-17556

[20] Li Ming, Weng Jian, Yang An-Jia, et al. CrowdBC: A blockchain-based decentralized framework for crowdsourcing. IEEE Transactions on Parallel and Distributed Systems, 2019, 30(6): 1251-1266

[21] Yang Meng-Meng, Zhu Tian-Qing, Liang Kai-Tai, et al. A blockchain-based location privacy-preserving crowdsensing system. Future Generation Computer Systems, 2019, 94(5): 408-418

[22] Wei Li-Jun, Wu Jing, Long Cheng-Nian, et al. A blockchain-based hybrid incentive model for crowdsensing. Electronics, 2020, 9(2): 1-18

[23] Pouryazdan M, Kantarci B, Soyata T, et al. Quantifying user reputation scores, data trustworthiness, and user incentives in mobile crowdsensing. IEEE Access, 2017, 5: 1382-1397

[24] Ding Yue-Jiao, Chen Zhong-Yu, Lin Fei-Long, et al. Blockchain-based credit and arbitration mechanisms in crowdsourcing//Proceedings of the 3rd International Symposium on Autonomous Systems. Shanghai, China, 2019: 490-495

[25] Zou Shi-IIong, Xi Jin-Wen, Wang IIong-Gang, et al. CrowdBLPS: A blockchain-based location-privacy-preserving mobile crowdsensing system. IEEE Transactions on Industrial Informatics, 2020, 16(6): 4206-4218

[26] Li Liang, Zhang Xin-Yue, IIou Rong-IIui, et al. Participant recruitment for coverage-aware mobile crowdsensing with location differential privacy//Proceedings of the 2019 IEEE Global Communications Conference. Waikoloa Village, USA, 2019: 1-6

[27] Wang Tao-Chun, Liu Ying, Jin Xin, et al. Research on location and data privacy protection method based on k-anonymity in crowdsensing. Journal of Communications, 2018, 39(Z1): 170-178(in Chinese)
(王涛春, 刘盈, 金鑫等. 群智感知中基于k-匿名的位置及数据隐私保护方法研究. 通信学报, 2018, 39(Z1): 170-178)

[28] Tian Ye, Li Xiong, Sangaish A K, et al. Privacy-preserving scheme in social participatory sensing based on secure multi-party cooperation. Computer Communications, 2018, 119(4): 167-178

[29] Zhao Ke, Tang Shao-IIua, Zhao Bo-Wen, et al. Dynamic and privacy-preserving reputation management for blockchain-based mobile crowdsensing. IEEE Access, 2019, 7: 74694-74710

[30] Ma Rong, Chen Lei, Xiong Jin-Bo, et al. A personalized privacy protection data uploading scheme for mobile crowdsensing//Proceedings of the 2019 IEEE International Conference on Industrial Internet. Orlando, USA, 2019: 227-232

[31] Zhang Jun-Wei, Ma Jian-Feng, Yang Chao, et al. Universally composable secure positioning in the bounded retrieval model. Science China Information Sciences. 2015, 58(11): 1-15

[32] Zhang Jun-Wei, Lu Ning, Ma Jian-Feng, et al. Universally composable secure geographic area verification without pre-shared secret. Science China Information Sciences, 2019, 62(3): 1-15

[33] Popa R A, Li F II, Zeldovich N. An ideal-security protocol for order-preserving encoding//Proceedings of the 2013 IEEE Symposium on Security and Privacy. San Francisco, USA, 2013: 463-477

[34] Ji Ya-Xian, Zhang Jun-Wei, Ma Jian-Feng, et al. BMPLS: Blockchain-based multi-level privacy-preserving location sharing scheme for telecare medical information systems. Journal of Medical Systems, 2018, 42(8): 1-13

[35] Pedersen T P. Non-interactive and information-theoretic secure verifiable secret sharing//Proceedings of the Advances in Cryptology-CRYPTO'91. Santa Barbara, USA, 1992: 129-140

[36] Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philosophical Transactions of the Royal Society of London, 1825, 115: 513-583

[37] Mechanical Turk. 2020-12-05. https://www.mturk.com/mturk/

[38] Paillier P. Public-key cryptosystems based on composite degree residuosity classes//Proceedings of the Advances in Cryptology-EUROCRYPT'99, International Conference on the Theory and Application of Cryptographic Techniques. Prague, Czech Republic, 1999: 223-238

GAO Sheng, Ph. D., associate professor. His research interests include blockchain technology and applications, data security and privacy computing.

CHEN Xiu-Hua, M. S. candidate. Her research interests include crowdsensing and blockchain.

ZHU Jian-Ming, Ph. D., professor. His research interests include digital economics, financial information security and blockchain.

YUAN Li-Ping, M. S. candidate. Her research interests include federate learning and blockchain.

MA Xin-Di, Ph. D., lecturer. His research interests include data security and privacy protection.

LIN Hui, Ph. D., professor. His research interests include blockchain, machine learning and mobile edge computing.

Background

The problem studied in this paper belongs to the worker selection problem in crowdsensing in the field of privacy and security. Based on the limited task budget, to ensure the task completion quality, worker selection is an important issue that cannot be ignored in crowdsensing. In crowdsensing, workers and task publishers focus on different requirements. On the one hand, workers expect to participate in sensing tasks without revealing location privacy and data privacy, and then obtain fair rewards in data quality evaluation. On the other hand, task publishers hope to select a set of appropriate workers under a limited task budget to obtain high-quality sensing data as much as possible. Hence, the different requirements of workers and task publishers need to be considered when designing worker selection schemes.

At present, there are some works that focus on worker selection in traditional crowdsensing and blockchain-based crowdsensing, respectively. For worker selection in traditional crowdsensing, most work relies on a centralized sensing platform to perform worker selection. However, since the centralized sensing platform has absolute control over worker selection, it is vulnerable to attacks and destruction. For worker selection in blockchain-based crowdsensing, existing work adopts the blockchain to construct crowdsensing system models to solve the security problems caused by centralization. However, they lack consideration of location privacy leakage or data privacy security in quality evaluation.

To tackle the above issues, this paper proposes a Location Privacy-preserving Worker Selection scheme under a limited budget for the blockchain-based crowdsensing system (LPWS). Firstly, combining order-preserving encryption and Merkle tree to provide workers with personalized location privacy protection. Second, under a limited task budget, LPWS takes location distance and reputation level as worker selection indicators and then models worker selection as the target optimization problem. Based on dynamic programming, LPWS can determine a set of suitable workers to maximize the possibility of obtaining high-quality sensing data. Finally, LPWS uses random disturbance and Pedersen's commitment to achieve the privacy and fairness in quality evaluation, and then based on evaluation results to perform remuneration payment and reputation update, thereby encouraging workers to actively provide high-quality data. Simulation results indicate that compared with related work, under the guarantee of task completion quality, LPWS not only realizes the safe and fair worker selection, but also ensures the privacy and fairness of quality evaluation, thereby increasing the participation enthusiasm of workers and task publishers.

This work is supported by the National Natural Science Foundation of China (Nos. 62072487, 61902290), the Beijing Natural Science Foundation (No. M21036), and the National Statistical Science Foundation of China (No. 2020LD01).