Noname manuscript No.
(will be inserted by the editor)

Recruitment Algorithms for Vehicular Sensor Networks

**Fabio Campioni · Salimur Choudhury ·
Usman Tariq · Ali Kashif Bashir**

Received: date / Accepted: date

Abstract Vehicular crowdsensing allows the rapid, predictable movement of vehicles, as well as their wide variety of sensors, to gather sensing data in crowdsensing applications. Recruitment algorithms are used to select a subset of participants in an area that will provide the most complete coverage. In this paper, we explore two variations of the vehicular recruitment problem. In the first problem, which we refer to as the priority based vehicle recruitment problem, we consider coverage areas in which subsets must be covered. In the multisensor variation, we consider coverage areas which require different types of sensors, in which participating vehicles have one or more sensor types onboard. For each, we implement a mixed integer programming model which returns optimal solutions, as well as a heuristic for obtaining approximate solutions. In the unbudgeted priority vehicular recruitment performance evaluation, our heuristic on average obtains only 0.05% lower utility at 1.78% higher recruitment cost. In the budgeted runs, our heuristic obtains on average only 0.02% lower utility at 0.59% higher recruitment costs. In the unbudgeted multisensor vehicular recruitment performance evaluation, our heuristic obtains only 0.04% lower utility at 1.10% higher recruitment cost, and in the

F. Campioni
School of Computing, Queen's University, Kingston, ON, Canada
E-mail: campioni@cs.queensu.ca

S. Choudhury
Department of Computer Science, Lakehead University, Thunder Bay, ON, Canada
E-mail: salimur.choudhury@lakeheadu.ca U. Tariq
College of Computer Science and Engineering, Prince Sattam bin Abdulaziz University, Saudi Arabia
E-mail: u.tariq@psau.edu.sa

A. Bashir
Center for Advanced Computation Science, Manchester Metropolitan University, Machester, UK
E-mail: dr.alikashif.b@ieee.org

budgeted runs we obtain 11.33% lower utility at 0.27% higher recruitment cost.

1 Introduction

Compared to existing wireless sensor networks, participatory sensing utilizes the presence of ordinary users in order to capture sensing data. Using wireless networks, this data is shared among users, or sent to a central service which utilizes the data in an application [6]. The use of incentives encourages users to participate in crowdsensing networks. Participatory sensing does not assume all users will participate and has to receive confirmation from a participant that it will contribute data. Opportunistic sensing, however, assumes all users automatically participate, without any explicit authorization. [2]. Existing works have almost exclusively considered mobile devices.

Devarakonda et. al [4] place sensing units in vehicles and use them to measure air quality throughout an area of interest. Rana et. al [11] utilize the microphones available on smartphones to measure noise pollution through an area. Hull et. al [9] propose a traffic monitoring system where smartphones located in vehicles can measure congestion, delays, and conditions using their sensors.

However, mobile devices are not always ideal for crowdsensing applications due to their limited resources [2]. The sensors and network modems can especially consume processing power and battery life, as well as cost associated with data transmission. Modern vehicles, however, are being built with several onboard processors, storage space, and larger batteries (especially electric vehicles). Their rapid and predictable movement allows few participants to cover a large area. There are over 70 sensors installed in new vehicles as of 2013 [5], which makes them ideal for crowdsensing.

Several existing works all describe a similar framework for vehicular crowdsensing. A road side service provider receives sensing tasks from applications. Having knowledge of vehicle positioning throughout an area, it recruits vehicles that can obtain the sensing data, by providing incentives to the vehicles, obtaining and aggregating the data, and then using it within an application.

In this paper we present several contributions:

- Propose two new variations of the vehicular recruitment problem referred to as the priority based vehicular recruitment problem where coverage areas are assigned different priorities, and the multisensor vehicular recruitment problem where sensing tasks require multiple types of sensors
- Develop a MIP (mixed integer programming) model for providing solutions to the priority based vehicular recruitment problem, as well as a heuristic for obtaining approximate solutions
- Propose a MIP model for the multisensor vehicular recruitment problem, as well as a heuristic for the multi sensor vehicular recruitment problem

The organization of the paper is as follows: in section 2 we describe some most relevant prior work. Our priority based vehicle recruitment problem is introduced in section 3 along with our proposed algorithm and related experimental results. We introduce multi sensor vehicle recruitment problem in chapter 4. It also includes our proposed algorithm and related experimental results.

2 Related Work

We review several existing recruitment algorithms for crowdsensing applications. Existing crowdsensing works utilized mobile devices extensively. Reddy et. al [12] describe several approaches for selecting mobile participants. They utilize several factors including past participation and availability to determine how reliable a participant's data is. The system requires that participants' trajectories are known or predetermined at the time of recruitment, and that its reputation score meets the requirement. Coverage is defined as the number of equally sized geographic areas covered by at least one selected participant. They compare three approaches; a random approach, where participants are selected at random until the budget is exhausted; the naive approach, which continuously selects the participant providing the most coverage; and the heuristic approach which is identical to the naive approach but updates the unique coverage of other participants after a participant is selected. Each of the techniques describe select participants until the total incentive budget has been exhausted. The simulation results show that the greedy approach outperforms the other two approaches.

Abdelhamid et. al [7] propose a framework for selecting an optimal subset of vehicular participants in crowdsensing networks. They describe three different objectives; one objective to maximize coverage, another to minimize cost, and finally one which minimizes coverage overlap. The maximum coverage model aims to achieve the highest coverage possible with no regard to cost or redundancy. The minimum cost model minimizes the total recruitment cost while still meeting the optimal level of coverage found in the previous model. The minimum overlap model minimizes redundant coverage while again still maintaining optimal coverage obtained by the first maximum coverage model. The authors perform simulations on randomly generated areas of interest and divide the area into equal sized sections. Each vehicle is assigned a reputation value at random between 0 and 1. The particular problem is shown to be NP-hard in other related. Due to the speed at which vehicles move, a solution has a very short shelf life before it becomes too inaccurate. Because of this, the authors mention the need for heuristics to obtain approximate solutions in polynomial time.

Abdelhamid et. al [1] also propose several heuristic algorithms for the vehicular recruitment problem. The paper builds upon their work in [7]. The algorithms attempt to obtain as much coverage as possible in polynomial time, as the optimal solution requires exponential time in the worst case. They first remove any trajectories with a reputation less than the threshold. The algorithm repeatedly selects the vehicle covering the most uncovered area and subtracting its incentive cost from the budget. Vehicles providing no coverage are removed from consideration. The algorithm runs until all area is covered, there are no more vehicles to consider, or the budget has been met. The RBMC algorithm makes no attempt to minimize incentive cost, so the authors propose a second algorithm, RBMC-MC, with the additional goal of minimizing cost. The algorithm works the same as RBMC, but when selecting the vehicle covering the most area, uses cost as a tie breaker. In their performance evaluation, they compare the performance of their heuristic algorithms to solutions obtained by the optimal framework in [7]. They run various scenarios with different number of vehicles, as well as with a variety of reputation thresholds and budgets. Their results show that they are able to achieve nearly the same coverage as optimal with equal or slightly higher recruitment costs overall. However, Campioni et al [3] show that in the worst case their algorithm is unbounded and they propose a new algorithm to solve the problem.

Yi et. al [15] propose the Fast-VPR algorithm. They consider a two dimensional grid of equal sized cells over a period of time rather than a single dimension. They define coverage as the union of grid cells covered across all time periods. Instead of using a budget constraint, they introduce a parameter λ which is used to restrict the total recruitment cost of a solution. The algorithm aims to maximize coverage by recruiting a a subset of vehicles with the lowest possible recruitment cost. The algorithm starts from a set containing all vehicles and an empty set and compares the change in utility from removing the vehicle from the initial full set and adding it to the initial empty set and probabilistically adding it to the solution. The performance evaluations utilize both synthetic, randomly generated vehicle trajectories, as well as real world trajectories from a traffic dataset. The authors also evaluate performance using both randomized pricing models as well as a pricing model where the incentive cost of a participant is related to the number of grid cells it covers. They also compare their algorithms using randomly generated trajectories, as well as real world trajectory datasets. They compare their algorithm to the Greedy-SC algorithm proposed by He et. al [8]. Both algorithms are compared in terms of spatiotemporal coverage. Their results show that while they achieve slightly less coverage than Greedy-SC, FastVPR returns solutions in less time.

For the multisensor vehicular recruitment problem in section 4, we draw inspiration from Liu et al. [10]. They propose both an optimal participant recruitment mechanism as well as a greedy participant recruitment algorithm for the heterogeneous vehicular recruitment problem. They consider a sensing area where different grid cells require different types of sensors, such as vibration sensors, brightness sensors, and temperature sensors. A number of vehicles travel throughout the sensing area, each with a subset of the sensor types onboard, as well as a predefined trajectory. They propose a multi objective optimization model to minimize the difference between required and collected data matrices for all sensing tasks. However, they note that this problem is difficult to solve, so they break the problem down by sensing tasks and assign a weight to each task. They then propose the HPR (heterogeneous participant recruitment) strategy. Essentially, the algorithm repeatedly selects the participant with the highest utility, and pays it an incentive. This repeats until all incentives are depleted or all participants have been evaluated. The algorithm loops through all time periods. In their performance evaluation, they test the algorithms using the T-drive trajectory dataset [16], which contains weekly trajectories of over $10,000$ taxis. Using a variety of budgets, the authors found that the proposed greedy strategy could collect 85.4 % of sensing data with 34 % incentive budget.

Xiao et. al [14] propose the DUR (Deadline-sensitive User Recruitment) problem for mobile crowdsensing networks. The problem is similiar to most other crowdsensing user recruitment algorithms, although they consider probabilistic coverage, as participants do not always follow their announced trajectory. Although they do not specifically consider vehicles, the concepts can be adapted to vehicular crowdsensing networks. They prove that the DUR problem can be formulated as a non-trivial set cover problem. They also propose a greedy algorithm, called gDUR, to obtain approximate solutions to the problem. They also consider a variation of the problem which considers sensing duration deadlines, called dDUR. Since their problem formulation differs slightly from existing works, they summarize the various techniques and develop two more algorithms for comparison to their algorithms, called MCUR (Minimum Cost User Recruitment) and MCURP (MCUR with Probabilistic Mobility). They compare the algorithms on both real world and synthetic mobility trace datasets. They use a variety of time durations and deadlines. In all cases, their gDUR and dDUR algorithms have vastly improved recruitment cost (67 to 96%) compared to MCURP and 166 to 227% larger successful processing ratios compared to MCUR. However, they found that MCUR achieved lower costs compared to their algorithms, because it did not consider the deadline constraint, so would select fewer users, which resulted in lower successful processing ratios. Also, in some cases, MCURP achieved larger successful processing ratios, because the algorithm usually selected many more users, but this resulted in much larger recruitment costs. In the synthetic traces, they found that gDUR and dDUR had between 59 to 67.6% fewer recruitment costs and 12.2 to 17 times larger successful processing ratios compared to MCUR.

Xiao et. al [13] explore a game theory strategy for the vehicular recruitment problem. The authors formulate the problem as a game. Each vehicle selects a different sensing strategy based on the sensing cost, its radio channel state, and the expected payment from the mobile crowdsensing server. Vehicles are paid by the server based on the accuracy of their sensing data. They implement learning strategies so that the vehicles and server can estimate system paremeters over time without explicitly knowing the system model. They compare various learning strategies compared to a greedy strategy and compare them in random generated simulations. They also compare all strategies to a random selection method. Their results show that a PDS learning technique improves learning speed, greatly increases the utility of cars, and reduces energy usage compared to the random and Q-learning methods.

3 Priority Based Vehicle Recruitment Problem

In this section we propose a new variation of the vehicle recruitment problem. We start from the same formulation as mentioned in section 3. The main objective function is still to maximize utility. However, we consider slightly different problem scenarios. In addition, we develop a heuristic to obtain approximate solutions and compare its performance to the MIP model. In this formulation, we consider maximizing coverage of an area where some important areas must be covered. That means, there are some areas (priority areas) that has to be covered. In addition, we need to cover other areas as much as we can. In all other existing variations of this problem, each section of coverage area is equal in coverage priority. However, in this problem instance, we design a framework that ensures high priority areas are covered, by adding a constraint to the formulation such that certain grid cells must be covered at specified time periods. A grid cell refers to a single section of area requiring sensing coverage. The size of the grid cell would depend on the sensing application and would be decided by those deploying the crowdsensing system.

3.1 Problem Definition

In the real world, not all sensing areas are equally important. For example, an area might not normally be sensed except in the case of a situation or emergency. Or, an area may not need to be sensed except for certain times of day, and then in off peak hours priority can be given to other areas.

It is critical for efficient vehicular crowdsensing that service providers be able to efficiently recruit vehicles based on coverage areas with various levels of priority. Doing so ensures cheaper and more useful sensing data.

We now describe the problem formulation. V refers to the set of available vehicles, and S refers to the selected subset of V that are chosen to provide coverage. U refers to the utility of a selected subset S, which is defined in equation 3.

The objective function, which is maximized, aims to maximize utility:

$$\max U(S) \tag{1}$$

such that:

$$S \subseteq V \tag{2}$$

The utility is calculated as follows:

$$U(S) = f(S) - cost(S) \tag{3}$$

f refers to the coverage obtained by a solution, and *cost* refers to the recruitment cost associated with that selection of vehicles.

R_{vt} indicates which grid cells a vehicle v covers at time period t.

We calculate coverage for each grid cell g at each time period t in equations 4 and 5. c_{gt} is equal to the number of vehicles in the current solution covering a grid cell g at time period t, and c'_{gt} is equal to 1 if c_{gt} is greater than 0 and 0 otherwise.

$$c_{gt} = \{v \in V : v \in S, g \in R_{vt}\} \tag{4}$$

$$c'_{gt} \leq |c_{gt}|; \quad \forall g, t \in G, T \tag{5}$$

$$f(s) = \sum_{t,g \in T,G} c'_{gt} \tag{6}$$

We also add a constraint that ensures certain grid cells are covered during certain time periods:

$$c'_{gt} \geq 1; \quad \forall (g,t) \in G_{req} \tag{7}$$

G_{req} contains ordered pairs (g,t) if a grid cell g must be covered at time period t.

3.2 Heuristic Priority-Based Vehicle Recruitment

In this section we describe a heuristic algorithm to solve the priority-based vehicle recruitment problem. Since the MIP formulation described in the above section is NP-hard, we propose a heuristic solution to produce approximate solutions in polynomial time.

The heuristic algorithm works as follows. The algorithm iterates through each time period $t \in T$. For each time period, it calculates the set of grid cells that must be covered at this time. It then calculates the grid cells covered by the current solution. It then subtracts the covered cells from the required cells to determine which cells still need to be covered. It iterates through the set of available vehicles, choosing the one adding the most utility to the current solution. Like some of our existing heuristics, it calculates the cost per unit of utility increase and chooses the cheapest one. When a vehicle is selected, the cells it covers are removed from the set of uncovered cells and added to the set of covered cells. This process repeats until all required cells have been covered, and then the algorithm repeats this process, selecting vehicles to cover as many non required grid cells as possible. The budgeted version works in the same way, but only considers vehicles that can be afforded under the current budget.

Algorithm 1 Unbudgeted Required Coverage Heuristic

Set of vehicles V
Set of grid cells G
Set of time periods T
Set $Greq_{gt}$indicating required coverage for grid cell g at time t
Set of required cells RC
Set of covered cells CC
Set of uncovered cells UC
Set R_{vt} containing grid cells covered by v at time t
Set of costs C for each vehicle v
Set of selected vehicles S
for $t \in T$ **do**
    $RC = \{g \in G : g \in Greq_{gt}\}$
    $CC = \{g \in G : g \in R_{vt} \ \forall \ v \in S\}$
    $UC = RC - CC$
    while $UC \neq \emptyset$ **do**
        $covering\_vehicles$ =vehicles not currently selected and which cover one or more uncovered required cells at this time period
        $p_v = \frac{U(S+v)-U(S)}{C_v}$ for each covering vehicle
        max_v = vehicle with highest p_v
        for each grid cell g covered by max_v **do**
            $CC = CC \cup g$
for $t \in T$ **do**
    $CC = \{g \in G : g \in R_{vt}, v \in S\}$
    $UC = G - CC$
    while $UC \neq \emptyset$ **do**
        $covering\_vehicles$ =vehicles not currently selected and which cover one or more uncovered grid cell at this time period
        $p_v = \frac{U(S+v)-U(S)}{C_v}$ for each covering vehicle
        max_v = vehicle with highest p_v
        for each grid cell g covered by max_v **do**
            $CC = CC \cup g$

3.3 Performance Evaluation

In this section we use our MIP framework as a benchmark to provide an upper bound on utility so that we can evaluate the solutions produced by our heuristic. We generate random scenarios with a variety of number of vehicles. We demonstrate algorithm performance with up to 500 vehicles as we feel this is a reasonable number of vehicles to expect at most in a particular area for a sensing application. We also randomly assign required coverage to grid cells at a time period with a 0.01 probability. Our reasoning for selecting a low probability of required coverage is that many of the randomly generated problem instances were infeasible with a higher probability. Each vehicle is assigned a random recruitment cost between 0 and 1.

In order to ensure feasible runs, we set the budget as a function of the number of vehicles, as a scenario with a large number of vehicles and a very small budget is often found to be infeasible by the MIP model.

Table 1 shows the average recruiter utility and recruitment cost across all scenarios. In the scenarios with small numbers of vehicles, the heuristic returns

	MIP RC		RC_Heuristic	
Number of Vehicles	Average of Utility	Average of Cost	Average of Utility	Average of Cost
50	927.5	23.98	927.5	23.98
100	1876.5	45.99	1876.35	46.08
150	2557.35	65.23	2557	65.41
200	3364.25	84.31	3364	84.55
250	3649.6	92.93	3648.25	94.23
300	4078.5	105.92	4076.8	107.74
350	4189.2	117.27	4186.5	119.93
400	4498.65	125.85	4496.1	128.44
450	5240.1	131.43	5236.25	135.16
500	5232.7	136.79	5228.65	140.76

Table 1 Unbudgeted priority based vehicular recruitment

solutions very close to optimal. In the 50 vehicle scenario the heuristic achieves the same utility with only 0.01% higher recruitment cost. In the 100 to 200 vehicle scenarios, the heuristic achieves 0.01% lower utility, at between 0.20 to 0.27% higher recruitment costs. In 250 and 300 vehicle scenarios, the heuristic achieves only 0.04% lower utility at 1.41 and 1.72% higher recruitment costs. With 350 and 400 vehicle scenarios, the heuristic achieves only 0.06% lower utility at 2.06 and 2.26% higher recruitment cost. With 450 vehicles, the heuristic achieves 0.07% lower utility at 2.84% higher recruitment cost. In the final scenario with 500 vehicles, the heuristic achieves just 0.08% lower utility with only 2.9% higher recruitment cost. While the performance differences increase as the number of vehicles increases, the heuristic still achieves solutions very close to optimal.

3.4 Budgeted Priority Based Vehicle Recruitment Problem

In this section we modify the problem introduced in the previous section to add a budget constraint. Sometimes a service provider may have a limited amount of incentives to provide to participants but still need to provide sensing coverage for an area. By introducing a budget the service provider can limit the amount of incentives it has to pay out to participants while still attaining the best coverage possible. We use the same MIP framework as the normal priority based vehicle recruitment problem, with the exception of the following constraint which limits the recruitment cost of a solution S to a budget constraint B:

$$\sum_{v \in S} C_v \leq B \tag{8}$$

The budgeted version works the same way, but keeps track of the current solution cost and only considers vehicles that can be added to the solution while still staying under the specified budget.

Algorithm 2 Budgeted Required Coverage Heuristic

Set of vehicles V
Set of grid cells G
Set of time periods T
Set R_v^t containing grid cells covered by v at time t
Set of costs C for each vehicle v
Budget B
Set of selected vehicles S
$current\_cost = 0$
for $t \in T$ **do**
    $RC = \{g \in G : g \in Greq_{gt}\}$
    $CC = \{g \in G : g \in R_{vt} \ \forall \ v \in S\}$
    $UC = RC - CC$
    while $UC \neq \emptyset$ **do**
        $covering\_vehicles$ = vehicles not currently selected which cover one or more uncovered required cells, and which can be afforded under the current budget
        $p_v = \frac{U(S+v)-U(S)}{C_v}$ for each covering vehicle
        max_v = vehicle with largest p_v
        $S = S \cup max\_v$
        current_cost = current_cost + C_{max_v}
        for each grid cell g covered by max_v **do**
            $CC = CC \cup g$
for $t \in T$ **do**
    $CC = \{g \in G : g \in R_{vt}, v \in S\}$
    $UC = G - CC$
    while $UC \neq \emptyset$ **do**
        $covering\_vehicles$ = vehicles not currently selected and which cover the current grid cell at this time period, which can be afforded by the current budget
        $p_v = \frac{U(S+v)-U(S)}{C_v}$ for each covering vehicle
        max_v = vehicle with highest p_v
        for each grid cell g covered by max_v **do**
            $CC = CC \cup g$

		MIP RC Budgeted		RC Heuristic Budgeted	
Budget	Number of vehicles	Average Utility	Average Cost	Average Utility	Average Cost
50	25	910.00	23.78	910.00	23.78
	50	1242.00	25.76	1242.00	25.76
	100	1046.00	24.52	1046.00	24.69
100	50	2211.00	45.82	2211.00	45.97
	100	2002.00	46.96	2001.00	47.02
	200	1868.00	46.11	1868.00	46.11
200	100	3036.00	82.90	3036.00	82.95
	200	3039.00	83.36	3039.00	83.54
	400	3209.00	77.78	3208.00	78.43
250	125	3523.00	90.96	3520.00	93.04
	250	3539.00	97.53	3538.00	98.22
	500	3396.00	101.48	3396.00	101.86

Table 2 Budgeted Priority Based Vehicular Recruitment

The performance differences are very similar for the budgeted problem. In the 50 vehicle scenarios, the heuristic is only on average 0.23% more expensive than the optimal solution while achieving the same coverage in all budgets. In the 100 vehicle scenarios, the heuristic only returns 0.02% lower utility on average with only 0.16% higher recruitment costs. In the 200 vehicle scenario, the heuristic solution is 0.36% more expensive than optimal with only 0.01% lower utility on average. In the final scenarios with 250 vehicles, the heuristic has 0.04% lower utility with 1.90% higher recruitment costs on average.

4 Multi Sensor Vehicle Recruitment Problem

In this section we propose a framework for solving multi sensor coverage. Sensors are generally designed to sense a particular phenomenon and nothing else. However, different areas contain different sensing tasks that only certain sensors can measure. Additionally, not all vehicles contain every type of sensor in order to perform these sensor tasks. Therefore, we reformulate the original framework to consider multiple sensors per vehicle, with each portion of coverage area requiring one or more types of sensors and only consider an area covered if a selected vehicle covers the area with the correct type of sensor.

We draw inspiration from Liu et al. [10] . They also consider coverage of areas which require multiple sensor types. They propose a utility scheme to determine how much utility a vehicle can add by being selected, and also propose a participant recruitment algorithm. However, they consider uncertain trajectories and use probabilistic methods of determining where vehicles move, whereas we use a fixed model where the location of vehicles is already known. They also propose an optimal recruitment algorithm to provide a benchmark to compare their greedy partipant recruitment algorithm. During performance of their participant recruitment algorithms they compare them using a real world trajectory dataset and vary both budgets and distribution of sensors among vehicles.

However, due to the fact that Liu et. al [10] consider vehicles with uncertain trajectories and we only consider vehicles with certain trajectories, we do not compare to their algorithm in our performance evaluation.

4.1 Problem Definition

The objective function still aims to maximize the utility function, U. S refers to the set of selected vehicles, which is a subset of V, which is the set of all available vehicles.

$$\max U(S) \tag{9}$$

such that:

$$S \subseteq V \tag{10}$$

The utility is calculated as follows:

$$U(S) = f(S) - cost(S) \tag{11}$$

When calculating coverage, we only consider vehicles that have that grid cell's sensor type in the calculation.

For each grid cell g at time period t, coverage c is calculated as the number of selected vehicles which cover that cell at that time period, but also have the sensors required to provide coverage for that grid cell. Then, c' is calculated as 1 if $|c| > 0$ for a particular grid cell g at time t, and 0 otherwise.

Therefore, the coverage calculation changes to:

$$c_{(g,t)} = \{v \in S : g \in R_{vt} \quad Gtype_{gt} \in Vtype_v\} \tag{12}$$

$$c'_{gt} \leq |c_{gt}|; \quad \forall g, t \in G, T \tag{13}$$

$$f(s) = \sum_{t,g \in T,G} c'_{gt} \tag{14}$$

For each vehicle $v \in V$, $Vtype_v$ contains the types of sensors vehicle v has. A vehicle can have one or more sensors onboard. Gtype contains the type of sensor required for grid cell g to be covered at time period t.

4.2 Heuristics

We also propose several heuristic algorithms for the multisensor vehicular recruitment problem. We propose both an unbudgeted and budgeted version.

The unbudgeted heuristic works as follows. The algorithm iterates over each time period $t \in T$ and grid cell $g \in G$. It first checks if this grid cell is covered by any vehicles already selected, and skips it if so. It then calculates which vehicles can cover this cell based on whether the vehicle is located at the grid cell at the specified grid cell at the time period and contains the type of sensor required by the grid cell. It calculates the cost per unit of utility gained and chooses the vehicle with the cheapest ratio and adds it to the solution.

The budgeted version works the same as the unbudgeted version, but keeps track of recruitment cost and only considers vehicles that can be added to the solution set without exceeding the specified budget.

Algorithm 3 Unbudgeted Multisensor Vehicle Recruitment Heuristic

Set of vehicles V
Set of sensor types $Sensor$
Set of grid cells G
Set of time periods T
Set R_{vt} containing grid cells covered by v at time t
Set of costs C for each vehicle v
Set of selected vehicles S
for $t \in T$ **do**
    for $g \in G$ **do**
        $coverage_{gt}$ = vehicles covering this grid cell at this time period with the required sensor type
        if $coverage_{gt} \neq \emptyset$ **then**
            Continue
        covering vehicles = vehicles not currently selected which cover grid cell g and contain the required type of sensor to cover g
        $p_v = \frac{U(S+v)-U(S)}{C_v}$ for each covering vehicle
        max_v = vehicle with largest p_v
        $S = S \cup max\_v$

Algorithm 4 Budgeted Multisensor Vehicle Recruitment Heuristic

Set of vehicles V
Set of sensor types $Sensor$
Set of grid cells G
Set of time periods T
Set R_{vt} containing grid cells covered by v at time t
Set of costs C for each vehicle v
Set of selected vehicles S
$current\_cost = 0$
Budget B
for $t \in T$ **do**
    for $g \in G$ **do**
        $coverage_{gt}$ = grid cells covered by current solution at this grid cell and time period which can be afforded under the current budget
        if $coverage_{gt} \neq \emptyset$ **then**
            Continue
        covering vehicles = vehicles covering this grid cell at this time period with the required sensor type and which can be afforded under the current budget
        $p_v = \frac{U(S+v)-U(S)}{C_v}$ for each covering vehicle
        max_v =vehicle with largest p_v value
        $S = S \cup max\_v$
        $current\_cost = current\_cost + C_{max\_v}$

	MIP Multisensor		Multisensor Heuristic	
Number of Vehicles	Average of Utility	Average of Cost	Average of Utility	Average of Cost
50	691.05	21.47	691.05	21.48
100	1513.05	45.47	1513.05	45.53
250	2904.35	106.04	2904.05	106.41
500	4441	151.52	4437.9	154.64

Table 3 Unbudgeted multi sensor vehicle recruitment

		MIP Multisensor Budgeted		Multisensor Heuristic Budgeted	
Budget	Number of vehicles	Average Utility	Average Cost	Average Utility	Average Cost
50	25	743	21.03	743	21.03
	50	679	20.87	679	20.87
	100	818	22.51	818	22.51
	200	705	23.77	705	23.80
100	25	1581	24.97	1419	25.00
	50	1591	49.13	1591	49.13
	100	1570	42.00	1570	42.00
	200	1270	42.70	1270	42.70
250	25	2347	24.99	1568	24.99
	50	2799	49.97	2212	50.00
	100	2728	99.85	2709	99.74
	200	2587	95.19	2587	95.72
500	25	3043	24.99	1865	25.00
	50	3700	49.99	2916	49.98
	100	4066	99.89	3678	100.00
	200	4197	153.93	4195	155.62

Table 4 Budgeted multisensor vehicle recruitment

4.3 Performance Evaluation

In the first set of runs (table 3) we do not impose a budget constraint on either of the algorithms. In the 50 and 100 vehicle scenarios the heuristic achieves the same utility that the MIP model does. Recruitment costs are only 0.04 and 0.13% higher, respectively. In the 250 vehicle scenario the heuristic utility is 0.01% lower with 0.35% higher recruitment cost. In the final scenario with 500 vehicles the heuristic utility was on average 0.07% lower with 2.06% higher recruitment costs.

In the budgeted runs (table 4), we run a variety of budgets for each scenario with a varying number of vehicles. In small budgets, the biggest performance difference can be seen between the heuristic and MIP model. In the runs with a budget of 25, the heuristic achieves the same utility in the 50 vehicle scenario but this decreases as the number of vehicles increases, to 10.25, 33.19, and 38.71% lower utility, respectively. Recruitment costs are similiar (between 0.01 and 0.1% higher).

In the runs with a budget set to 50 , utility is the same for 50 and 100 vehicle scenarios, and 20.97 and 21.19 percent lower in 250 and 500 vehicle scenarios. In the 100 budget runs, the performance gap continues to further decrease. The heuristic utility is optimal in 50 and 100 vehicle scenarios and only 0.07% lower in the 250 vehicle scenario. Recruitment costs are also the same as optimal in 50 and 100 vehicle scenarios. In the largest scenario with 500 vehicles, the heuristic utility is only on average 9.54% lower than optimal with only 0.10% higher recruitment cost. In the largest budget run, with a budget of 250, the heuristic performs very closely to the MIP model. In all scenarios but the largest, the heuristic achieves optimal utility at very similar recruitment costs. In the largest scenario with 500 vehicles the heuristic results in 0.05% lower utility at 1.10% higher recruitment cost.

5 Conclusion

In this paper, we have proposed both a mixed integer programming model for a new variant of the vehicular recruitment problem, which we call the priority based vehicular recruitment problem. We do the same for an existing variation, also known as the multisensor or heterogeneous vehicular recruitment problem, where vehicles contain different types of sensors for different sensing areas. For both problems, we also develop a heuristic which we compare to solutions obtained by the optimal models. Results show we obtain very similar coverage at similar recruitment costs. In the multisensor performance evaluation, our heuristic underperforms compared to the MIP model in low budgets with large numbers of vehicles, but the performance gap lessens as the budget parameter increases.

References

1. S Abdelhamid, H S Hassanein, and G Takahara. Reputation-Aware, Trajectory-Based Recruitment of Smart Vehicles for Public Sensing. *IEEE Transactions on Intelligent Transportation Systems*, PP(99):1–14, 2017.
2. Andrew T. Campbell, Shane B. Eisenman, Nicholas D. Lane, Emiliano Miluzzo, Ronald A. Peterson, Hong Lu, Xiao Zheng, Mirco Musolesi, Kristf Fodor, and Gahng Seop Ahn. The Rise of People-Centric Sensing. *IEEE Internet Computing*, 12(4):12–21, 7 2008.
3. F. Campioni, S. Choudhury, K. Salomaa, and S. G. Akl. Improved recruitment algorithms for vehicular crowdsensing networks. *IEEE Transactions on Vehicular Technology*, pages 1–1, 2018.
4. Srinivas Devarakonda, Parveen Sevusu, Hongzhang Liu, Ruilin Liu, Liviu Iftode, and Badri Nath. Real-time air quality monitoring through mobile sensing in metropolitan areas. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing - UrbComp '13*, 2013.
5. William J. Fleming. New automotive sensors - A review. *IEEE Sensors Journal*, 8(11):1900–1921, 2008.
6. R.K. Ganti, F. Ye, and H. Lei. Mobile crowdsensing: Current state and future challenges. *IEEE Communications Magazine*, 49(11):32–39, 2011.
7. S A Hamid, H Abouzeid, H S Hassanein, and G Takahara. Optimal recruitment of smart vehicles for reputation-aware public sensing. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 3160–3165, 4 2014.
8. Zongjian He, Jiannong Cao, and Xuefeng Liu. High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility. *Proceedings - IEEE INFOCOM*, 26:2542–2550, 2015.
9. Bret Hull, Vladimir Bychkovsky, Yang Zhang, Kevin Chen, Michel Goraczko, Allen Miu, Eugene Shih, Hari Balakrishnan, and Samuel Madden. CarTel : A Distributed Mobile Sensor Computing System. *The 4th ACM Conference on Embedded Networked Sensor Systems*, pages 125–138, 2006.
10. Yazhi Liu and Xiong Li. Heterogeneous participant recruitment for comprehensive vehicle sensing. *PLoS ONE*, 10(9):1–19, 2015.

11. Rajib Rana, Chun Tung Chou, Nirupama Bulusu, Salil Kanhere, and Wen Hu. Ear-Phone: A context-aware noise mapping using smart phones. *Pervasive and Mobile Computing*, 17(PA):1–22, 2015.
12. Sasank Reddy, Deborah Estrin, and Mani Srivastava. Recruitment framework for participatory sensing data collections. *Proc. Int. Conf. on Pervasive Computing (Pervasive)*, 2010.
13. L. Xiao, T. Chen, C. Xie, H. Dai, and H. V. Poor. Mobile crowdsensing games in vehicular networks. *IEEE Transactions on Vehicular Technology*, 67(2):1535–1545, Feb 2018.
14. M. Xiao, J. Wu, He Huang, L. Huang, and Chang Hu. Deadline-sensitive user recruitment for mobile crowdsensing with probabilistic collaboration. In *2016 IEEE 24th International Conference on Network Protocols (ICNP)*, pages 1–10, Nov 2016.
15. Kefu Yi, Ronghua Du, Li Liu, Qingying Chen, and Kai Gao. Fast participant recruitment algorithm for large-scale Vehicle-based Mobile Crowd Sensing. *Pervasive and Mobile Computing*, 38:188–199, 7 2017.
16. Jing Yuan, Yu Zheng, Chengyang Zhang, and Wenlei Xie. T-Drive : Driving Directions Based on Taxi Trajectories. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 99–108, 2010.