

面向智慧矿山的移动群智感知覆盖及能效优化

王朝炜, 刘婷, 王天宇, 王卫东

(北京邮电大学电子工程学院, 北京 100876)

摘要: 移动群智感知 (MCS, mobile crowd sensing) 是一种有效利用智能移动终端协同采集环境数据的技术, 集成多种传感器的移动载体 (如车辆) 越来越多地被当作参与者来承担各种感知任务。在智慧矿山物联网 (IoT, Internet of things) 中, 为了更好地感知人一机一环的实时信息, 支撑安全生产顺利进行, 基于 MCS 思想对矿山环境下移动感知节点的覆盖质量和能耗优化进行研究, 提出了一种综合考虑覆盖率 (CP, coverage percentage) 和覆盖密度 (CD, coverage density) 的效用函数 $F(\Omega)$ 来衡量 MCS 的覆盖质量。为了获得最优的覆盖质量, 针对参与感知的车辆选择问题提出了一种改进的贪婪算法——覆盖质量优化 (CQO, coverage quality optimization) 算法来优化覆盖质量, 并使用真实的车辆轨迹数据集对所提出的算法进行评估, 研究了影响覆盖质量的几个因素。实验结果表明, 该算法具有较好的覆盖质量。

关键词: 智慧矿山; 移动群智感知; 覆盖率; 覆盖密度

中图分类号: TN92

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2020.00191

Mobile crowd sensing coverage and energy optimization in smart coalmine

WANG Chaowei, LIU Ting, WANG Tianyu, WANG Weidong

School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract: Mobile crowd sensing (MCS) is a promising diagram for the environmental information collection based on the smart mobile terminal. Nowadays, vehicles with multiple embedded sensors are increasingly being considered as participants to complete various sensing tasks. In order to better perceive the data in the coalmine environment, the coverage quality and energy consumption of the perception data of sensing terminals were studied based on MCS. A new sensing coverage quality indicator called utility function $F(\Omega)$ jointly considering the coverage percentage and coverage density was defined. The selection of vehicles as an optimization problem to improve the coverage quality was formulated, then an improved greedy algorithm—coverage quality optimization (CQO) algorithm was proposed. The proposed algorithm with the real vehicle trajectory dataset was formulated and several factors influencing the coverage quality were studied. The experiment results show that the proposed algorithm achieves a better coverage quality.

Key words: smart coalmine, mobile crowd sensing, coverage percentage, coverage density

1 引言

我国的能源结构中, 煤炭资源的消耗占了很大的比例^[1]。然而, 矿井工作环境恶劣, 煤矿开采

力度的加大使得开采作业愈发困难, 保障井下开采作业过程中的人身安全是重中之重。为了保障井下人员、设备的安全和采矿工作的顺利进行, 采用了多种通信和网络技术。如在矿井中安装有有线电话系

收稿日期: 2020-04-27; 修回日期: 2020-06-08

通信作者: 王卫东, wangweidong@bupt.edu.cn

基金项目: 国家重点研发计划 (No.2017YFC0804404)

Foundation Item: The National Key R&D Program of China (No.2017YFC0804404)

统^[2],可以加快信息调度的效率;通过建设基站,无线通信网络和有线电话系统有效连接,能实现无线通信、信息调度等各项基本功能;矿井网络互联,对煤矿井下电力设备、瓦斯浓度、通风情况进行实时监控,及时预警灾害情况,达到降低井下事故发生概率的目的。随着移动互联网、IoT和无线传感器网络(WSN, wireless sensor network)等技术的快速发展,IoT在矿山上的应用也得到大规模拓展,形成了面向人一机一环多元信息采集、处理与交互的矿山IoT体系。矿山IoT主要以工业以太网为骨干网,承载视频业务、语音业务、控制命令和传感数据等的传输,向上支撑态势评估、安全生产等业务。在矿山环境下联网节点的数量呈指数级增长,所感知的数据呈海量、多源、异构等特点。

将海量的感知数据经高速互联网连接,可以实时与云端进行信息交互。但是将收集的全部海量数据传输到云端是不切实际的,不仅操作难度大,而且会造成相当多的资源浪费^[3]。加快对矿山无线网络数据感知收集技术的探究迫在眉睫,因此,为了高效地收集矿山环境的感知数据,可以引入MCS技术来改善移动感知节点的覆盖性能,提升信息收集的能力。

MCS技术与传统的WSN相比具有很多优势,如MCS在收集数据的过程中,不需要事先在目标区域部署相关的感知设备,充分利用矿山车辆移动范围广而灵活的特性,将其作为嵌入多种传感器的移动载体,对矿山环境中的感兴趣区域(ROI, regions of interest)进行多元数据的感知和收集^[4-5]。而传统的WSN为了感知、收集ROI的数据,需要在所有ROI部署多种传感器,成本与效果受诸多因素限制。MCS利用动态选择的车辆进行数据的收集,可以获得较好的覆盖质量。而且,由于车辆的维护性,不易出现类似WSN中传感器节点能量耗尽导致的数据收集失败的问题。

文献[6]旨在减少感知数据的冗余,通过选择最少的车辆来达到最大的时空覆盖度(STC, spatial-temporal coverage)。文献[7]将MCS服务平台与候选车辆之间的交互建模为马尔可夫决策过程(MDP, Markov decision process),使用深度强化学习最大化STC。由于仅根据当前位置选择参与车辆,因此,仿真得到的覆盖效果不一定达到最优。文献[8]提出了一种基于预测轨迹的车辆选择算法。文献[4,9-10]通过路径规划或路径分段的方法完成

感知任务。Gong等^[9]设置了车辆行驶距离预算,提出了一种基于遗传的路由(GR, genetic-based routing)算法,并根据GR算法设计了4种在线任务分配算法。文献[11-13]从经济角度解决了MCS的激励机制问题,其中文献[12]考虑了MCS中平台、参与者和任务发起者3方面的利益。此外,为了平衡平台与参与者之间的效用,文献[14]提出了一个以用户为中心的基于声誉的招募策略。文献[15]介绍了多臂赌博机(MAB, multi-armed bandit)模型并使用有界 ϵ -first算法,有界 ϵ -first算法将总预算 B 中的 ϵB 部分用来获取评估质量特性,而剩余的 $(1-\epsilon)B$ 部分则用于优化效用函数。在保证服务质量(QoS, quality of service)方面,文献[16]提出了一种GR算法,文献[17]采用了矩阵补全技术。文献[18]设计了一种高效节能的MCS任务分配框架。文献[19-21]讨论了感知过程中的能耗问题和感知数据的质量问题。

本文建立了一个MCS数据模型,主要提出CP和CD两个概念来衡量车辆对ROI的覆盖质量。使用真实的车辆轨迹数据集进行仿真评估,结果表明,所提出的CQO算法可以获得较高的覆盖质量。同时,研究了成本预算、候选车辆数量和感知次数等变量对覆盖质量的影响。

2 智慧矿山MCS系统模型

2.1 MCS架构背景

本文提出的智慧矿山MCS系统模型如图1所示,此系统由一个云平台 and 配备多个传感器的车辆组成,地面场景如图1(a)所示,井下场景如图1(b)所示。云平台的工作是将感知任务分发给候选车辆,并根据本文提出的策略选择合适的车辆来执行感知任务。当车辆完成感知任务后,将感知数据上传到云平台。

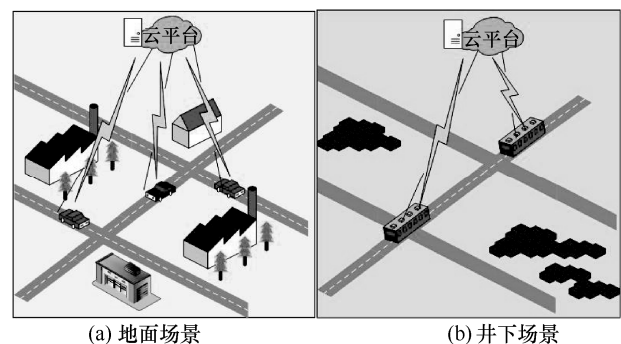


图1 智慧矿山MCS系统模型

需要指出的是，在现实中，不同区域的感知需求是不同的。某些感知任务的感知结果是“是”或“否”的二值（0 或 1）类型，如果只根据一辆车的感知数据来判断得出结论，会造成较大的偏差，因此，收集具有一定冗余度的数据是相当有必要的。为了提高相同位置感知数据的准确度，需要进行多次采集以保证数据的准确性。

2.2 MCS 系统模型

基于 MCS 的场景，本文将目标区域划分为 10×10 的网格单元，区域重要度示意图如图 2 所示。文献[22]以矿工为主要检测目标，井下矿工尤其是某些区域的矿工并非均匀分布在巷道或井下其他位置，也就是说 ROI 可分为热点区域和非热点区域（普通区域）。在图 2 中，A、G、H、J 为权重为 2 的热点区域，其他区域（B、C、D、E、F、I、K、L）为权重为 1 的非热点区域，此处权重值的设定主要是为了区分热点区域和非热点区域对覆盖质量的影响。若将热点区域的权重值设置为 10，非热点区域权重值设置为 1，10:1 的比例会对本文提出的覆盖质量造成过大的影响；若将热点区域和非热点区域的权重比例设置得过小，则无法显示本文提出的非均匀覆盖问题以及无法表示热点区域较非热点区域可以对覆盖质量做出更大的贡献。借鉴了文献[23]中视觉敏感度函数模型的思想，视觉对比敏感度反映了人类视觉对不同频率的图像信息接受程度的不同，即人类视觉对分布在不同频带上的图像信息具有不同的视觉敏感程度。本文为了区分不同区域的重要程度，同时保证不同区域的权重不至于失衡，所以设置热点区域和非热区域权重比例为 2:1。

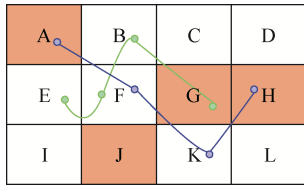


图2 区域重要度示意图

W_i 表示 i 区域的权重值，则有

$$\begin{aligned} W_A = W_G = W_H = W_J = 2 \\ W_B = W_C = W_D = W_E = W_F = W_I = W_K = W_L = 1 \end{aligned} \quad (1)$$

图 2 中的蓝线、绿线代表两辆车的轨迹，轨迹上的点代表每一次记录时车辆的地理位置。本文假设有 m 辆车，由集合 $V = \{v_1, v_2, v_3, \dots, v_m\}$ 表示，在 n 个时间段 $T = \{t_1, t_2, t_3, \dots, t_n\}$ 内运动。

2.2.1 车辆轨迹

车辆轨迹 $L(V)$ 为 $m \times n$ 矩阵，表示为

$$L(V) = \begin{bmatrix} l_{11} & l_{12} & \dots & l_{1n} \\ l_{21} & l_{22} & \dots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ l_{m1} & l_{m2} & \dots & l_{mn} \end{bmatrix} \quad (2)$$

$$R = \{r_1, r_2, r_3, \dots, r_K\}$$

其中， R 为 ROI 的集合， K 为 ROI 的数量， l_{ij} 代表区域 R 内 v_i 在时间 t_j 的位置， $l_{ij} \in R$ 。根据车辆在各时刻的经纬度确定车辆的位置，构造候选车辆 $L(V)$ 的轨迹矩阵。 $L(V)$ 的行表示车辆在感知时间内的行驶轨迹，列表示某一时间段内所有候选车辆的位置。

2.2.2 覆盖次数的统计矩阵

$P(V)$ 是候选车辆到达 ROI 的次数的统计矩阵，表示为

$$P(V) = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1k} \\ p_{21} & p_{22} & \dots & p_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \dots & p_{mk} \end{bmatrix} \quad (3)$$

其中， $P(V)$ 的行表示车辆到达 k 个区域的次数统计值。以图 2 中两辆车的轨迹为例，简单起见，假设车辆从左向右行驶。 $L(v_1) = [A F K H]$ ， $L(v_2) = [E F B G]$ ，时间段为 $t_1 \sim t_4$ ，并且有

$$R = \{A, B, C, D, E, F, G, H, I, J, K, L\}$$

$$L(v_1, v_2) = \begin{bmatrix} A & F & K & H \\ E & F & B & G \end{bmatrix} \quad (4)$$

根据 $P(V)$ 的定义， $P(v_1) = [1 0 0 0 0 1 0 1 0 0 1 0]$ ， $P(v_2) = [0 1 0 0 1 1 1 0 0 0 0 0]$ ，则有

$$P(v_1, v_2) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (5)$$

设一个变量 φ 表示候选车辆 v_i 是否被选中，即

$$\varphi_i = \begin{cases} 1, & v_i \in \Omega \\ 0, & \text{其他} \end{cases} \quad (6)$$

其中， Ω 是从候选车辆集合 V 中被选出的车辆的集合， $\Omega \subseteq V$ 。

将矿工聚集的区域视作热点区域，对于热点区域，要进行重点的数据收集；而对于非热点区域，如果收集的数据量和热点区域的数据量一样，不仅

会造成数据的冗余，通信开销增大，还会过度消耗额外的能量。所以本文假设条件之一就是ROI进行非均匀覆盖。一方面减少能源的消耗，另一方面实现对热点区域的重点覆盖，同时对ROI整体的覆盖范围进行评估。本文提出了一个新的指标来衡量覆盖质量，考虑了CP和CD。其中，CD可以看作所选车辆到达热点区域而不是所有ROI的平均次数。

2.2.3 CD

CD的计算式为

$$CD = \frac{\sum_{i=1}^{t=n} \sum_{v_i \in \Omega} p'_{ij}}{K_1} \quad (7)$$

其中， K_1 表示ROI中热点区域的数量， p'_{ij} 表示选中车辆 v_i 到达热点区域的次数， $v_i \in \Omega$ ，并且有

$$R = \begin{Bmatrix} A & B & C & D & E & F & G & H & I & J & K & L \end{Bmatrix}$$

$$P(v_1) = \{ \begin{matrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{matrix} \} \quad (8)$$

$$P(v_2) = \{ \begin{matrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \end{matrix} \}$$

CD是车辆到达热点区域的平均累积次数(并非所有ROI)。为了便于说明，图2中的热点区域为A、G、H、J，所以 $K_1=4$ ， $K=12$ 。为了更清楚地表达CD的含义，相应的 p'_{ij} 显示在式(8)的实线框中。

通过计算可知， $CD(v_1, v_2) = (1+0+1+0+0+1+0+0)/4 = 0.75$ ，则在 $t_1 \sim t_4$ 期间， v_1 、 v_2 到达热点区域的平均次数为0.75。

CD的引入是为了表示所选车辆在热点区域覆盖的情况，对热点区域覆盖的次数越多，则所收集的数据的准确度也会越高，即CD越大热点区域的数据也就越可信。文献[21]研究了WSN节点对区域的覆盖情况的影响，当至少存在 K 个节点覆盖该区域时，视为该区域被有效覆盖，研究表明，只有较少的区域可以满足有效覆盖的要求。文献[21]设计了一种离线计算的过程来达到选择最少的节点数来实现有效覆盖区域的CP要求，这一点和本文提出的问题不谋而合，因为本文考虑的热点区域覆盖次数问题和文献[21]提出的有效覆盖是同样的问题，本文考虑的热点区域的有效覆盖用CD表示。

2.2.4 CP

CP的计算式为

$$CP = \frac{\sum_{v_i \in \Omega} P_{ij} \times W_k}{K} \quad (9)$$

其中， W_k 表示ROI中每个区域的权重， P_{ij} 表示选中车辆 v_i 到达ROI的次数， $v_i \in \Omega$ 。CP是选定的车辆集合 Ω 到达ROI的数量与感知区域总数 K 的比例，表示对ROI的覆盖范围，不同的任务可能需要不同类型的覆盖。如有些任务需要ROI的全覆盖，这时CP作为衡量标准比较适合；而CD衡量的是ROI热点区域的数据的可信度，当CD达到设定值时，认为数据可信。

应该注意的是，当多于一辆车经过同一区域时， P_{ij} 计数一次。以图2中的 v_1 、 v_2 为例，它们在 t_2 时刻到达F，此时，仍然认为F被覆盖一次，因为本文关心的是总覆盖情况，那么重复覆盖只计数一次，在下面的虚线框中表示。

$$CP(v_1, v_2) = (1 \times 2 + 1 \times 1 + 1 \times 2 + 1 \times 1 + 1 \times 1 + 1 \times 1 + 1 \times 1 + 1 \times 2) / 12 = 10 / 12 = 83.3\% \quad (10)$$

F区域：忽略不计

另外，车辆的感知成本记为 C ， $C = \{c_1, c_2, c_3, \dots, c_m\}$ ， Ω 表示选中车辆的集合， $C(\Omega)$ 表示所选车辆的总成本，所以 $C(\Omega) = [C, \varphi]$ 。本文的 $C_m \sim U(1, 5)$ 服从均匀分布。

2.2.5 覆盖质量的效用函数

本文衡量的覆盖质量主要包括CP和CD两个方面。覆盖质量的效用函数 $F(\Omega)$ 的优化目标是在感知成本预算 B 有限的约束条件下，实现MCS框架的高覆盖质量，计算式为

$$\max F(\Omega) = \alpha \times CP + \beta \times \lg(CD) = \alpha \times \frac{\sum_{v_i \in \Omega} P_{ij} \times W_k}{K} + \beta \times \lg \left\{ \frac{\sum_{i=1}^{t=n} \sum_{v_i \in \Omega} p'_{ij}}{K_1} \right\} \quad (11)$$

s.t. $C(\Omega) \leq B$

其中， α 和 β 分别是CP、CD的权重系数。CP代表覆盖范围的大小，即有多少ROI可以实现覆盖，单位是百分比。CD表示覆盖密度，CD越大，得到的感知数据越可信。为了设定一个合理的 α 、 β 值，本文利用文献[4]中Leave-one-out方法调整 α 和 β ，在每次迭代中，只有一辆车不参与迭代，并从给定的 α 、 β 参数范围中选择可以最大化 $F(\Omega)$ 的 α 和 β 值。

文献[4]对均匀覆盖和最大范围覆盖问题进行了研究,量化了覆盖质量,覆盖质量 $\Gamma = E_{cp} + \delta E_{ud}$ 由 E_{cp} (CP) 和 E_{ud} (覆盖均匀度) 表示,覆盖均匀度表示的是节点空间的分布情况。当前大多数相关工作集中于如何在预算内获得 ROI 的最大 CP,却较少关注热点区域的覆盖质量(可信度)问题,本文覆盖质量的效用函数 $F(\Omega)$ 的引入是为了在获得较大的 CP 的同时,使热点区域的覆盖质量也能够得到保障。

3 覆盖质量优化算法

本文提出的问题是在有限的成本范围内选择合适的车辆去达到较优的覆盖质量,这是 NP 难问题。为了证明该问题是 NP 难问题,首先证明它属于 NP 问题。假设有可能的解决方案 Ω' , Ω' 可以在多项式时间内验证,并且检查算法的复杂度为 $O(n)$,因此该问题为 NP 问题。为了进一步证明该问题是 NP 难问题,对于给定集合 $S = \{S_1, S_2, \dots, S_n\}$, 每一个 S_i 都有对应的感知收益 F_i 和感知成本 C_i ,目标是找到 $S' \subseteq S$, 满足相应的预算成本限制并获得较大的感知收益。此时问题也就转换为文献[24]中预算内最大覆盖范围问题 (MCP, maximum coverage problem), 而 MCP 已被证明为 NP 难问题,因而该问题为 NP 难问题。文献[24-25]详细地论证了该类问题是 NP 难问题。由于很难在多项式时间内找到一个 NP 完全问题的精确解,因此设计了近似算法。贪婪算法通常被用来解决该类问题,因此本文采用了贪婪算法,并对贪婪算法的指标进行了改进和优化。

本文使用了 3 种算法来优化效用函数 $F(\Omega)$, 它们是随机选择算法、贪婪算法^[25]和本文提出的改进贪婪算法 (CQO 算法)。随机选择算法是从候选车辆中随机选取参与车辆进行计算仿真;贪婪算法的选择策略是选择尽可能多的车辆,从感知成本最低的车辆开始,直到达到成本预算的上限,将所选中的车辆添加到 Ω 集合中。

定义边际效用函数 E_i 为

$$E_i = \frac{\Delta F(v_i)}{c_i} = \frac{F(\Omega') - F(\Omega)}{c_i} \quad (12)$$

其中, E_i 是车辆在单位感知成本条件下获得的边际效用函数 $F(v_i)$ 的增量, $F(\Omega')$ 是 Ω' 的效用函数, $\Omega' = \{\Omega \cup v_i\}$, $v_i \in V - \Omega$ 。CQO 算法在成本预算限制下选择 $F(\Omega)$ 与感知成本比例最高的车辆参与感知任务,所以选择车辆的标准是从 Ω 中选择 E_i

最大的车辆。 $F(\Omega)$ 是被选定车辆集 Ω 的效用函数。 $\Delta F(v_i)$ 是由 v_i 得到的效用增量,它由 $F(\Omega)$ 和 $F(\Omega')$ 计算得出。本文提出的 CQO 算法优先选择拥有最高边际效用函数 E_i 的车辆来优化覆盖质量。

CQO 算法的主要思想是将所有 m 辆候选车辆分成两组,一组是选定的集合,另一组是未被选择的集合。在每一轮选择中,从未被选择的集合中选出一辆新车并将其添加到选定的集合中;在每次选择迭代后计算效用函数 $F(\Omega)$,每次总是选择 E_i 最大的车辆,不断重复此过程直到成本预算 B 用完。根据选择策略,必须在每次迭代时遍历未选择集合中的每一辆车。在最坏的情况下,一共有 m 次选择迭代,因此计算复杂度为 $O(m^2n)$ 。CQO 算法流程如算法 1 所示。

算法 1 CQO 算法

Input: $V = \{v_1, v_2, v_3, \dots, v_m\}$, $C = \{c_1, c_2, c_3, \dots, c_m\}$, B

Output: Ω

$\Omega = \emptyset$; $S = \emptyset$; $C_{temp} = 0$;

repeat

for $S' \subseteq V$, $C(S') \leq B$ do

$S = S'$; $\max = 0$;

for $v_i \in V - S$ do

$S' = \{S' \cup v_i\}$, $\Delta F(v_i) = F(S') - F(S)$;

$E_i = \frac{\Delta F(v_i)}{c_i}$

if $E_i > \max$ and $C(S') \leq B$

$S = S'$

$\max = E_i$

end if

if $F(S) > F(\Omega)$

$\Omega = S$

end if

end for

end for

4 实验验证

在模拟实验中,本文采用了基于 Intel Core i7 3470 CPU 的服务器,内存大小为 12 GB。仿真工具为 Python 3.7 和 MATLAB R2017b。

实验是在真实数据的基础上进行的,由于井下矿工的移动轨迹数据获取难度较大,因此对于本文提出的模型采用了文献[26]中的车辆轨迹数据集进行模拟,该数据集为一天内约 4 300 辆车的 GPS 坐

标、时间、行驶速度等数据，如[id:105, time: 2007-02-20 00:00:48, longitude:121.466600, latitude: 31.220800]。

在将数据集导入仿真之前，首先对车辆的轨迹进行预处理，舍弃了采样频率过低的车辆。值得指出的是，文献[24-25]中启发式贪婪算法在每次循环中选择最大化 $\frac{w'_i}{c_i}$ 的对象，这与本文中车辆选择标准 $E_i = \frac{\Delta F(v_i)}{c_i}$ 是相同的。假设有两辆车，即 $F(v_1)=1$ 的 v_1 和 $F(v_2)=h$ 的 v_2 。设 $c_1=1, c_2=h+1 (h>0)$ ， $B = h+1$ 。 $E_1 = \frac{1}{1} > E_2 = \frac{h}{1+h}$ 。选择 v_1 是满足预算限制条件且 E_i 最大化的合理选择。 $c_1 + c_2 = h + 2 > h + 1$ 。而最优解选择 v_2 时，其效用函数 $F(\Omega)$ 的增益大于选择 v_1 ，也满足要求。文献[24-25]更详细地说明了这个问题。所以 CQO 算法在参与者有限的情况下，所获得的结果并非总是最优的。

4.1 CD 比较

在本节讨论 CD。将算法 1 的第 8 行、第 9 行的 $\Delta F(v_i)$ 用 CD 替代，也就是选择车辆的指标是 $E_i = \frac{\Delta CD(v_i)}{c_i}$ 。

车辆轨迹的示意图如图 3 所示，图 3 展示了 100 辆车在采样次数设为 24 次时的运动轨迹。车辆的 CD 分布如图 4 所示，图 4 为 ROI 内 100 辆车一天内采样 24 次轨迹的 CD 分布图，图例表示该区域被覆盖的次数。可以看到，实际的车辆覆盖并不是均匀地分布，而是以热点区域为中心分散分布，覆盖的不平衡性表明热点区域的覆盖质量可能优于其他区域，而不只是随机或均匀分布。离热点区域越远，该区域的覆盖次数越少。实际上，CD 与车辆的采

样次数有关。较大的采样次数会使该区域更容易被覆盖，因此 CD 越大，成本越高。

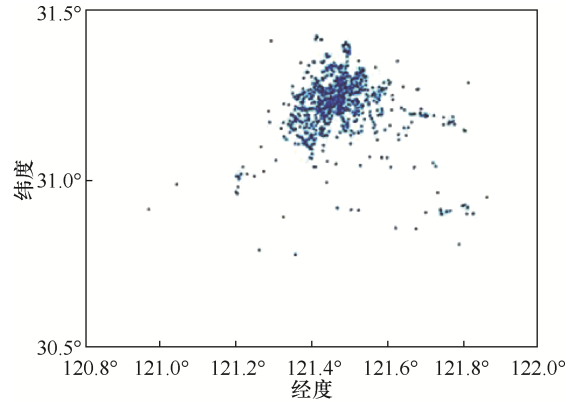


图 3 车辆轨迹的示意图

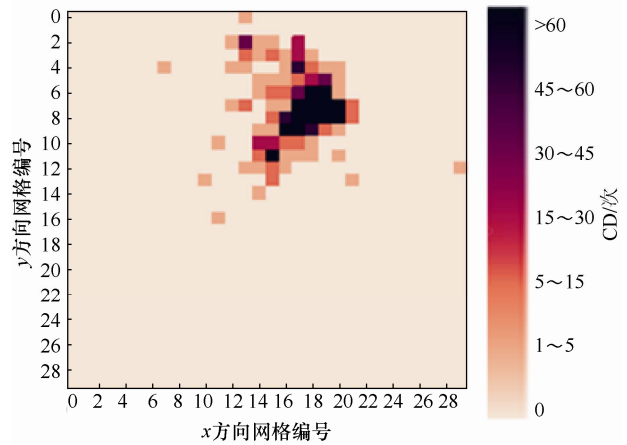


图 4 车辆的 CD 分布

经过算法仿真，CD 在不同预算成本、候选车辆数、感知采样次数下的对比如图 5 所示。

图 5(a)展示了在不同预算成本下 3 种车辆选择算法的 CD 性能。横轴表示预算成本，纵轴表示 CD。可以看出，CD 随着感知预算成本的增加而增加，预算成本越高意味着更多的车辆会被选中。

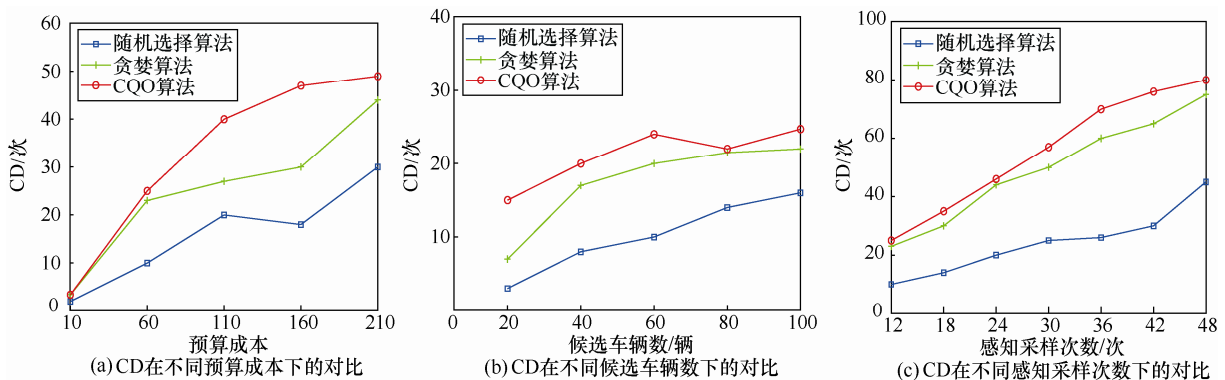


图 5 CD 在不同预算成本、候选车辆数、感知采样次数下的对比

图 5(b)是在固定预算成本下 3 种算法的候选车辆数对 CD 的影响。随着候选车辆数的增加, CD 也呈现逐渐增加的趋势。当候选车辆数较少时, 覆盖 ROI 的车辆数量有限。因此, 候选车辆越多, 能够覆盖 ROI 的次数也就越多。然而, 由于感知预算成本有限, 当候选车辆数达到 80 辆时, CD 增加趋势不明显, 这是因为之前所选的车辆几乎耗尽了感知成本。

图 5(c)展示了 3 种算法在不同感知采样次数下的 CD 性能。横轴表示一天之内的车辆感知采样次数。T=12 表示在一天内每 2 h 采样一次车辆轨迹, T=48 表示每 30 min 采样一次。实验结果趋势可以近似视作直线, 这一点很容易理解。当车辆感知次数增加时, 收集的区域覆盖数据会增多, 那么车辆覆盖 ROI 的次数也会增加, 采样频率越高, 采集和处理传感数据所需的能量就越多, 如何在减小能源消耗和得到更高的 CD 之间获得平衡是一个亟待解决的问题。

4.2 CP 比较

CP 是 STC 的另一种形式, CP 将 STC 百分比化。CP 与 CD 的不同之处在于, 本文赋予热点区域比非热点区域更大的权重值 W 。经过前期研究发现, CP 的实验结果随着成本预算、候选车辆数量和感知次数变化, CQO 算法的性能也优于其他算法的性能, 这与文献[6]中 STC 以及文献[4]中平均 CP 的性能相似, 故 CP 的比较不在本文中赘述。

4.3 效用函数比较

在本节中, 将讨论包括 CD 和 CP 在内的效用函数 $F(\Omega)$ 。为了在保证 CP 的同时最大化 CD, 所以 α 设置为 1。为了确定 β 的值, 参考了文献[4]中的 Leave-one-out 方法。在经过计算后, 当 $\beta=0.3$ 时,

CQO 算法可以得到最大 $\frac{F_{CQO} - F_{benchmarks}}{F_{benchmarks}}$ 增益。因此, 在 $F(\Omega) = \alpha \times CP + \beta \times \lg(CD)$ 中, 设 $\alpha=1, \beta=0.3$, 即 $F(\Omega) = CP + 0.3 \lg(CD)$ 。将贪心算法作为对比算法, 所以此式等于 $\frac{F_{CQO} - F_{greedy}}{F_{greedy}}$ 。

$F(\Omega)$ 在不同预算成本、候选车辆数、感知采样次数下的对比如图 6 所示。从图 6(a)可以看到, 当预算成本低于 60 时, 贪婪算法和 CQO 算法的性能相近, 原因在于当成本较低时, 只有更少的车辆可以被选中。然而, 随着预算成本的增加, CQO 算法的性能优于其他算法的性能。从图 6(b)可以看出, 在候选车辆数量从 40 辆增加到 60 辆期间, $F(\Omega)$ 迅速增加, 增速明显大于车辆数目从 60 辆增加到 80 辆和从 80 辆增加到 100 辆期间的增速。通过分析车辆编号为 40~60 的候选车辆轨迹, 发现这些车辆到达热点区域的频率比其他车辆高得多, 所以这些车辆的 CD 要比其他车辆的 CD 大, $F(\Omega)$ 也是同样的道理。由此可见, ROI 的热点区域设置和候选车辆的行驶轨迹对实验结果有很大影响。图 6(c)显示了不同算法在不同感知采样次数下的性能。总的来说, 与图 5(c)的仿真结果有些相似, 都表现出线性增长的趋势。CQO 算法同其他算法相比, 性能依旧较好。感知采样次数的增加将不可避免地导致更高的成本, 如通信成本和汽油成本。覆盖质量的提高, 如感知数据数量和精度的提高, 会消耗更多的能量。

在模拟实验过程中, 发现了一些其他问题。本文所分析的车辆数据集中, 有些车辆只是停滞在某个区域, 并没有移动。当这个区域正好是热点区域时, 那么 CD 是很大的; 反之, CD 就会很小。

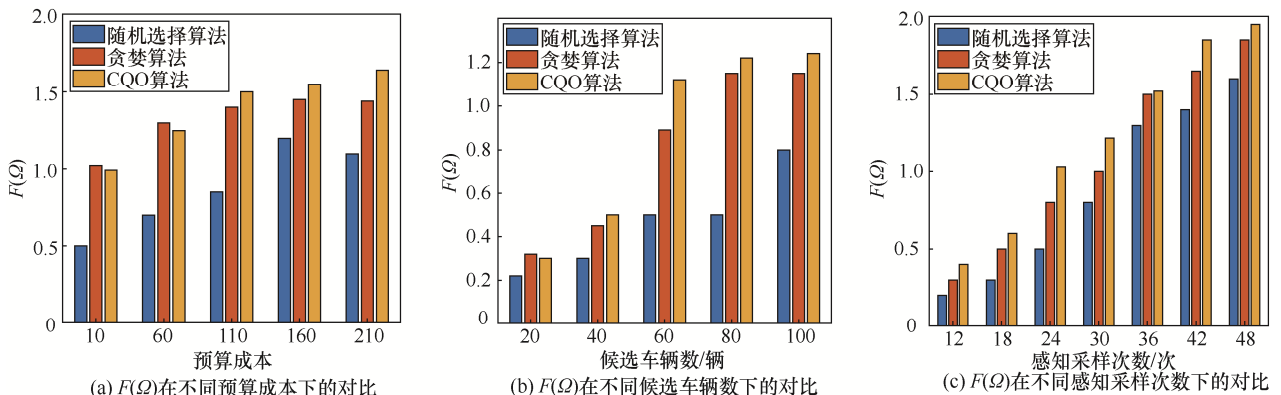


图 6 $F(\Omega)$ 在不同预算成本、候选车辆数、感知采样次数下的对比

通过设置适当的 β 值可以缓解该问题对 $F(\Omega)$ 的影响。

5 结束语

本文旨在分析在智慧矿山中数据采集的重要意义, 基于矿山中车辆移动的特性, 针对 MCS 模型, 提出了一个综合考虑 CP 和 CD 的效用函数 $F(\Omega)$ 来衡量覆盖质量。本文要解决的问题是在有限预算成本下获得较好的覆盖质量, 该问题已被证明是 NP 难问题, 因而提出改进后的贪婪算法——CQO 算法。仿真验证表明, 本文提出的 CQO 算法在仅考虑 CD 的情况下和在联合考虑 CD 和 CP 的情况下, 与贪婪算法和随机选择算法两种算法相比, 都达到了更好的覆盖质量。CD 和效用函数 $F(\Omega)$ 在预算成本、候选车辆数和感知采样次数参数改变时都能够达到较好的稳定性, 进一步研究了 CQO 算法的合理性。本文对效用函数 $F(\Omega)$ 的研究为智慧矿山 IoT 面向人一机一环海量数据收集的性能优化提供了理论参考。

参考文献:

- [1] 袁亮, 俞啸, 丁恩杰, 等. 矿山物联网人一机一环状态感知关键技术研究[J]. 通信学报, 2020, 41(2): 1-12.
YUAN L, YU X, DING E J, et al. Research on key technologies of human-machine-environment states perception in mine Internet of things[J]. Journal on Communications, 2020, 41(2): 1-12.
- [2] 吴喜雄. 通信技术在矿山无线通信的应用研究[J]. 中国战略新兴产业, 2019(24): 109.
WU X X. Application research of communication technology in mine wireless communication[J]. China's Strategic Emerging Industries, 2019(24): 109.
- [3] 牛志升, Sherman SHEN, 张钦宇, 等. 面向沉浸式体验的空天地一体化车联网体系架构与关键技术[J]. 物联网学报, 2017, 1(2): 17-27.
NIU Z S, SHEN S, ZHANG Q Y, et al. Space-air-ground integrated vehicular network for immersive driving experience[J]. Chinese Journal on Internet of Things, 2017, 1(2): 17-27
- [4] CHEN Y Y, LYU P, GUO D K, et al. Trajectory segment selection with limited budget in mobile crowd sensing[J]. Pervasive and Mobile Computing, 2017, 40: 123-138.
- [5] HOWE J. The rise of crowdsourcing[EB/OL]. WIRED, 2006.
- [6] WANG C W, LI C S, QIN C, et al. Maximizing spatial-temporal coverage in mobile crowd-sensing based on public transports with predictable trajectory[J]. International Journal of Distributed Sensor Networks, 2018, 14(8): 155014771879535.
- [7] WANG C W, GAIMU X G, LI C S, et al. Smart mobile crowdsensing with urban vehicles: a deep reinforcement learning perspective[J]. IEEE Access, 2019, 7: 37334-37341.
- [8] HE Z J, CAO J N, LIU X F. High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility[C]//2015 IEEE Conference on Computer Communications (INFOCOM). IEEE, 2015: 2542-2550.
- [9] GONG W, ZHANG B X, LI C. Location-based online task assignment and path planning for mobile crowdsensing[J]. IEEE Transactions on Vehicular Technology, 2019, 68(2): 1772-1783.
- [10] MASUTANI O. A sensing coverage analysis of a route control method for vehicular crowd sensing[C]//2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops). IEEE, 2015: 396-401.
- [11] XU J L, WANG S G, ZHANG N, et al. Reward or penalty: aligning incentives of stakeholders in crowdsourcing[J]. IEEE Transactions on Mobile Computing, 2019, 18(4): 974-985.
- [12] HU A D, GU Y G. Mobile crowdsensing tasks allocation for multi-parameter bids[C]//2017 IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC). IEEE, 2017: 489-493.
- [13] LUONG N C, HOANG D T, WANG P, et al. Data collection and wireless communication in Internet of things (IoT) using economic analysis and pricing models: a survey[J]. IEEE Communications Surveys & Tutorials, 2016, 18(4): 2546-2590.
- [14] SURYADASARI V, POURYAZDAN M, KANTARCI B. On the impact of selective data acquisition in mobile crowd-sensing performance[C]//2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE). IEEE, 2018: 1-4.
- [15] TRAN-THANH L, STEIN S, ROGER S A, et al. Efficient crowdsourcing of unknown experts using bounded multi-armed bandits[J]. Artificial Intelligence, 2014, 214: 89-111.
- [16] AL-TURJMAN F, KARAKOC M, GUNAY M, et al. Routing mobile data couriers in smart-cities[C]//2016 IEEE International Conference on Communications (ICC). IEEE, 2016: 1-6.
- [17] REN Y Y, LIU Y X, ZHANG N, et al. Minimum-cost mobile crowdsourcing with QoS guarantee using matrix completion technique[J]. Pervasive and Mobile Computing, 2018, 49: 23-44.
- [18] XIONG H Y, ZHANG D Q, CHEN G L, et al. iCrowd: near-optimal task allocation for piggyback crowdsensing[J]. IEEE Transactions on Mobile Computing, 2016, 15(8): 2010-2022.
- [19] SIM I, CHOI K, KWON K, et al. Energy efficient cluster header selection algorithm in WSN[C]//2009 International Conference on Complex, Intelligent and Software Intensive Systems. IEEE, 2009: 584-587.
- [20] LEE H J, WICKE M, KUSY B, et al. Data stashing: energy-efficient information delivery to mobile sinks through trajectory prediction[C]//The 9th ACM/IEEE International Conference on Information Processing in Sensor Networks. ACM, 2010: 291-302.
- [21] ZHAO D, MA H D, LIU L. Energy-efficient opportunistic coverage for people-centric urban sensing[J]. Wireless Networks, 2014, 20(6): 1461-1476.
- [22] 唐志博. 基于矿工行为分析的 WMSN 网络优化方法[D]. 徐州: 中国矿业大学, 2015.

TANG Z B. WMSN network optimization method based on miners' behavior analysis[D]. Xuzhou: China University of Mining and Technology, 2015.

[23] 马丰原. 基于视觉感兴趣区域的图像质量评价算法研究[D]. 西安: 西安科技大学, 2013.

MA F Y. Research on image quality assessment based on visual regions of interest[D]. Xi'an: University of Science and Technology, 2013.

[24] KHULLER S, MOSS A, NAOR J S. The budgeted maximum coverage problem[J]. Information Processing Letters, 1999, 70(1): 39-45.

[25] SVIRIDENKO M. A note on maximizing a submodular set function subject to a knapsack constraint[J]. Operations Research Letters, 2004, 32(1): 41-43.

[26] NI L M, CHEN L, QU H, et al. Smart city GPS data set[DB/OL]. 2007.

[作者简介]



王朝炜(1982-), 男, 陕西西安人, 博士, 北京邮电大学电子工程学院副教授, 主要研究方向为下一代移动通信技术、无线传感器与IoT技术等。



刘婷(1995-), 女, 江苏淮安人, 北京邮电大学电子工程学院硕士生, 主要研究方向为MCS、边缘计算等。



王天宇(1996-), 男, 河南郑州人, 北京邮电大学电子工程学院硕士生, 主要研究方向为MCS、移动节点能效优化技术等。



王卫东(1967-), 男, 内蒙古包头人, 博士, 北京邮电大学电子工程学院教授, 主要研究方向为卫星移动通信、下一代移动通信技术、IoT技术等。