基于博弈分析的众包交通监测隐私保护机制

何云华[①②] 孙利民*[①②] 杨卫东[②] 李 红[②]
[①](西安电子科技大学计算机学院 西安 710071)
[②](中国科学院信息工程研究所北京市物联网安全重点实验室 北京 100093)

摘 要：众包交通监测利用移动终端上传的 GPS 位置信息实时感知交通状况，具有广阔的应用前景。然而，上传的 GPS 信息会泄露用户隐私。该文基于博弈论分析用户上传行为，提出隐私保护的优化上传机制。首先建立用户上传行为与路况服务质量和隐私泄露之间的关系，据此构建不完全信息博弈模型，以便分析用户上传行为；然后，根据用户上传博弈纳什均衡，提出用户终端可控的隐私保护优化上传机制。理论分析表明，该文提出的上传机制最大化用户效用，具有激励相容特性；通过真实数据实验验证，上传机制能够提高用户的隐私保护度，以及算法的激励相容特性。

关键词：众包监测；位置隐私；不完全信息博弈；均衡
中图分类号： TP309.2 **文献标识码**： A **文章编号**：1009-5896(2016)02-0340-07
DOI: 10.11999/JEIT150721

Enhancing Privacy Preserving for Crowdsourced Monitoring — A Game Theoretic Analysis Based Approach

HE Yunhua[①②] SUN Limin[①②] YANG Weidong[②] LI Hong[②]
[①](*School of Computer Science*, *Xidian University*, *Xi'an* 710071, *China*)
[②](*Beijing Key Laboratory of Internet of Things Security*, *Institute of Information Engineering*, *Chinese Academy of Sciences*, *Beijing* 100093, *China*)

Abstract: Crowdsourcing traffic monitoring is a promising application, which exploits ubiquitous mobile devices to upload GPS samples to obtain live road traffic. However, uploading the sensitive location information raises significant privacy issues. By analyzing the upload behavior of mobile users, this paper designs a privacy preserving traffic data collection mechanism. Using the relationships among the traffic service quality, privacy loss and the upload behavior, an incomplete information game is built to analyze the upload behavior of users. Based on the existence and uniqueness of Nash equilibrium in this game, a user-centric privacy preserving traffic data collection mechanism is proposed, which can maximize the utilities of users, and this mechanism has a feature of incentive compatible. Finally, the experimental results on real world traffic data confirm the effectiveness of privacy protecting and the feature of incentive compatible.

Key words: Crowdsourced monitoring; Location privacy; Incomplete information game; Equilibrium

1 引言

智能手机在人们生活中迅速发展，预计 2015 年全球智能手机用户持有量将达到 98.2 亿。智能手机大都嵌入 GPS、摄像头、加速度等传感器，能够采集人类活动和周围环境多类数据。利用大量的个人智能手机和无线网络基础设施，来收集和分析超大规模感知数据的方式，称为众包[1]。众包方式将颠覆传统的实时路况信息采集技术。例如，Google 服务器利用 Android 手机上的地图软件，获取 GPS 抽样信息，估计道路上的交通状况。相对于传统路况感知方式，如探测车辆、地感线圈、交通照相机等，众包交通监测具有低代价、无需部署、高覆盖区域的特点，有望成为未来交通监测的主要手段。

众包交通监测的广泛应用必须考虑以下 3 个方面：(1)用户隐私；(2)交通状况预测准确性；(3)用户上传激励机制。其中用户隐私与交通状况预测是一对矛盾体，即道路上上传用户越多，且每个用户上传次数越多，路况估计越准确；但是用户上传越

收稿日期：2015-06-15；改回日期：2015-09-17；网络出版：2015-11-19
*通信作者：孙利民 sunlimin@iie.ac.cn
基金项目：国家自然科学基金(61472418, 61202099)，中国科学院先导专项基金(XDA06040100)
Foundation Items: The National Natural Science Foundation of China (61472418, 61202099), The Strategic Priority Research Program of the Chinese Academy of Sciences (XDA06040100)

频繁，也越容易泄露自己的位置隐私。

然而，目前的众包交通监测系统，如文献[2,3]都采用匿名技术保护用户隐私。MONTJOYE 等人[4]和 CHRIS 等人[5]对移动用户轨迹的研究表明，尽管匿名消除了明显的标识符，但利用用户移动特性和上传数据的时空特性仍然能够跟踪用户。跟踪攻击防御方法可分为集中式和分布式两种。集中式防御由中心服务器对轨迹数据做处理，如减少记录数据[6]、添加噪声数据[7]、空间混淆[8]等。然而，一旦中心服务器受到攻击，用户隐私将会泄露[6]。分布式防御方法，不依赖于中心隐私服务器。Mix 域方法让用户进入 Mix 域时更新假名，不发送位置信息到服务器，使得攻击者难以区分 Mix 域中的用户，PPLANISAMY 等人[9]考虑车辆密度、Mix 形状、位置粒度、移动限制等因素，建立适应于道路网的 Mix 域模型。LIU 等人[10]提出了最优放置 Mix 域的方法。然而，Mix 域中的用户不上传位置信息，严重影响了路况估计质量。文献[11]提出节点在路段上设定的标记点更新位置，标记点的放置确保了节点隐私保护的最小需求距离，并且避免一些隐私特别敏感的位置。由于不同节点在同一位置的隐私级别不同，而且节点隐私随时间和位置变化，因此标记点很难满足所有用户隐私的要求。

本文兼顾路况服务质量和用户个性化隐私需求，基于博弈理论分析了终端用户上传行为与路况服务质量和隐私泄露之间的联系，提出了终端用户隐私保护的优化上传机制。由于用户通常不知道周围道路上其他用户的隐私保护度，因此我们采用不完全信息博弈对位置隐私泄漏建模。根据博弈纳什均衡存在性与唯一性和服务器提供的路况服务反馈信息，提出了移动用户的优化上传策略。该机制不依赖中心隐私服务器，每个用户权衡路况估计准确性和位置隐私，独立地决定是否上传。该机制在满足基本路况估计需求的情况下，最大化用户效用，为用户提供强隐私保护机制。

2　模型假定与问题描述

图 1 为实时路况监测系统的架构。移动终端用户实时将 GPS 抽样信息<location, speed, direction, timestamp>上传服务器；服务器根据用户上传信息估计当前交通状况，并为用户提供实时路况信息和导航服务。系统将路网分成单个路段，特别划分了直道、岔口、十字路口等 3 种典型路段。服务器对路况估计的准确性 Q，依赖于路段上移动终端用户上传的数量 k[11]。假定用户集 $P=\{1,2,\cdots,n\}$ 愿意提供 GPS 抽样，因为它们期望得到较好的 Q 值。由

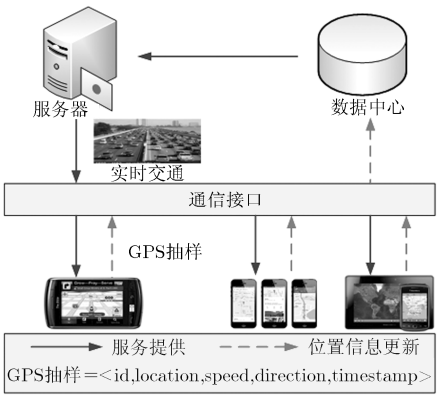


图 1 众包交通监测模型

于移动终端被不同的个体拥有，假定用户具有不同的隐私级 LP 是合理的，并且用户上传 GPS 抽样会带来一定的位置隐私损失 c。

2.1 路况服务

路况估计准确性 Q 与对应路段的上传用户数量 k 相关，k 越大，Q 值越大。通过真实数据[12]的实验发现，某一路段 i 上速度估计的均方差(RMS)与该路段的车辆上传数量 k_i 呈单调递减关系。令路段 i 上的路况估计准确性 $Q_i=1-1_{\mathrm{RMS}}$，其中 1_{RMS} 表示归一化的速度均方差，Q_i 与 k_i 的关系可表示为

$$Q_i=\log_\alpha(1+k_i\beta) \tag{1}$$

其中，α,β 的值可根据实验获取的 k_i 和 Q_i 值，通过 log 函数拟合得到。如图 2 所示，直道 $\alpha=5.949\times10^5$, $\beta=1.858\times10^4$，岔口 $\alpha=9.397\times10^5$, $\beta=1.855\times10^4$，十字路口 $\alpha=1.330\times10^6$, $\beta=1.855\times10^4$。

我们建立上传策略与服务质量之间的关系式。每个用户有 2 种可能的策略 s_i：上传(Y)或不上传(N)，根据式(1)，得到路况服务质量：

$$Q=\log_\alpha\left(1+\beta\sum_{i=1}^{n}I(s_i,Y)\right) \tag{2}$$

其中当 $x=y$ 时，$I(x, y)=1$，否则 $I(x, y)=0$。

2.2 位置隐私

2.2.1 跟踪攻击　攻击者从多个用户的匿名 GPS

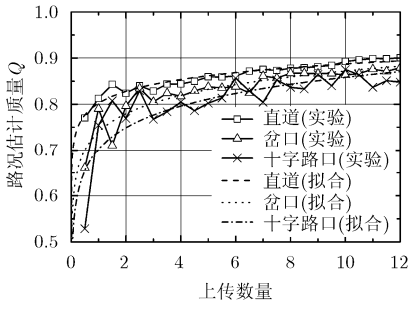


图 2　Log 函数拟合

抽样序列中，提取由同一用户产生的 GPS 抽样序列。攻击者关联先前的上传和下一次上传，找出最接近于预测或者最有可能的上传数据，可表示为[5]

$$\arg\max_{x} p(x \mid x_{i-1}) \tag{3}$$

其中，$p(x \mid x_{i-1})$ 表示为前一次上传位置为 x_{i-1}，本次上传位置为 x 的条件概率。

2.2.2 身份标识攻击 攻击者在给定的某条轨迹情况下，根据边信息[5]推断出发布该轨迹的用户。给定某条轨迹 L，可计算[5]

$$\arg\max_{i} \Pr(r_i(t_k), k = 1,2,\cdots \mid L) \tag{4}$$

其中，$r_i(t_k)$ 表示攻击者收集的关于用户 i 的边信息，也即用户 i 在时刻 t_k 的位置。

2.2.3 位置隐私量化 用户面临着跟踪攻击和身份标识攻击，因此位置隐私可量化为跟踪攻击不正确性[13]和身份标识的不确定性[14,15]。采用正确位置 x_i 与基于 $p(x \mid x_{i-1})$ 的估计值之间的期望距离，来量化跟踪攻击的不正确性(uncorrectness)[13]。表示如式(5)：

$$\sum_{x} p(x \mid x_{i-1}) \left\| x - x_i \right\| \tag{5}$$

当且仅当 $x = x_i$ 时，距离 $\left\| x - x_i \right\|$ 等于 0；否则，距离 $\left\| x - x_i \right\|$ 等于 1。

设 $p(P = \mathrm{ID}_i \mid\mid L)$ 为轨迹 L 的拥有者为 ID_i 的概率，采用分布 $p(P = \mathrm{ID}_i \mid L)$ 的熵来量化身份攻击的不确定性[15]，表示如式(6)：

$$H = \sum_{i} p(P = \mathrm{ID}_i \mid L) \log_2 \frac{1}{p(P = \mathrm{ID}_i \mid L)} \tag{6}$$

熵 H 给出了在 P 中查明唯一结果 ID_i 的难易程度。熵值越高，攻击者的确定性(certainty)越低。当 $p(P = \mathrm{ID}_i \mid L)$ 满足均匀分布时，熵值达到最大值 $\log_2 n$。

归一化得到，用户 i 在决定上传与否前的位置隐私：

$$\mathrm{LP}_i^- = \frac{1}{2}\left(\frac{H}{\log_2 n} + \sum_{x \in R} p(x \mid x_{i-1}) \left\| x - x_i \right\| \right) \tag{7}$$

用户上传数据会带来一定的位置隐私损失。假设用户 i 上传带来的隐私损失为 c_i，用户隐私度可表示为

$$\mathrm{LP}_i(s_i) = \begin{cases} \mathrm{LP}_i^- - c_i, & s_i = Y \\ \mathrm{LP}_i^-, & s_i = N \end{cases} \tag{8}$$

2.3 问题描述

给定某一路段上路况估计质量 Q 的最低需求 $Q_{\min}$ 和路段上每个用户 i 的隐私度 LP_i^-。优化目标是寻找上传方案 $s = (s_1, s_2, \cdots, s_n)$ 使得整体隐私 $\sum_i \mathrm{LP}_i(s_i)$ 最大化，且满足 $Q \geq Q_{\min}$。必须考虑如下两个因素：

(1)用户 i 可能不知道路段上其他用户的隐私度；

(2)如何估计路段上的最低服务质量需求 $Q_{\min}$。

针对因素(1)，我们采用不完全信息博弈[14]引入自然作为参与者，为路段上用户分配一个类型 θ，服从同一概率分布 $f(\theta)$，类型 θ 表示用户在该路段上的隐私度。每个用户知道路段上其他用户的隐私度分布，但不知道某个用户的隐私度。针对因素(2)问题，我们利用服务器的全局视角，根据路段类型和 GPS 抽样的平均速度，实时估算各路段的最低服务质量需求。虽然我们能够采用全局算法使得用户隐私度达到整体最优，但可能存在某个用户为增加其效用而改变上传策略，因此所设计的机制需满足激励相容特性[1]。

3 博弈建模

博弈论适合描述、预测和解释参与者的行为，逐渐被应用于解决移动网络中的安全与隐私问题[14,16,17]。黄等人[16]基于演化博弈研究无线网络物理层的安全协作方法。Freudiger 等人[14]采用不完全信息博弈，研究 Ad hoc 网络中的位置隐私保护机制 Mix-zone 中节点的非合作行为。SHOKRI 等人[17]采用 STACKELBERG 博弈优化设计 LBS 位置隐私保护机制。本文则采用不完全信息博弈论研究众包交通监测中终端用户的上传行为，优化设计终端用户的隐私保护上传机制。

参与者集合 $P = \{1,2,\cdots,n\}$ 对应于某一路段上的当前所有移动终端的用户集合。每个参与者有两种可能的策略 s_i：上传(Y)或不上传(N)。将用户 i 的收益函数定义为

$$\begin{aligned} & u_i(s_i(\theta_i), s_{-i}(\theta_{-i})) \\ & \quad = wQ_i(s_i(\theta_i), s_{-i}(\theta_{-i})) + \mathrm{LP}_i(s_i(\theta_i)) \end{aligned} \tag{9}$$

其中，服务质量 $Q_i(s_i, s_{-i})$ 由用户 i 和其对手 $-i$ 的策略决定，位置隐私 $\mathrm{LP}_i(s_i, s_{-i})$ 由用户策略对跟踪攻击不正确性和身份攻击的不确定性表示，参数 w 可视为用户对 Q 的期望。θ_i 是参与者类型，服从同一概率密度函数 $f(\theta_i)$，它表示博弈前参与者的隐私级，θ_i 决定其策略函数，表示为 $s_i : \theta_i \to \{Y, N\}$。

3.1 纳什均衡

由于用户上传博弈是不完全信息博弈，因此，首先引入节点上传行为博弈的贝叶斯纳什均衡[14]的概念：

定义 1 策略集 $s^* = \{s_i^*(\theta_i), s_{-i}^*(\theta_{-i})\}$ 是贝叶斯纯策略均衡，如果对每个参与者 i 有

$$s_i^*(\theta_i) \in \arg\max_{s_i \in \{Y,N\}} \sum_{\theta_{-i}} f(\theta_{-i}) u_i(s_i, s_{-i}^*(\theta_{-i})),\ \forall \theta_i \quad (10)$$

用户上传博弈的贝叶斯均衡，可以通过比较上传和不上传的平均收益得到，如式(11)所示。

$$\left.\begin{aligned} E[u_i(Y, s_{-i})] &= wE[Q(Y, s_{-i}(\theta_{-i}))] + \mathrm{LP}_i^- - c_i \\ E[u_i(N, s_{-i})] &= wE[Q(N, s_{-i}(\theta_{-i}))] + \mathrm{LP}_i^- \end{aligned}\right\} \quad (11)$$

当 $c_i < w(E[Q(Y, s_{-i}(\theta_{-i}))] - E[Q(N, s_{-i}(\theta_{-i}))])$ 时，上传(Y)为用户的纳什均衡策略；当 $c_i \geq w(E[Q(Y, s_{-i}(\theta_{-i}))] - E[Q(N, s_{-i}(\theta_{-i}))])$ 时，不上传(N)为用户的纳什均衡策略。

定义用户 i 上传的概率为 $q_i = \int_{\tilde{\theta}_i}^{1} f(\theta_i)\mathrm{d}\theta_i$，其中 $\tilde{\theta}_i$ 为用户 i 上传所需的最低隐私级。设 P_Y 为元素个数为 k 的上传用户集，那么上传用户数等于 k 的概率可表示为 $\Pr(K=k) = \prod_{i \in P_Y} q_i \prod_{j \in P - P_Y} (1 - q_j)$，从而得到期望的路况估计准确性 Q 为

$$E(Q) = \sum_{k=1}^{n} \Pr(K=k) \log_\alpha (1 + \beta k) \quad (12)$$

且存在 $\hat{k}$，使得 $E(Q) \approx \log_\alpha(1 + \hat{k}\beta)$。近似得到

$$E[Q(Y, s_{-i}(\theta_{-i}))] - E[Q(N, s_{-i}(\theta_{-i}))] \approx \log_\alpha \frac{1 + \beta\left(1 + \hat{k}\right)}{1 + \beta\hat{k}} \quad (13)$$

即 $w(\log_\alpha(1 + \beta(1 + \hat{k})) - \log_\alpha(1 + \beta\hat{k}))$ 为用户的上传均衡阈值。

4　上传机制

本节设计强隐私保护的位置信息上传机制 UploadGame，不仅提供需求的路况估计质量 Q，而且提供强隐私保护 LP，达到用户隐私保护度和路况质量的整体最优性。该机制利用服务器为用户提供路况质量需求反馈，实时调整路况估计质量，避免了因节点自由决定是否上传不能满足路况估计需求的缺陷。另外，该算法以用户为中心，节点能够自主决定是否上传，满足其个性化的隐私需求。UploadGame 算法分为 2 个阶段，服务器反馈阶段和节点选择上传阶段。

(1)服务器反馈阶段：服务器根据历史交通状况，确定当前路况估计需求，为路段上用户提供上传数量需求的反馈。服务器首先计算最低路况估计质量需求 $Q_{\min}$，然后由式(4)可求得需要上传节点数 $k = (\alpha^{Q(v)} - 1)/\beta$。假定 $Q_{\min}$ 与历史估计的平均速度 v 之间的关系满足正态分布，那么 $Q_{\min}$ 可由式(14)给出：

$$Q_{\min} = Q(v) = \frac{\rho}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(v-\mu)^2}{2\sigma^2}} \quad (14)$$

其中 $\rho > 0$ 为系统参数，μ 和 σ 分别表示速度 v 的平均值和标准差。式(14)反映了节点速度很低时，路况估计质量需求不高，这是由于拥塞时速度不会产生很大的变化；速度很高，路况估计质量需求也不高，因为此时交通顺畅，不需要进行精确估计。当然也可采用其他函数关系进行估计，如用户更关心交通事故发生时，可采用 Q 与 Δv 之间的函数关系来捕捉路况变化。

(2)节点选择上传阶段：节点基于服务器的反馈，计算隐私损失阈值 $c_{\max}$，然后根据该阈值来决定是否上传。如果节点知道其他节点的隐私损失，如 $c_1 \leq c_2 \leq \cdots \leq c_n$，容易得到隐私损失阈值 $c_{\max} = c_k$。然而，节点通常不知道其他节点的隐私度和隐私损失，因而需要估计出 c_k 的值。假定隐私度与隐私损失满足 $\mathrm{LP}_i = \lambda / c_i$，该式反映了用户隐私保护度越高其上传隐私损失越小，也即长时间未上传的节点上传所带来的隐私损失低。由节点隐私度服从概率分布 $f(\theta_i)$ 可得到

$$\frac{k}{n} = \int_{\underline{\theta}}^{1} f(\theta_i)\mathrm{d}\theta_i \quad (15)$$

其中 n 是周围车辆辆数。根据式(15)，求得 $\underline{\theta} = F^{-1}(1 - k/n)$，从而得到 c_k 的估计值：

$$\hat{c}_k = \lambda / \underline{\theta} = \frac{\lambda}{F^{-1}(1 - k/n)} \quad (16)$$

算法 1　UploadGame 算法。

//阶段 1　服务器确定上传用户数量 k

(1)根据上次估计的平均速度计算服务质量需求。根据式(14)得到服务质量需求 $Q_{\min} = Q(v) = \frac{\rho}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(v-\mu)^2}{2\sigma^2}}$。

(2)计算上传用户数量 k。根据式(1)以及当前的服务质量需求 $Q(v)$，计算 $k = (\alpha^{Q(v)} - 1)/\beta$。

//阶段 2　用户上传决策

(3)根据式(16)以及阶段 1 求得的上传用户数 k，计算上传阈值 $c_{\max} = \lambda / F^{-1}(1 - k/n)$。

(4)如果 $c_i < c_{\max}, s_i = Y$。如果节点隐私损失 c_i 小于阈值，节点上传位置信息。

(5)else $s_i = N$。否则节点不上传位置信息。

4.1 博弈分析

4.1.1 纳什均衡　纳什均衡的存在性和唯一性，能够说明算法 1 的收敛性，由定理 1 给出。

定理 1　算法 1 存在唯一的纳什均衡。

证明　设用户的隐私损失满足 $c_1 \leq c_2 \leq \cdots \leq c_n$，$s^* = \{Y^1, Y^2, \cdots, Y^k, N^1, \cdots, N^{n-k}\}$ 为通过算法 1 得到的用户上传结果，先证纳什均衡存在性。设 $g(x) = \log_\alpha(1 + \beta x) - \log_\alpha(1 + \beta(x-1))$ 在[1,n]上连续，其

一阶导数为

$$g'(x)=\frac{-\beta^2}{(1+\beta x)(1+\beta(x-1))\ln\alpha}<0 \tag{17}$$

因此函数 $g(x)$单调递减。设 $c(x)$为[1,n]上的连续函数，且满足 $c(i)=c_i$，$c'(x)\geq 0$。函数 $G(x)=wg(x)-c(x)$ 的一阶导数 $G'(x)=wg'(x)-c'(x)<0$。由算法 1 得到 $G(1)<0<G(n)$，根据介值定理，存在 $\tilde{k}$，使得 $G(\tilde{k})=0$。

当 $i\leq k=\lfloor\tilde{k}\rfloor$ 时，有

$$\begin{aligned}&u_i(Y,s_{-i})-u_i(N,s_{-i})\\&=w\log_\alpha\frac{1+\beta\tilde{k}}{1+\beta(\tilde{k}-1)}-c_i\\&=w\log_\alpha\frac{1+\beta\tilde{k}}{1+\beta(\tilde{k}-1)}-c_{\tilde{k}}+c_{\tilde{k}}-c_i\\&=G(\tilde{k})+(c_{\tilde{k}}-c_i)>0\end{aligned} \tag{18}$$

所以当 $i\leq k$ 时，上传总为优势策略。

当 $i>k=\lfloor\tilde{k}\rfloor$ 时，有

$$\begin{aligned}&u_i(Y,s_{-i})-u_i(N,s_{-i})\\&=w\log_\alpha\frac{1+\beta(\tilde{k}+1)}{1+\beta\tilde{k}}-c_i\\&=G(\tilde{k}+1)+(c_{\tilde{k}+1}-c_i)\leq 0\end{aligned} \tag{19}$$

所以不上传总为最优策略。

综上所述，$s^*=\{Y^1,Y^2,\cdots,Y^k,N^1,\cdots,N^{n-k}\}$ 是上传博弈的纳什均衡。

再证唯一性。假设 $s^*=\{Y^1,Y^2,\cdots,Y^h,N^1,\cdots,N^{n-h}\}$ 也为纳什均衡，根据算法 1，有 $h\leq k$。当 $h<k$ 时，总存在 c_{h+1} 满足 $c_{h+1}<w\log_\alpha\frac{1+\beta(1+\tilde{k})}{1+\beta\tilde{k}}$，也就是说存在用户 h+1 可以改变其策略来增加效用，不能满足纳什均衡的条件。从而证明了 $s^*=\{Y^1,Y^2,\cdots,Y^k,N^1,\cdots,N^{n-k}\}$ 的唯一性。

4.1.2 激励相容 下面将证明算法 1 具有激励相容特性[1]。激励相容是机制设计中最基本的概念，保证参与者根据真实类型作出决策是其最优反应策略，消除用户对其他用户操纵市场给其带来较大损失的担忧。

定义 2 一个机制是激励相容的，如果直接显示机制存在纯策略均衡 $s^*(\cdot)=(s_1^*(\cdot),\cdots,s_n^*(\cdot))$，其中 $s_i^*(\theta_i)=\theta_i,\forall\theta_i\in\Theta_i,\forall i\in N$。

换句话说，每个用户通过报道真实隐私损失，能够最大化其效用。

定理 2 算法 1 具有激励相容特性。

证明 设节点隐私损失为 c_i，节点选择隐私损失 c_i'。当 $c_i<w\log_\alpha\frac{1+\beta(k+1)}{1+\beta k}$ 时，如果 $c_i'<c_i$ 或 $c_i<c_i'<w\log_\alpha\frac{1+\beta(k+1)}{1+\beta k}$，将不会影响用户的上传策略，其效用不变；如果 $c_i\geq w\log_\alpha\frac{1+\beta(k+1)}{1+\beta k}$，节点将不上传，其效用

$$\begin{aligned}u_i(N,s_{-i}^*)&=w\log_\alpha(1+\beta k)+\mathrm{LP}_i^-\\&\leq w\log_\alpha(1+\beta(k+1))+\mathrm{LP}_i^--c_i\\&=u_i(Y,s_{-i}^*)\end{aligned} \tag{20}$$

此时节点能够改变其策略来增加效用，因此节点谎报隐私损失会降低其效用。

当 $c_i\geq w\log_\alpha\frac{1+\beta(k+1)}{1+\beta k}$ 时，如果 $c_i'>c_i$ 或 $c_i>c_i'\geq w\log_\alpha\frac{1+\beta(k+1)}{1+\beta k}$，用户将不上传，其效用不变；如果 $c_i'<w\log_\alpha\frac{1+\beta(k+1)}{1+\beta k}$，节点将上传，其效用

$$\begin{aligned}u_i(Y,s_{-i}^*)&=w\log_\alpha(1+\beta(k+1))+\mathrm{LP}_i^--c_i\\&<w\log_\alpha(1+\beta k)+\mathrm{LP}_i^-u_i(N,s_{-i}^*)\end{aligned} \tag{21}$$

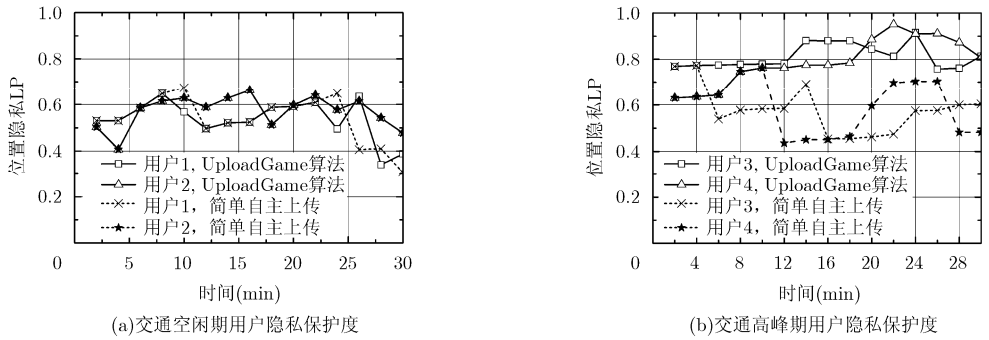
此时节点能够改变其策略来增加效用，因此节点谎报隐私损失会降低其效用。

综上所述，算法 1 具有激励相容特性。

5 仿真实验

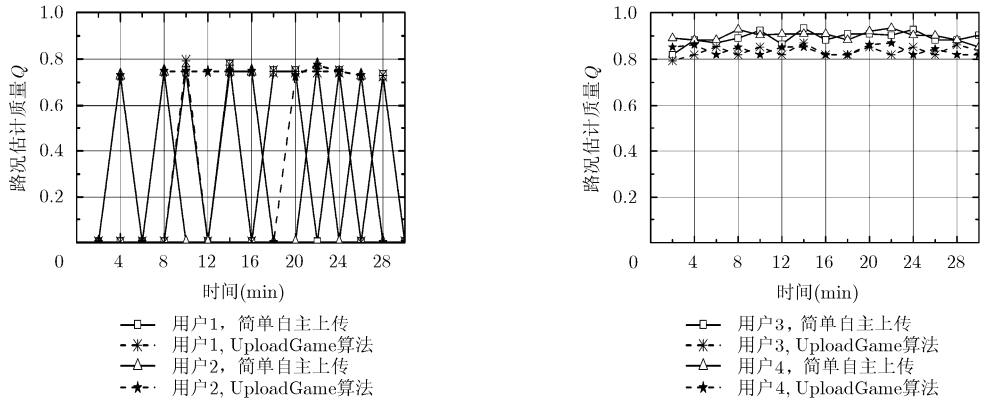
本文采用北京 2009 年 3 月出租车收集的 GPS 轨迹数据[12]作为仿真数据集，选取交通高峰期和交通空闲期 2 种场景。我们将研究的地理区域递归分为 4 个方格，每个方格用三元组<level, x, y>表示，其中 level 表示递归层数，x,y 分别表示相对左上角的偏移量。这里我们取 level=3，也即将区域分为 8×8 个块。

我们比较了 UploadGame 与简单的自主上传机制在隐私保护和路况估计质量上的性能。在简单的自主上传机制中，用户根据具有固定 w 值的上传阈值来决定是否上传。在交通空闲期，用户 1 沿着路径 1 行驶从<3, 8, 1>到<3, 2, 7>，用户 2 沿着路径 2 行驶从<3, 1, 8>到<3, 8, 3>；在交通高峰期，用户 3 沿着路径 1 行驶，用户 4 沿着路径 2 行驶。用户隐私级和路况估计质量每隔 2 min 计算一次，结果在图 3、图 4 中给出。观察到，用户隐私级在交通高峰期比交通空闲期要高，因为交通高峰期有更多的用户在道路上，增加了跟踪和身份识别的难度。正如所期望的，在交通高峰期，UploadGame 机制下用户的隐私保护度高于简单机制(>25%)；在交通空闲期，UploadGame 机制提供的隐私保护度与简单机制相当。图 3 中的突然上升的点表示用户正经



(a)交通空闲期用户隐私保护度　　(b)交通高峰期用户隐私保护度

图 3 不同交通场景下用户的隐私级



(a)交通空闲期路况估计质量　　(b)交通高峰期路况估计质量

图 4 不同交通场景下用户获取的路况估计质量

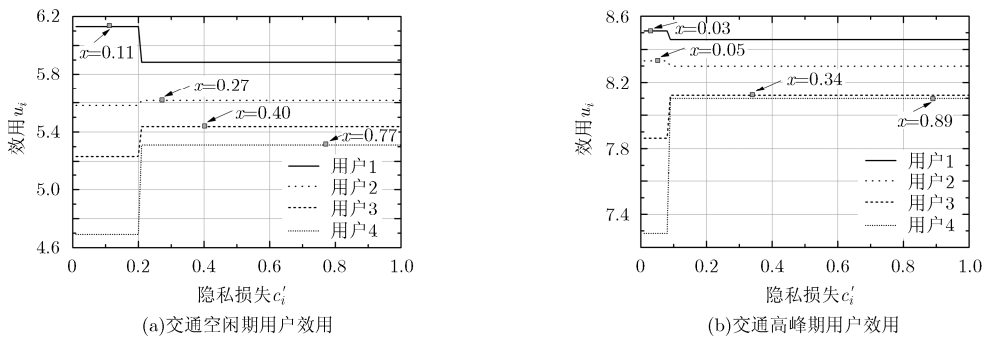
过高度混淆区域，如十字路口；突然下降点，表示用户在此时上传了自己的位置，其隐私泄露风险增加。

图 4 阐述了图 3 对应的交通高峰期和交通空闲期，用户获得的路况估计质量变化。交通高峰期的路况估计质量高于交通空闲期，因为交通高峰期更多用户上传。虽然交通空闲期，路况估计质量达不到 83%，但是此时用户仅关心交通是空闲的，不关心估计的准确性。但为了处理一些意外事件，UploadGame 机制鼓励用户上传。因此图 4 中的 UploadGame 机制的路况估计质量高于简单机制。在交通高峰期，UploadGame 机制能够满足最低路况估计质量需求(QoS=0.83)。

我们还验证了 UploadGame 机制的激励相容特性。分别在交通空闲和交通高峰期的<3,3,4>块中随机选取了 4 个用户，并允许隐私损失不同于其真实隐私损失。如图 5 所示，用户根据真实隐私损失进行决策时，其效用达到最优。图 5 中标记点是用户真实隐私损失所对应的效用值，从图中可以看出用户选取其他隐私损失时，不能增加其效用。

6 总结

本文基于博弈论分析设计众包监测系统的隐私



(a)交通空闲期用户效用　　(b)交通高峰期用户效用

图 5　UploadGame 的激励相容特性

保护机制，满足基本路况估计质量的同时，提高用户隐私保护度，最大化用户效用。首先建立用户上传行为与路况服务质量和隐私泄露之间的关系，构建终端用户上传博弈模型；然后，根据博弈纳什均衡，设计用户终端可控的隐私保护优化上传机制 UploadGame，理论分析该机制的收敛性和激励相容特性。最后通过真实数据的仿真实验分析 UploadGame 的性能优势，验证了 UploadGame 的激励相容特性。

参 考 文 献

[1] YANG D J, XUE G L, FANG X, *et al*. Incentive mechanisms for crowdsensing: crowdsourcing with smartphones[J]. *IEEE/ACM Transactions on Networking*, 2015, 99: 1–13.

[2] GAO S, MA J F, SHI W S, *et al*. TrPF: A trafectory privacy-preserving framework for participatory sensing[J]. *IEEE Transactions on Information Forensics and Security*, 2013, 8(6): 874–887.

[3] MOHAN P, PADMANABHAN V N, and RAMJEE R. Nericell: rich monitoring of road and traffic conditions using mobile smartphones[C]. Proceedings of the ACM Conference on Embedded Networked Sensor Systems, North Carolina, 2008: 323–336.

[4] MONTJOYE Y A, HIDALGO C A, VERLEYSEN M, *et al*. Unique in the crowd: The privacy bounds of human mobility[R]. Nature Science Report, Cambridge, 2013.

[5] CHRIS M, DAVID Y, and NUNG Y. Privacy vulnerability of published anonymous mobility traces[J]. *IEEE Transactions on Networking*, 2013, 21(3): 720–733.

[6] 孙利民, 李红, 王笑寒, 等. 物联网位置隐私保护综述[J]. 软件学报, 2014, 25(s1): 1–10.
SUN Limin, LI Hong, WANG Xiaohan, *et al*. Survey on the location privacy preservation in the internet of things[J]. *Journal of Software*, 2014, 25(s1): 1–10.

[7] SHI E, CHAN T H, RIEFFEL E, *et al*. Privacy-preserving aggregation of time-series data[C]. Proceedings of 18th Network & Distributed System Security Symposium, California, 2011: 1–17.

[8] HOH B, GRUTESER M, XIONG H, *et al*. Achieving guaranteed anonymity in gps traces via uncertainty-aware path cloaking [J]. *IEEE Transactions on Mobile Computing*, 2010, 9(8): 1089–1107.

[9] PALANISAMY B and LIU L. Attack-resilient mix-zones over road networks: architecture and algorithms[J]. *IEEE Transactions on Mobile Computing*, 2015, 14(3): 495–508.

[10] LIU X X, ZHAO H, PAN M, *et al*. Traffic-aware multiple mix zone placement for protecting location privacy[C]. Proceedings of the 31rd Annual IEEE International Conference on Computer Communications, Florida, 2012: 972–980.

[11] HOH B, IWUCHUKWU T, JACOBSON Q, *et al*. Enhancing privacy and accuracy in probe vehicle based traffic monitoring via virtual trip lines[J]. *IEEE Transactions on Mobile Computing*, 2012, 11(5): 849–864.

[12] DATATANG Company. Taxi GPS data of one city in north of china (200903)[OL]. http://dx.doi.org/10.1287/mnsc.1040.0270.2014.12.

[13] SHOKRI R, THEODORAKOPOULOS G, BOUDEC J L, *et al*. Quantifying location privacy[C]. Proceedings of IEEE Symposium on Security and Privacy, California, 2011: 247–262.

[14] FREUDIGER J, MANSHAEI M H, HUBAUX J P, *et al*. Non-cooperative location privacy[J]. *IEEE Transactions on Dependable and Secure Computing*, 2013, 10(2): 84–98.

[15] 葛国栋, 郭云飞, 刘彩霞, 等. 内容中心网络中面向隐私保护的协作缓存策略[J]. 电子与信息学报, 2015, 37(5): 1220–1226. doi: 10.11999/JEIT140874.
GE Guodong, GUO Yunfei, LIU Caixia, *et al*. A collaborative caching strategy for privacy protection in content centric networking[J]. *Journal of Electronics & Information Technology*, 2015, 37(5): 1220–1226. doi: 10.11999/JEIT-140874.

[16] 黄开枝, 洪颖, 罗文宇, 等. 基于演化博弈机制的物理层安全协作方法[J]. 电子与信息学报, 2015, 37(1): 193–199. doi:10.11999/JEIT140309.
HUANG Kaizhi, HONG Ying, LUO Wenyu, *et al*. A method for physical layer security cooperation based on evolutionary game[J]. *Journal of Electronics & Information Technology*, 2015, 37(1): 193–199. doi:10.11999/JEIT140309.

[17] SHOKRI R, THEODORAKOPOULOS G, TRONCOSO C, *et al*. Protecting location privacy: Optimal strategy against localization attacks[C]. Proceedings of 19th ACM Conference on Computer and Communications Security, California, 2012: 617–627.

何云华： 男，1987 年生，博士生，研究方向为车载网安全与隐私.

孙利民： 男，1966 年生，研究员，研究方向为物联网安全、工业控制安全及位置隐私等.

杨卫东： 男，1977 年生，副教授，研究方向为车载网安全.

李 红： 男，1989 年生，博士生，研究方向为位置隐私.