

Article

Not peer-reviewed version

Bridging the Gap: An Algorithmic Framework for Vehicular Crowdsensing

[Luis G Jaimes](#)^{*}, Craig White, [Paniz Abedin](#)

Posted Date: 10 September 2024

doi: 10.20944/preprints202409.0701.v1

Keywords: Vehicular Crowdsensing; Greedy Algorithms; Recurrent Reverse Auctions



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Bridging the Gap: An Algorithmic Framework for Vehicular Crowdsensing

Luis G Jaimes ^{*,†}, Craig White [†] and Paniz Abedin [†]

Florida Polytechnic University; cwhite@floridapoly.edu (C.W); pabedin@floridapoly.edu (P.A.)

* Correspondence: ljaimes@floridapoly.edu

† These authors contributed equally to this work.

Abstract: In this paper, we investigate whether greedy algorithms, traditionally used for pedestrian-based crowdsensing, remain effective in the context of vehicular crowdsensing (VCS). Vehicular crowdsensing leverages vehicles equipped with sensors to gather and transmit data to address several urban challenges. Despite its potential, VCS faces issues with user engagement due to inadequate incentives and privacy concerns. In this paper, we use a dynamic incentive mechanism based on a recurrent reverse auction model, incorporating vehicular mobility patterns and realistic urban scenarios using the Simulation of Urban Mobility (SUMO) traffic simulator and OpenStreetMap (OSM). By selecting a representative subset of vehicles based on their locations within a fixed budget, our mechanism aims to improve coverage and reduce data redundancy. We evaluate the applicability of successful participatory sensing approaches designed for pedestrian data and demonstrate their limitations when applied to VCS. This research provides insights into adapting greedy algorithms for the particular dynamics of vehicular crowdsensing.

Keywords: vehicular crowdsensing; greedy algorithms; recurrent reverse auctions

1. Introduction

Mobile crowdsensing systems traditionally harness the sensing capabilities of cellular phone users to collect and transmit data on various environmental parameters to a central monitoring entity. These systems leverage the extensive user base to gather significant volumes of data from previously inaccessible locations, thereby addressing large-scale societal issues [1–3]. A common application involves using pedestrian crowdsensing to monitor air quality, thereby assessing pollution levels across different geographical scales, from local neighborhoods to entire countries.

With the advent of smart mobility (connected, autonomous, and semi-autonomous), there is growing interest in transitioning from pedestrian to vehicular crowdsensing. This shift capitalizes on the increasing prevalence of vehicles equipped with advanced sensing technologies, such as GPS and environmental sensors. Vehicular crowdsensing can provide more comprehensive spatial coverage and real-time data collection compared to pedestrian systems. For instance, vehicles can continuously monitor air quality over broader areas and varying traffic conditions, offering valuable insights into pollution patterns and traffic-related environmental impacts.

A primary challenge in both pedestrian and vehicular crowdsensing systems is ensuring active participant engagement. In these systems, users typically do not receive direct benefits from reporting data, which can impact their willingness to participate [4,5]. Effective incentive mechanisms are thus crucial to sustaining participation.

Reverse auctions are a prevalent method for incentivizing participants, including in vehicular crowdsensing scenarios [6]. This approach involves participants competing to offer their sensing data, with the auctioneer purchasing the cheapest m samples. This method helps manage costs but can lead to participant attrition if not properly managed. The development of incentive mechanisms is required to mitigate this risk and ensure system reliability.

Current incentive schemes, while effective in some contexts, often overlook critical factors such as the spatial distribution of data sources, coverage, and budget constraints. For example, purchasing the m least expensive samples in each auction round might lead to acquiring data from geographically

clustered sources, resulting in redundant information. To address this, incentive mechanisms should incorporate location and coverage constraints to ensure that data is not only cost-effective but also spatially representative. Additionally, budgetary constraints must be considered, as assuming an unlimited budget is impractical.

In this paper, we investigate the application of incentive mechanisms originally designed for mobile crowdsensing with pedestrian participants within the domain of vehicular crowdsensing (VCS). Utilizing open-street maps, the Simulation of Urban Mobility (SUMO) vehicular traffic simulator, and comprehensive simulations, we demonstrate that results and insights derived from pedestrian-based crowdsensing do not directly apply to the vehicular crowdsensing context. As a general framework for data purchasing, we use a recurrent reverse auction [6]. For acquisition optimization or recruitment, we use four different greedy strategies [7] and a random one. Finally, we show that the positive results of the strategy that overwhelmingly outperformed in the context of pedestrians did not translate in the context of vehicular data collection.

The remainder of this paper is organized as follows: Section 2 provides a brief review of the relevant literature. Section 3 introduces the system model. Section 4 details the performance evaluation within the context of pedestrian-based crowdsensing. Section 4.2 extends this evaluation to the context of vehicular crowdsensing. Finally, Section 5 summarizes the findings and offers recommendations for future research. Table 1 summarizes the terms and abbreviations used in this paper

Table 1. Common Terms and Abbreviations Used in this Paper.

Type of Incentive	Example
Users or participants	Private Owners of the Sensors, Pedestrians or Vehicles
ROI	Return of Investment
SUMO	Simulation of Urban Mobility
VCS	Vehicular Crowdsensing
MCS	Mobile Crowdsensing
OSM	Open Street Map
NE	Nash Equilibrium
RADP	Reverse Auction Dynamic Price
VPC	Virtual Participation Credit
RC	Recruitment
b_i^r	Bid Price of Participant i at Round r
v_i	Virtual Participation Credit of User i
GBMC	Greedy Budgeted Maximum Cover
L	Budget
BMCP	Budgeted Maximum Coverage Problem
GIA	Greedy Incentive Algorithm, cost-effective
G	Collection of Sets
GSC	Greedy Set Cover
Bid's price	we used interchangeably with participant true valuation
Participants	Vehicles or Pedestrians
u_i	Sensor owned by user i
S_i	Set of sensors within the disk centered at u_i
w_i	weight (cardinality) of S_i
c_i	Cost of the sample provided by the sensor u_i
W_i'	Total number of elements covered by set S_i but not covered by any set in G

2. Literature Review

This section provides a brief overview of prior work and advancements in Crowdsensing.

With the rapid advancements in Internet of Things (IoT) technology, a wide range of mobile devices such as smartphones, vehicle terminals, and drones are now equipped with sophisticated sensors and wireless communication modules, including microphones, GPS, cameras, and gyroscopes. This technological progress has given rise to a novel data-sensing approach known as mobile crowd-

sensing (MCS). MCS leverages mobile devices as primary sensing units, distributing sensing tasks and collecting data through the Internet. This approach facilitates the completion of various complex and large-scale sensing and computing tasks [8,9].

In "Sensing Data Market," Chou, Bulusu, and Kanhere [10] highlight the need for incentive mechanisms to motivate individuals to share their sensing data. Additionally, they foresee the implementation of mechanisms to ensure data quality, preventing redundancy and coverage issues. However, the authors only identify these requirements as essential for establishing a sensing data market, without detailing any specific mechanisms for their implementation.

Danezis et al. [11] employed economic and psychological methods to tackle the issue of incentivizing individuals to share and utilize their location data. Participants in their experiment were made to think that a sealed-bid second-price auction was being conducted to ascertain the price at which they would reveal their location (privacy). The research identified the rates of non-participation and dropout, which are crucial for understanding the value users place on their location data. This study serves as a reference for establishing the parameters of the reverse auction scheme proposed here.

Authors like Shaw et al. [12] and Yan et al. [13] examine crowdsourcing as a method for recruitment, while Reddy et al. [14] investigate the use of micro-payments—transactions where small tasks are paired with small payments, as a form of incentive. In this study, a pilot experiment was conducted with 55 participants, utilizing 5 different types of micro-payments: a lump sum payment (MACRO), medium micro-payment (MEDIUM), high micro-payment (HIGH), low micro-payment (LOW), and competition-based (COMPETE), with the latter being the only dynamic method. The experiment aimed to gain insights into recycling practices.

Participants were instructed to take photos of the contents of outdoor waste bins and label the images with tags indicating their contents. The MACRO scheme promised individuals \$50 to participate in the study. The MEDIUM, HIGH, and LOW schemes offered 20, 50, and 5 cents per valid submission, respectively. The COMPETE scheme's payment was based on peer ranking, determined by the number of samples taken, and ranged from 1 to 22 cents per valid submission. The total payout for the micro-payments was capped at \$50 per participant. The results showed that the HIGH and MEDIUM schemes were the most successful, while the MACRO and LOW schemes resulted in fewer photo submissions, and several participants in the COMPETE group dropped out. This study indicates that flat rates, like the MACRO scheme, do not encourage user participation, and dynamic schemes like COMPETE lead to fatigue and participant dropout.

Experiments conducted by Lee and Hoh [6] highlight the challenge of determining the appropriate amount for fixed pricing (micro-payments), as a high price can render the strategy economically unviable, while a low price can discourage user participation. They also found that dynamic schemes tend to lose participants over time. To address these issues, they propose the Reverse Auction Based Dynamic Price Scheme (RADP-VPC-RC), which allows individuals to sell their sensing data through a dynamic price reverse auction system. This scheme incorporates several mechanisms to maintain a minimum number of participants, accommodate people with different reward expectations, and recruit members who have dropped out, addressing common issues in recurrent reverse auction systems. Although RADP-VPC-RC reduces sample prices through participant competition, it does not tackle problems of data redundancy, coverage, and budget constraints, which are the focus of our study. The RADP-VPC-RC scheme is further detailed in the next section, as our GIA [15] scheme is based on it.

Yang et al. [16] introduced the concept of platform-based mobile crowdsensing. In this work, the authors model a crowdsensing system as a Stackelberg game, a sequential game where players are divided into leaders and followers. The game is structured as a two-level model, with the platform as the leader (first player) and the vehicles as the followers. The second level of the game corresponds to a non-cooperative scenario where the participants (players) compete against each other to maximize their profits. However, Yang's work considers only one platform and multiple players Chakeri et al. [17,18] extend Yang et al.'s work by developing a greedy algorithm that enables the implementation of

a crowdsensing system using multiple platforms and participants per platform. This foundational work would later become crucial for advancing vehicular crowdsensing systems.

Building upon the work of Chakeri et al. [17], another branch of vehicular crowdsensing research utilizes Nash Equilibrium (NE), a key concept in game theory. In this approach, the crowdsensing system is modeled as a game, where players assume two distinct roles: participant vehicles and the platform. The platform's role involves designating areas of sensing interest and assigning corresponding rewards, while the participants are responsible for data collection. Representative work in this field includes studies such as [5,19–21]. Unlike the work of Chakeri and Yang, where participants are static and the games are one-shot, this new body of work focuses on evolutionary games. In these games, vehicular players are dynamic, requiring them to adjust their strategies as the game progresses.

One of the other main research challenges in Mobile Crowd Sensing (MCS) involves ensuring sufficient coverage of sensing tasks, which reflects the effectiveness of the incentive mechanism designed to maximize this coverage [5]. Another key challenge is determining the optimal number of mobile users required to complete a given task within a defined area of interest [22]. Several studies have been conducted to improve coverage quality through the implementation of budget-constrained optimal participant selection strategies. The authors Zheng et al. [23] and Zhang et al. [24] proposed a greedy method to recruit an optimal set of mobile users who can cover the entire region of interest while keeping within the system's budget constraints. Meanwhile, Xu et al. [25] presented budget-feasible frameworks for selecting an ideal group of mobile users to maximize continuous-time interval coverage while adhering to budgetary limits. In this paper, we address the dual challenges of ensuring sufficient coverage of sensing tasks and effectively incentivizing participants in Mobile Crowd Sensing (MCS). We focus on developing strategies that not only maximize the coverage of sensing tasks but also create robust incentive mechanisms to engage and retain participants.

3. System Model

We model our VCS system as a game with two main players: A crowdsourcer or data buyer C_i , and a set of participant vehicles $V = \{v_1, v_2, \dots, v_N\}$. The crowdsourcer aims to collect a representative set of samples $S = \{s_1, s_1, \dots, s_M\}$ with $M \leq N$ of a given region G_i to reconstruct a variable of interest. For that purpose, the crowdsourcer uses a limited budget L to outsource the data collection to a set of interested participants V . Given the limited budget available, the crowdsourcer cannot purchase all the samples offered by participants at any given time, and its goal is to acquire a sub-collection of samples $\mathcal{S} = \{S_1, S_2, \dots, S_M\}$ such that $\bigcup_{S \in \mathcal{V}} S \approx V$ and $|\mathcal{S}|$ is within budget L . For that purpose, we use a model with three main components: A geometric sensor model, a multi-round reverse auction method for data buying or acquisition, and several recruitment approaches for selecting the winners of each round.

3.1. Geometrical Sensor Model

To address the coverage problem, we utilize the following geometric disk model:

$$f(d(u_i, u_j)) = \begin{cases} 1 & \text{if } d(u_i, u_j) \leq r \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

In this model, $d(u_i, u_j)$ denotes the Euclidean distance between sensor u_i and sensor u_j , and $r > 0$ is a constant that defines the coverage radius for each sensor. This function characterizes a disk centered at sensor u_i with radius r . Sensors located within this disk are assigned a coverage measure of 1 and are considered *covered* by sensor u_i . In contrast, sensors situated outside this disk have a coverage measure of 0 and are regarded as *not covered* by u_i . Figure 1 illustrates the coverage areas of sensors 1 and 7. The fundamental premise is that the sensing value of any environmental variable within a disk of radius r is uniform. Consequently, if sensing samples are obtained from sensors 1 and 7, it is redundant to collect additional samples from sensors 2, 3, 4, 5, and 6, as their data would be similar.

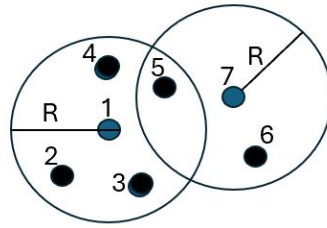


Figure 1. Example of coverage per user.

3.2. Data Purchasing Model

For data purchasing, we use the reverse Auction-based Dynamic Price with Virtual Participation Credit, and re-join (RADP-VPC-RC) [6,15,26,27]. Algorithm 1 describes a round of multi-round auction-based mechanism.

Algorithm 1 RADP-VPC-RC Incentive Mechanism

```

1: Initialize auction round  $r \leftarrow 1$ 
2: Initialize virtual participation credit  $v_i \leftarrow 0$  for all users  $i$ 
3: Set number of required participants  $Q$ 
4: while service quality is not guaranteed do
5:   for each user  $i$  do
6:     Submit actual bid  $b_i^r$ 
7:     Calculate competition bid  $b_i^r \leftarrow b_i^r - v_i$ 
8:   end for
9:   Sort users based on competition bids  $b_i^r$ 
10:  Select the top  $Q$  users as winners
11:  for each winner  $i$  do
12:    Pay  $b_i^r$  to participant  $i$ 
13:    Reset  $v_i \leftarrow 0$ 
14:    increase the  $b_i \leftarrow 1.1b_i$  with 50% of probability
15:  end for
16:  for each loser  $i$  do
17:    Update virtual participation credit  $v_i \leftarrow v_i + \alpha$ 
18:    decrease the  $b_i \leftarrow 0.8b_i$  with 50% of probability
19:    evaluates ROI if  $ROI \geq 0.5$  stays, otherwise it drops
20:  end for
21:  Notify winners and losers
22:  Send the highest bid price to the dropped users for recruiting
23:  Dropped users evaluate their ROI, if  $ROI \geq 0.5$ , return with 50% of probability in next round.
24:   $r \leftarrow r + 1$ 
25: end while

```

Line 1 initializes the first round, and line 2 sets to zero the virtual participation credit v_i for all participants i . In Lines 4 to 8 every active participant submits its bid's price for its sample, however, its competition bid's price corresponds to its bid's price minus its accumulated virtual participation credit. Lines 9 and 10 show how to select the winners for the current round r (the selection criteria could be modified according to the recruitment algorithm). Lines 10-14 describe how the virtual participation credit of winners is set to zero, and randomly half of the winners increase their bid's price by 10%. Lines 16-19 describe how the losers' virtual participation credit is increased, resulting in a more competitive bid price and more chances to win in the next round. In addition, randomly half of the losers decrease their bid's price to improve the winning chances for the next round. Finally lines 20-23 describe the re-join process. Here, the system sends to the dropped users the highest bid price which is used for them to evaluate their (Return of Investment) ROI, if it is greater than 0.5 then they randomly return to the next round with 50% of probability. Finally, line 24 signals the transition to the next round where the whole process re-takes place. A detailed description of some elements of this reverse auction such as the ROI and others can be found in [15].

3.3. Recruitment Model

We compare four approaches to select each round's winners, specifically, the participants from whom the samples are purchased. The first approach is the Greedy Budgetive Maximum Cover (GBMC) or cost-effective. The second approach corresponds to a pure greedy strategy that acquires the cheapest samples until all the samples are obtained or the budget is exhausted. The third approach uses a GBMC with enumeration with $k=1$ and $k=2$. Finally, the fourth approach is a random acquisition method in which no optimization strategy is used to select the winners for each round. The next sub-section briefly reviews each of the recruitment strategies.

3.3.1. Greedy Budgeted Maximum Coverage Algorithm (GBMC) or the Cost-Effective

GIA [15] uses RADP-VPC-RC for data acquisition, and the Greedy Budgeted Maximum Coverage Algorithm (GBMC) [7] for recruitment. For the reader's ease, we present the GBMC Algorithm 2.

Algorithm 2 The Greedy Budgeted Maximum Coverage Algorithm for GIA.

```

1: Input:  $S$  a collection of sets made up by the user locations
2: Output:  $S', C, S$ , covering set
3: function GREEDY AREA PRICE( $S$ )
4:    $G \leftarrow \emptyset$ 
5:    $C \leftarrow 0$ 
6:    $U \leftarrow S$ 
7:   while  $U \neq \emptyset$  do
8:     select  $S_i \in U$  that maximizes  $\frac{W'_i}{c_i}$ 
9:     if  $C + c_i \leq L$  then
10:       $G \leftarrow G \cup S_i$ 
11:       $C \leftarrow C + c_i$ 
12:     end if
13:      $U \leftarrow U \setminus S_i$ 
14:   end while
15:   return  $G$ 
16: end function
17: function GREEDY AREA( $S$ )
18:    $G' \leftarrow \emptyset$ 
19:    $U' \leftarrow S$ 
20:   while  $U' \neq \emptyset$  do
21:     select  $S_j \in U'$  that maximizes  $W'_j$ 
22:     if  $C' + c'_j \leq L$  then
23:        $G' \leftarrow G' \cup S_j$ 
24:        $C' \leftarrow C' + c'_j$ 
25:     end if
26:      $U' \leftarrow U' \setminus S_j$ 
27:   end while
28:   return  $G'$ 
29: end function
30: function GBMC
31:    $G \leftarrow$  GREEDY AREA PRICE( $S$ )
32:    $G' \leftarrow$  GREEDY AREA( $S$ )
33:   if  $w(G) \geq w(G')$  then
34:      $S' = G$ 
35:   else
36:      $S' = G'$ 
37:   end if
38:   return  $S'$ 
39: end function
40:

```

Algorithm 2 corresponds to the Greedy Budgeted Maximum Coverage algorithm for GIA. The input of GBMC corresponds to the collection of disks S_i with radius R located at the center of every sensor u_i . In addition, w_i is the associate weight w_i of S_i or the number of samples covered by S_i including u_i . The output corresponds to the subset $S' \subseteq S$ that maximizes the coverage within the budget L . The algorithm has two sub-routines: Greedy area price lines 3-16, and Greedy area lines 17-30. Greedy area price initializes the collection in G to empty in line 4 and sets the initial spending value C to zero in line 5. Line 6 sets the universal set to input collection S . Lines 7-13 present the optimization criteria to select a winner, namely buy the sample u_i with sample price c_i that maximizes $\frac{W'_i}{c_i}$ where W'_i is the number of samples within S_i including u_i but not covered by any set in G . In other words, select the sample price with the best coverage per dollar. Greedy area price, the second sub-routine follows the same logic as the first one, it only differs in the optimization criteria to select the winner, Here, we buy the sample u_j whose S_j maximizes W'_j , namely, maximizes the number of samples within S_i including u_i but not covered by any set in G . The GBMC runs the two sub-routines in parallel and selects the output that maximizes the number of samples covered. Finally, the GIA algorithm corresponds to a modified version of RADP-VPC-RC (Algorithm 1) where the criteria for winner selection (lines 9-10) is substituted by the output of the GBMC (Algorithm 2).

3.3.2. An Example of How GBMC for GIA Works

The following example shows how GBMC works and how it compares to the well-known Greedy Set Cover (GSC). Consider a scenario where N users are deployed in a target area with a given radius R . Let's take three users (sensors), u_1 , u_2 , and u_3 , covering 1, 5, and 7 users respectively, including themselves. In other words, their samples have weights $W'_1 = 1$, $W'_2 = 5$, and $W'_3 = 7$. Assign the costs $c_1 = 1$, $c_2 = 3$, and $c_3 = 6$ for these samples, and define the total budget $L = 6$.

Algorithm 3 Improved k -Greedy Algorithm

```

1: Input:  $S$  a collection of sets,  $L$  budget,  $k$  an integer
2: Output:  $H$  a collection of sets
3:  $H_1 \leftarrow \arg \max\{w(G) \mid G \subseteq S, |G| < k, \text{ and } c(G) \leq L\}$ 
4:  $H_2 \leftarrow \emptyset$ 
5: for all  $G \subseteq S$  such that  $|G| = k$  and  $c(G) \leq L$  do
6:    $U \leftarrow S \setminus G$ 
7:    $R \leftarrow \text{Greedy}(U, L, c(G))$ 
8:   if  $w(R \cup G) > w(H_2)$  then
9:      $H_2 \leftarrow R \cup G$ 
10:  end if
11: end for
12: if  $w(H_1) > w(H_2)$  then
13:   return  $H_1$ 
14: else
15:   return  $H_2$ 
16: end if

```

Greedy Set Cover Algorithm Solution

The Greedy Set Cover (GSC) algorithm tries to maximize $\frac{W'_i}{c_i}$ for $i = 1, 2, 3$:

$$\frac{W'_1}{c_1} = \frac{1}{1} = 1,$$

$$\frac{W'_2}{c_2} = \frac{5}{3} \approx 1.67,$$

$$\frac{W'_3}{c_3} = \frac{7}{6} \approx 1.17.$$

The GSC algorithm would first purchase u_2 because $\frac{W'_2}{c_2}$ is the highest. However, after purchasing u_2 , the remaining budget is $6 - 3 = 3$, which is not enough to purchase u_3 .

GBMC Algorithm Solution

In contrast, the GBMC algorithm evaluates two sub-routines (lines 32-33):

GREEDY AREA PRICE: Similar to GSC, it would select u_2 , covering 5 users.

GREEDY AREA: It selects u_3 because W'_3 is the highest without violating the budget. It covers 7 users.

GBMC then compares the results of these sub-routines and ultimately chooses u_3 , as it covers more users within the budget. Thus, the final decision by GBMC is to purchase u_3 , which covers 7 users, maximizing the coverage better than the GSC algorithm under the given budget constraint.

3.3.3. k -Greedy Algorithm

Algorithm 3 (k -greedy) enhances GIA by incorporating a brute-force component. Initially, initializes two sets H_1 , and H_2 (lines 3-4). It sets H_1 with the singleton subset that maximizes the weight (line 3), provided the cost is within the budget. Then, it starts H_2 by choosing a constant k and finds all k -size subsets of the participant group that can be afforded. For each subset, the GIA algorithm is executed, but it is forced to first select all participants in the k -subset. Finally, the algorithm compares the weights of H_1 , and H_2 (lines 12), and outputs the H value with the higher weight. Thus, higher values of k significantly slow down the algorithm but yield results closer to the optimal. Section 3.3.4 provides a numerical example that details the step-by-step process of how k -greedy works.

3.3.4. An Example About How k -Greedy Works

The following example illustrates the difference between GIA and greedy $k=2$. Let's $X = \{e_1, e_2, e_3\}$ be the set of samples associated to a set of mobile sensors (participants smartphones), and $S = \{S_1, S_2, S_3\}$ be the collection of sets, where each S_i corresponds to the set of samples inside the disk centered at sensor e_i . Following our convention about the weight of sets (number of elements), and its cost (cost of sample e_i at the center of the disk), then we have:

$$S_1 = \{e_1, e_2\} \text{ with weight } w(S_1) = 2 \text{ and cost } c(S_1) = 24$$

$S_2 = \{e_1, e_2, e_3\}$ with weight $w(S_2) = 3$ and cost $c(S_2) = 42$
 $S_3 = \{e_2, e_3\}$ with weight $w(S_3) = 2$ and cost $c(S_3) = 24$

Let $L=100$ be the budget. Then, the total cost of the acquired samples should not exceed 100.

GIA Execution:

Select the highest weight set: (Algorithm 2 - line 34)

$S_1 = \{e_1, e_2, e_3\}$ with weight $w(S_2) = 3$ and cost $c(S_2) = 2$.
 Current weight = 3.
 Remaining budget = 0.
 Remaining sets: $\{S_1, S_3, S_4\}$.

No budget remaining to select another set:

Total weight = 3 (from S_2).

K = 2-Greedy Execution:

Calculate H_1 :(Algorithm 3 -line 3)

Single sets: S_1, S_2, S_3, S_4 .
 Maximum weight from single sets = 3 from S_2 .
So, $H_1 = \{S_2\}$ with weight 3.

Calculate H_2 :(Algorithm 3 -lines 4-11)

Consider all subsets with cardinality exactly $k = 2$:

$\{S_1, S_3\}$ with weight $2 + 2 = 4$ and total cost $1 + 1 = 2$.
 $\{S_1, S_4\}$ with weight $2 + 1 = 3$ and total cost $1 + 1 = 2$.
 $\{S_3, S_4\}$ with weight $2 + 1 = 3$ and total cost $1 + 1 = 2$.

Maximum weight from these combinations = 4 from $\{S_1, S_3\}$.
So, $H_2 = \{S_1, S_3\}$ with weight 4.

Output the better solution between H_1 and H_2 :

H_1 has weight 3, and H_2 has a weight of 4.
Therefore, the output is $H_2 = \{S_1, S_3\}$ with weight 4.

Thus, while GIA ends up acquiring S_2 with a weight of 3, k -Greedy acquires $\{S_1, S_3\}$ with a weight of 4, maximizing weight and budget utilization.

3.3.5. Pure Greedy

This approach uses the RADP-VPC-RC for data acquisition and the Pure Greedy Algorithm for winner selection. This simple strategy repeatedly selects the cheapest sample regardless of its location until the budget is exhausted. For the ease of the reader, we sketch the Pure Greedy Algorithm. Here, the Cheapest-first Greedy Algorithm corresponds again to a modified version of RADP-VPC-RC (Algorithm 1) where the criteria for winner selection (lines 9-10) is substituted by the output of the Algorithm ??.

Algorithm 4 Pure Greedy Algorithm

```

1: Input:  $S$  a collection of sets made up by the user locations
2: Output:  $G, C$  covering set
3: function PURE GREEDY ( $S$ )
4:    $G \leftarrow \emptyset$ 
5:    $U \leftarrow S$ 
6:   while  $U \neq \emptyset$  do
7:     select  $S_i \in U$  with the lowest  $c_i$ 
8:     if  $C + c_i \leq L$  then
9:        $G \leftarrow G \cup S_i$ 
10:       $U \leftarrow U \setminus S_i$ 
11:     end if
12:   end while
13:   return  $G$ 
14: end function

```

3.3.6. Random Algorithm

We also tested a random selection of participants. This algorithm repeatedly chooses randomly from a uniform distribution among the participants that it can afford until the budget is exhausted. The algorithm does not have any consideration in terms of location, price, or any optimization criteria.

Algorithm 5 Random Selection Algorithm

```

1: Input:  $S$  a collection of sets made up by the user locations
2: Output:  $S, C, S_c$  covering set
3: function RANDOM SELECTION( $S$ )
4:    $G \leftarrow \emptyset$ 
5:    $C \leftarrow 0$ 
6:    $U \leftarrow S$ 
7:   while  $U \neq \emptyset$  do
8:     select  $S_i \in U$  based on a random selection of  $u_i$ 
9:     if  $C + c_i \leq L$  then
10:       $G \leftarrow G \cup S_i$ 
11:       $C \leftarrow C + c_i$ 
12:     end if
13:      $U \leftarrow U \setminus S_i$ 
14:   end while
15:   return  $G$ 
16: end function

```

4. Performance Evaluation

This section outlines the performance evaluation for two sets of experiments. In the first set, the participants are pedestrians, whereas in the second set, the participants are represented as vehicles.

4.1. Pedestrian-Based Crowdsensing

For pedestrians-based crowdsensing, we evaluate two recruitment models: cheapest first (pure-greedy) which corresponds to Algorithm ??, and cost-effective (Maximum budgeted greedy) represented by Algorithm 2 for recruitment, and the recurrent reverse auction RADP-VPC-RC presented in Algorithm 1 as a data purchasing model.

4.1.1. Experimental Setup

Table 2 outlines the parameters used in the experimental setup. The simulations take place within a square area of 100 x 100 units. The spatial distribution of participant locations follows three different scenarios: 2D normal distribution, uniform distribution, and exponential distribution. These distributions represent potential real-world situations. The 2D normal distribution for participant locations and bid prices simulates a stratified scenario, where a population is clustered in distinct neighborhoods (locations following a 2D normal distribution) with varying socioeconomic statuses (bid prices following a 1D normal distribution). In this case, participants in lower-income neighborhoods are expected to bid lower for their sensing samples compared to those in higher-income areas. Conversely, the uniform distribution for both participant locations and bid prices represents a scenario where neither location nor socioeconomic status has any influence on the participants' asking prices per sample. Finally, the exponential distribution for locations and bid prices simulates a scenario where participants are densely clustered in certain regions, indicating higher population concentrations or activity levels. In such areas, increased competition may drive down bid prices, while participants in less populated or active regions might bid higher due to lower competition. The experiment employs a synchronous random model to handle participant mobility. Initially, participants are randomly deployed. For each round, the mobility algorithm updates their locations by calculating a displacement vector (\vec{R}), which combines the X-axis displacement (\vec{X}) and Y-axis displacement (\vec{Y}), such that $\vec{R} = \vec{X} + \vec{Y}$. The displacement vectors for each axis are generated using a random function (i.e., $(-1)^{\text{uni}(0,2)}$), determining the movement direction. The speed is determined by the vector's magnitude, which is derived by multiplying the number of steps (randomly selected between 0 and 3) by the step length (randomly selected between 0 and 11).

4.1.2. Evaluated Metrics

We evaluate the following metrics: sensor radius, number of active participants, coverage, and percentage of budget utilization.

Sensor radius: We defined a sensor's radius as the constant r that defines the area of coverage of each sensor. Thus, the acquisition of any sample s_j in round k within the area coverage of sensor i is redundant if sample s_j was already acquired in the same round.

Coverage: Coverage is defined not by the area, but by the number of participants (mobile sensors) within the area covered by the combined range of individual sensors associated with the acquired samples.

Number of Active Participants: This metric represents the average number of participants who continue to engage after m rounds of recurrent reverse auctions. Any crowdsensing system must maintain a minimum number of active participants to function effectively. A common issue with recurrent reverse auctions is the decline in participant numbers over time. This reduction leads to a spike in bid prices due to a lack of competition.

Percentage of Budget Utilization per Round: This metric assesses the average percentage of the budget used per round. A common issue is that a significant portion of the budget remains unused each round, leading to poor overall budget optimization.

4.1.3. Experiments

This section describes a set of experiments conducted to test the performance of *GIA* optimizing a set of proposed metrics. The first set of experiments focuses on identifying the optimal radius length r . After establishing the ideal r , the next set of experiments involves comparing the average number of active participants and the average cost per round between the RADP-VPC-RC using a pure-greedy recruitment strategy and *GIA*. Several distributions for users' true valuations and participant's locations were used. The final set of experiments aims to showcase the *GIA* algorithm's superiority in terms of coverage, especially when participants are positioned in stratified locations within the area of interest.

4.1.4. Experiment 1: Determining the Ideal Length of Sensor Radius r

To explore the relationship between radius r , the number of active participants, and the average price per round, r was incrementally increased from 1 to 10. Each scenario involved 1,000 rounds, repeated 50 times, with results averaged under a fixed budget of 100 units per round. The average number of active participants was recorded for each value of r using three different distributions for users' true valuations and a uniform initial deployment. The goal was to identify the radius at which participant numbers declined significantly. As shown in Figure 2 (right), when r exceeds 5, the number of active participants decreases consistently across all distributions. This occurs because a larger radius can encompass the entire population within one user's sensing range, reducing the number of samples purchased and potentially leading to poor data representation. Conversely, a smaller radius may result in redundant data due to overlapping coverage. Thus, finding the right balance for r is crucial for optimizing data granularity and quality.

Figure 2(left) shows that when r exceeds 5, The *GIA* algorithm fails to fully use the allocated budget per round. However, when $r \leq 5$, the algorithm continues using the entire budget. By combining the findings from these figures, we conclude that $r = 5$ strikes a balance between maintaining a good level of participation and minimizing redundant data. In practice, however, the value of r should be adjusted based on how the variable of interest varies with distance.

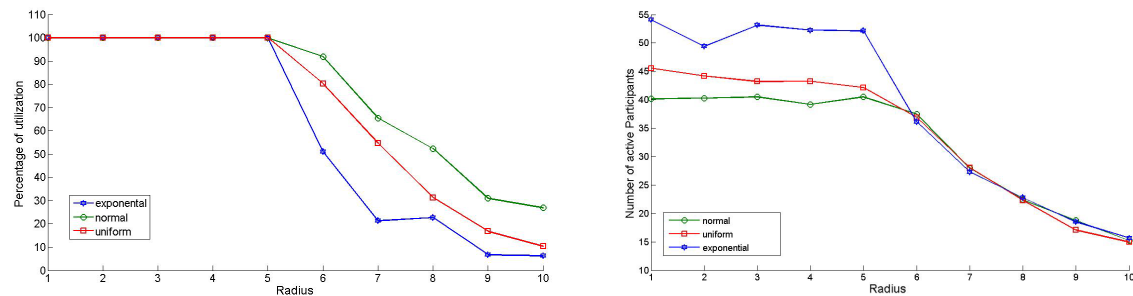


Figure 2. Radius vs Percent utilization (left) and Number of Participants (right).

Table 2. Parameters for Simulation Set 1.

Parameters	Experiment 1	Experiment 2	Experiment 3
Deployment area	100m x 100m		
Instances	100		
Deployment distribution	Uniform		
Deployment distribution	Four normal distributions with parameters:		
	$\mu_1 = (30, 80),$ $\mu_2 = (80, 80),$ $\mu_3 = (50, 50),$ $\mu_4 = (90, 30),$ $Cov = \begin{pmatrix} 40 & 0 \\ 0 & 40 \end{pmatrix}$		
Uniform true valuation distribution	[0,10]		No
Normal true valuation distribution	$\mu = 5, \sigma = 2$		Yes
Exponential true valuation distribution	$\mu = 5$		No
Normal true valuation distribution	$\mu_1 = 5, \mu_2 = 10, \mu_3 = 15, \mu_4 = 20, \sigma = 2$		
RADP-VPC-RC	No	Yes	Yes
GIA	Yes	Yes	Yes
Radius R	1:10	5	5
Budget per round	100	20:200	0:350
Beta	(3,7)	(3,7)	(3,7)
Alpha	Not used	7	Not used
ROI Threshold	0.5	0.5	0.5

4.1.5. Experiment 2: Comparing the Performance Metrics.

The purpose of this set of experiments is to quantify the number of active users within the system as the budget per round increases from 20 to 200 after 100 rounds. This analysis was conducted under three distinct user valuation distributions (sample bid prices), and participant distribution on the target area: a normal distribution with parameters $\mu = 5$ and $\sigma = 2$, an exponential distribution with $\mu = 5$, and a uniform distribution within the range [0,10]. The sensor radius was set to five

Figure 3 (left), Figure 3 (center), and Figure 3 (right) illustrate that GIA and RADP-VPC-RC-pure-Greedy algorithms exhibit comparable performance, acquiring a similar number of samples per

round across various budgets. These findings indicate that both algorithms effectively facilitate user participation and re-engagement within the system. However, despite the similar number of active users, the GIA algorithm adopts a distinct sampling strategy by avoiding the purchase of samples from proximate locations, thus preventing the acquisition of redundant data. Details of the experimental parameters are provided in Table 2.

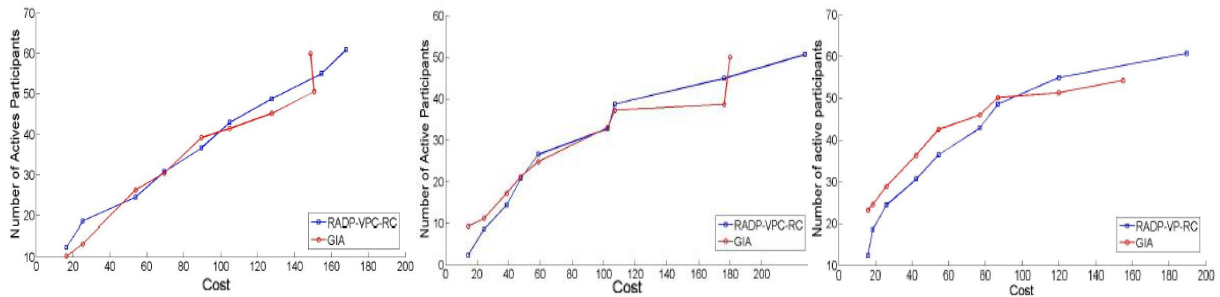


Figure 3. Cost vs Number of Active participants under Normal (left), exponential (center), and uniform (right) distributions.

4.1.6. Experiment 3: Coverage, Number of Active Participants, and Cost

Here, we want to evaluate the performance of the GIA and RADP-VPC-RC-pure-greedy algorithms in terms of coverage, number of active participants, and cost versus the number of acquired samples. To achieve this, a stratified scenario is simulated, representing a situation where participants from varying socioeconomic statuses, each with distinct true valuations, reside in different regions of a city. In this context, the RADP-VPC-RC-pure-greedy algorithm tends to select samples that are not representative of the entire population and instead acquires redundant data. Conversely, the GIA algorithm successfully obtains non-redundant samples from each user and location type. The initial user deployment locations were generated using a 2D normal distribution, forming four clusters with means $\mu_1 = (30, 80)$, $\mu_2 = (80, 80)$, $\mu_3 = (50, 50)$, and $\mu_4 = (90, 30)$, and a common covariance matrix $\text{Cov} = \begin{pmatrix} 40 & 0 \\ 0 & 40 \end{pmatrix}$. Each cluster comprised 25 locations, totaling 100 locations. Users' true valuations were generated from normal distributions with $\mu_1 = 5$, $\mu_2 = 10$, $\mu_3 = 15$, $\mu_4 = 20$, and $\sigma = 2$. Each valuation cluster was associated with a corresponding location cluster to simulate the expected scenario.

Figure 4 (left) illustrates the percentage of area covered, Figure 4 (center) shows the number of active participants, and Figure 4 (right) displays the total cost per round as the number of acquired samples ranges from 5 to 50 in increments of 5. From Figures 4 (center) and 4 (right), it is evident that both algorithms maintain a similar number of active participants and incur comparable expenditures to acquire their samples. However, Figure 4 demonstrates that the GIA algorithm selects samples from more widely distributed users across the area of interest, thereby achieving better coverage and minimizing redundant data acquisition. For instance, Figure 4 (left) indicates that when fifty samples are acquired, the GIA algorithm covers up to 64% more area than RADP-VPC-RC-pure-greedy.

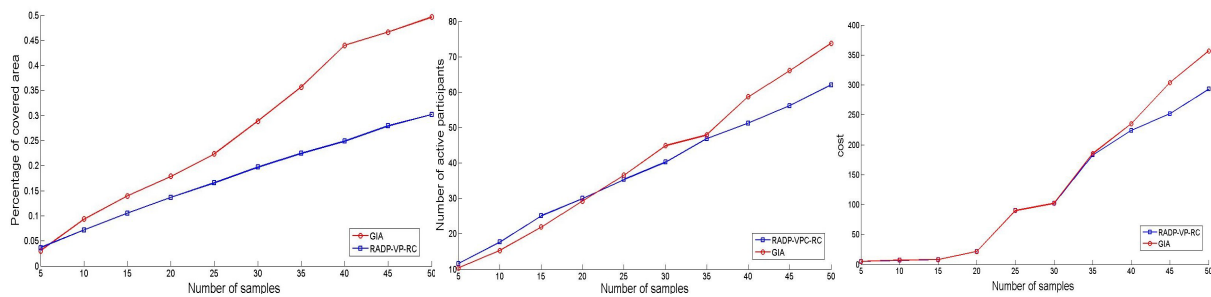


Figure 4. Number of samples vs Percentage area coverage (left), number of active participants (center), and cost (right).

4.2. Performance Evaluation for Vehicular Crowdsensing

In this section, we assess the effectiveness of four different recruitment or optimization approaches: Pure-Greedy (Section 3.3.5), Cost-Effective or GIA (Section 3.3.1), k-Greedy (Section 3.3.3), and Random Selection (Section 3.3.6). We use the RADP-VPC-RC as our purchasing model (Section 3.2) and the sensor model described in Section 3.1.

We evaluate the following metrics: Number of active participants (i.e., users who remain engaged after 100 rounds of the recurrent reverse auction), coverage, and average total expenditure. The number of active participants reflects the sustained engagement and interest of users over time. Coverage assesses how effectively the auction mechanism captures a representative set of samples of the target areas of interest. Average total expenditure offers insights into the financial implications of the auction process. By analyzing these metrics, we aim to gain a better understanding of their performance and impact within the context of vehicular crowdsensing.

4.2.1. Experiments Setup

The elements of our simulation environment are illustrated in Figure 5. This environment comprises a map, a vehicular traffic simulator, and data on streets' velocity distributions. We selected a densely mapped subsection of London's drivable streets, enabling participants to take alternate routes and reach their destinations through different trajectories. The mapping is based on OpenStreetMap (OSM) [28], while the traffic simulation is conducted using SUMO [29]. SUMO provides realistic vehicle movement and routing algorithms applied to OSM data. Before importing a map into SUMO, we preprocess it using a discretization approach to correct any inconsistencies, following the method described by Goss et al. [30,31]. The parameters for this set of experiments can be found in Table 3. Additionally, we utilize the Uber Movement dataset [32] to estimate the velocity distribution of streets, which allows us to calculate average commuting times between any two locations on the map [33].

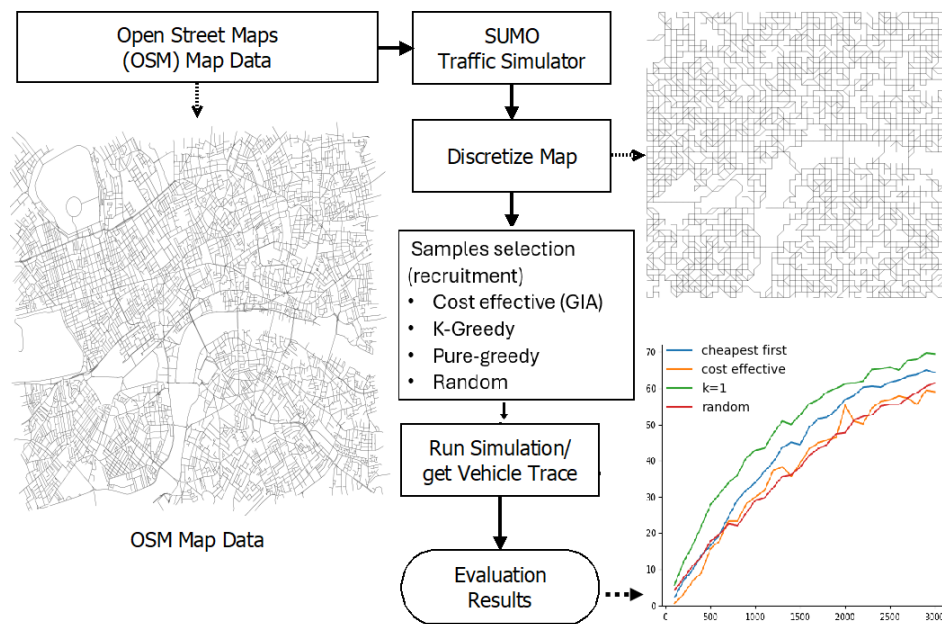


Figure 5. Simulation components

Table 3. Vicular-based Crowdsensing Simulation Parameters

Parameters		Parameters (cont.)	
Target Area	5200 x 5200 m	Area Coordinates	
Cell Size	100 x 100 m	<i>Low Lat</i>	51.480861
Reward Distribution	normal, uniform, exp distributions	<i>Low Lon</i>	0.176478
Vehicles trajectory deployment		<i>High Lat</i>	51.539079
<i>Amount</i>	100	<i>High Lon</i>	-0.057695
<i>Source</i>	2D normal, uniform, exp distributions		
<i>Destination</i>	2D normal, uniform, exp distributions		
Algorithms	<i>Budget for ExperimentS 1-5</i>	<i>Experiment 50</i>	
<i>Cheapest-first</i>	range(start=100, end=3000, step=500) . . .		—
<i>cost-effective</i>	range(start=100, end=3000, step=500) . . .		—
<i>k-greedy, k=1</i>	range(start=100, end=3000, step=500) . . .		—
<i>random</i>	range(start=100, end=3000, step=500) . . .		—

In these experiments, vehicles continuously traverse the map, joining and leaving the participation pool as they evaluate their return on investment (ROI) after each round. For each participant, we selected two points (a source and a destination) on a road map of Cologne, Germany, and had them drive between these two locations in a loop. In all experiments, the participants were split equally between the various districts, which corresponded to the centers of the multivariate exponential or normal distribution clusters used to place the participants. Unlike the previous pedestrian-based crowdsensing experiments where participant locations were drawn from a random distribution, it is not possible to sample two points from the distribution in this case, as most points will not fall directly on a road. Instead, we find the midpoint of each road, then evaluate the probability density function at that point and use that number as a weight for the road. From there, the probability that a road will be sampled is its weight divided by the sum of all weights. The covariance matrix used for the 2D normal distribution is $\begin{bmatrix} 150 & 0 \\ 0 & 150 \end{bmatrix}$. For the exponential distribution, each road midpoint's distance from the

center of the cluster was calculated, and then the weight was calculated by evaluating the probability density function of an exponential distribution. The density function used was: $\frac{(e^{-d})}{10}$ where d is the distance

We conducted experiments to test the performance of our four recruitment approaches under five different scenarios. These scenarios corresponded to various combinations of participant trajectory locations and participant true valuation (participant sample expected value) distributions. In one scenario, we assumed that participants' true valuation is heavily influenced by their location (location of their vehicles' trajectories). This scenario may correspond to a city which divided into socioeconomically stratified neighborhoods or districts with different income levels. We modeled this by deploying vehicle trajectories following a 2D normal distribution (clusters) and generating participant true valuations also following a normal distribution. In a different scenario, we assumed participants' true valuations were not influenced by the locations of vehicle trajectories. We simulated this by using a 2D normal distribution and a uniform distribution for the deployment of vehicle trajectories while generating participant true valuations from a uniform distribution. The first case may represent a city with an equitable income distribution, where regular vehicle trajectories are not correlated with participants' true valuation. In the second sub-scenario, while participant locations remain unimportant, society may be stratified into different socioeconomic levels. This reflects a scenario in which city planners distribute public or subsidized housing units evenly throughout the city. Finally, we represent a scenario in which traffic density may influence the participants' true valuations. The idea is that in areas with heavy traffic, the competition among participants' vehicles may drive sample prices down, while the opposite may happen in areas with less vehicular activity. We simulate this scenario by using an exponential distribution for both the deployment of vehicle trajectories and the generation of participant true valuations.

4.3. Experiment 2 - Impact of Budget Allocation on Coverage, Participant Engagement, and Budget Utilization with Correlated Bid Prices and Trajectory Location

This experiment assesses the impact of budget allocation on coverage, the active number of participants, and the efficiency of budget utilization after 100 rounds of RADP-VPC-RC. We tested four recruitment approaches: pure greedy (selecting the cheapest options first), cost-effective (GIA), K-greedy (K=1), and random selection. The experiment tested every budget level from 100 to 3000, in 100 increments, conducting 20 trials for each budget level. Figure 6 shows the results of evaluating these three metrics when both vehicle trajectories deployment and participant true valuations follow normal distributions.

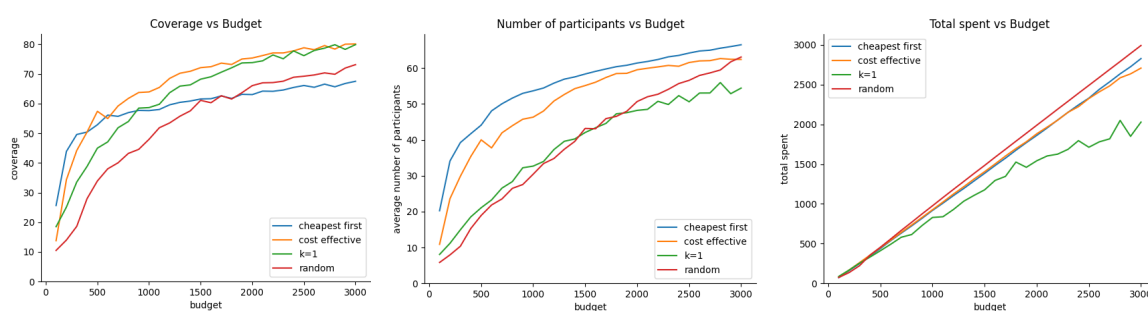


Figure 6. Normal distribution for trajectory distribution and participant true valuations.

On the other hand, Figure 7 shows the performance of our three metrics when vehicle trajectory deployment and participant bid prices follow exponential distributions.

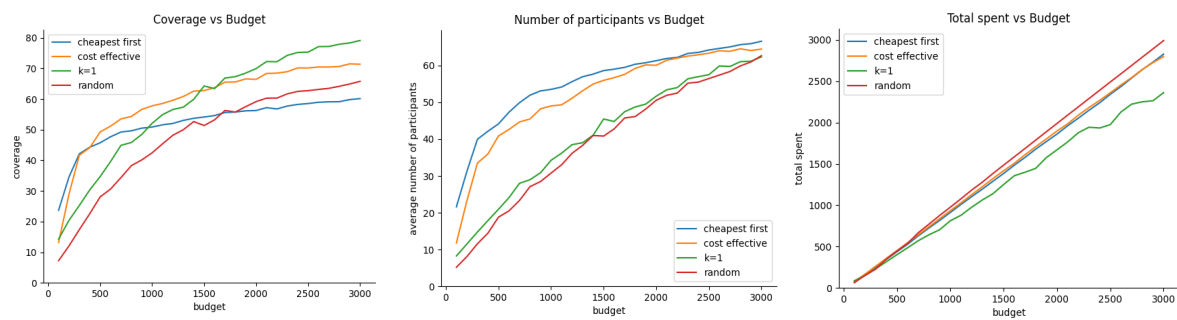


Figure 7. Exponential distribution for trajectory distribution and participant true valuations..

Figure 6 shows that the cost-effective or GIA approach performs better than other recruitment approaches for budget values greater than 300. However, the cost-effective approach is closely followed by k -greedy. This result found on vehicle simulations is consistent with the one found with pedestrians as shown in Figure 4 (left). In terms of the average number of participants after 100 rounds, the cheapest-first is slightly better than cost-effective which is not very different from the result with pedestrians. Finally, the random-acquisition seems to be the most efficient approach in terms of efficient budget expenditure. Figure 7 shows the simulation results of our three metrics under exponential distribution of vehicle trajectories, and participant true valuations. The result shows a similar trend in terms of the average number of active participants after 100 rounds, and the efficiency of budgeted utilization. However, there are small differences in terms of coverage where cost-efficient outperforms k -greedy until a budget of 1700, and then k -greedy surpasses cost-effective for budget values greater than 1700.

In conclusion, for these two distributions selecting a vehicle i may result in other vehicles falling within the area of influence (or coverage) of u_i . This performance can be attributed to the cost-efficient and k -greedy algorithms' ability to optimize selection based on both cost and potential coverage, making it particularly effective in scenarios with clustered distributions of participants.

4.4. Experiment 2 - Impact of Budget Allocation on Coverage, Participant Engagement, and Budget Utilization with Uncorrelated Bid Prices and Trajectory Location

The purpose of this experiment is to gain insight into how budget allocation affects coverage, the average number of active participants, and efficiency in budget utilization. All of them under the assumption of no correlation between vehicle trajectory location and density, and sample bid prices. Here, we test three scenarios. In the first scenario vehicle trajectories are distributed evenly through the city, and participant true valuations are uniformly distributed. The second scenario corresponds to vehicle trajectories uniformly distributed, but participant true valuations are normally distributed. Finally, the third scenario corresponds to vehicle trajectories normally distributed, namely people leaving and moving in well-defined districts while their true valuation is uniformly or evenly distributed.

Figure 8 shows the effect of budget allocation on coverage, average number of active participants, and efficient budget utilization. Here, we assume that vehicle trajectories and participant true valuations follow uniform distributions. Figure 9 presents the same comparison but this time when vehicular trajectories are uniformly distributed under the target area while the participant's true valuation is normally distributed. Finally, Figure 10 corresponds to the case when the vehicle trajectories are grouped following a normal distribution while participant true valuations are uniformly distributed.

In the scenario where there is no apparent correlation between vehicle trajectory locations and participant true valuations, a few key patterns emerge. First, in terms of coverage, the k -greedy, cost-effective, and cheapest-first approaches appear to perform at a similar level. Second, the cheapest-first approach outperforms the other methods in maintaining a high number of active participants, which is consistent with the findings from the previous experiment. Finally, continuing the trend observed in

prior experiments, the random acquisition approach seems to be the most efficient in terms of budget utilization.

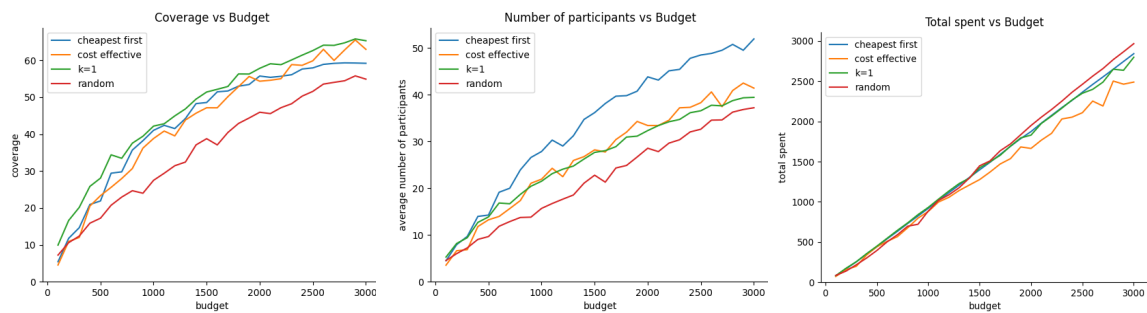


Figure 8. Budget vs coverage, number participants, and budget utilization under uniform distribution for trajectory locations and participant true valuations.

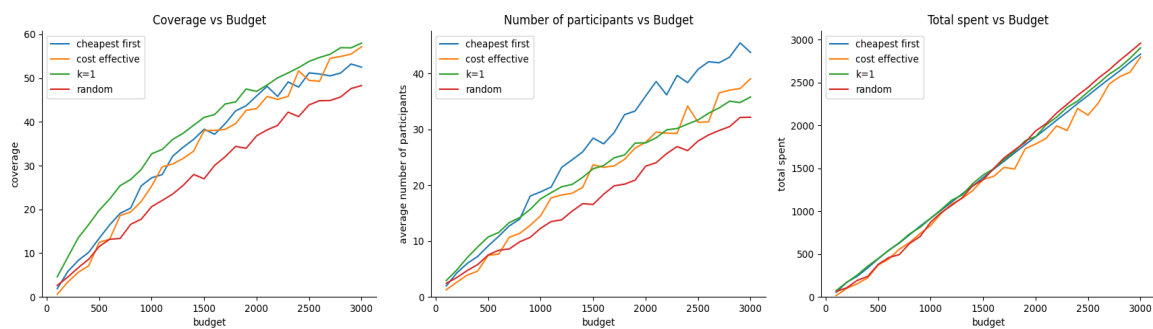


Figure 9. Budget vs coverage, number participants, and budget utilization under uniform, and normal distribution for trajectory locations and participant true valuations respectively

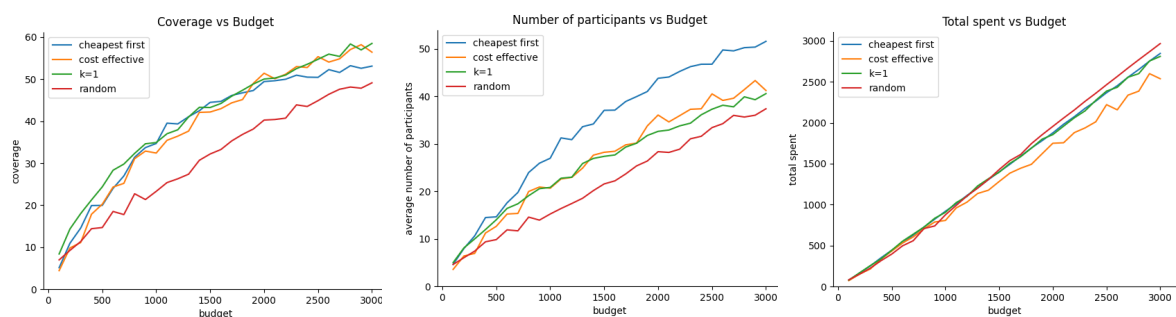


Figure 10. Budget vs coverage, number participants, and budget utilization under normal, and uniform distribution for trajectory locations and participant true valuations respectively

5. Conclusion and Future Work

This paper introduces an incentive mechanism for vehicular crowdsensing based on greedy algorithms and a recurrent-based auction. By integrating dynamic incentive mechanisms, leveraging vehicular mobility patterns, and addressing budgetary constraints, our approach enhances data collection efficiency by reducing redundancy and ensuring broad spatial coverage, a high number of active participants, and efficient budget utilization. We show that greedy algorithms, while effective in pedestrian-based crowdsensing, require significant adaptation to perform well in the vehicular context. Using simulations grounded in realistic urban environments, benchmarking four recruitment strategies under different vehicular crowd distribution scenarios, and using extensive simulations

our methodology outperformed traditional methods by optimizing participant selection based on both geographical and bid sample prices. Going forward, future work will focus on refining these techniques further, particularly in dynamic environments with fluctuating participant availability and evolving coverage demands. In addition, we would like to perform experimentation in terms of the optimal sensor radius in the vehicular context. By continually improving our participant selection strategies, we aim to unlock the full potential of edge-assisted vehicular crowdsensing and deliver enhanced data collection capabilities.

References

1. Perez, A.; Labrador, M.A.; Barbeau, S. G-Sense: A Scalable Architecture for Global Sensing and Monitoring. *IEEE Network* **2010**, *24*, 57–64.
2. Kanjo, E. NoiseSPY: A Real-Time Mobile Phone Platform for Urban Noise Monitoring and Mapping. *MONET* **2010**, *15*, 562–574.
3. Deng, L.; Cox, L.P. LiveCompare: grocery bargain hunting through participatory sensing. Proceedings of HotMobile, 2009.
4. Huang, K.L.; Kanhere, S.S.; Hu, W. Preserving privacy in participatory sensing systems. *Computer Communications* **2010**, *33*, 1266–1280.
5. Jaimes, L.G.; Chintakunta, H.; Abedin, P. SenseNow: A Time-Dependent Incentive Approach for Vehicular Crowdsensing. *IEEE Open Journal of Intelligent Transportation Systems* **2024**.
6. Lee, J.S.; Hoh, B. Dynamic pricing incentive for participatory sensing. *Pervasive and Mobile Computing* **2010**, *6*, 693–708.
7. Khuller, S.; Moss, A.; Naor, J. The Budgeted Maximum Coverage Problem. *Inf. Process. Lett.* **1999**, *70*, 39–45.
8. Wu, E.; Peng, Z. Research Progress on Incentive Mechanisms in Mobile Crowdsensing. *IEEE Internet of Things Journal* **2024**.
9. Guo, B.; Liu, Y.; Liu, S.; Yu, Z.; Zhou, X. CrowdHMT: crowd intelligence with the deep fusion of human, machine, and IoT. *IEEE Internet of Things Journal* **2022**, *9*, 24822–24842.
10. Chou, C.T.; Bulusu, N.; Kanhere, S. S.: Sensing data market. Proceedings of the 3rd IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS), 2007.
11. Danezis, G.; Lewis, S.; Anderson, R. How Much is Location Privacy Worth? Proceedings of the Fourth Workshop on the Economics of Information Security, 2005.
12. Shaw, A.D.; Horton, J.J.; Chen, D.L. Designing incentives for inexpert human raters. Proceedings of CSCW, 2011, pp. 275–284.
13. Yan, T.; Kumar, V.; Ganesan, D. CrowdSearch: exploiting crowds for accurate real-time image search on mobile phones. Proceedings of MobiSys, 2010, pp. 77–90.
14. Reddy, S.; Estrin, D.; Hansen, M.H.; Srivastava, M.B. Examining micro-payments for participatory sensing data collections. Proceedings of UbiComp, 2010, pp. 33–36.
15. Jaimes, L.G.; Vergara-Laurens, I.; Labrador, M.A. A location-based incentive mechanism for participatory sensing systems with budget constraints. 2012 IEEE International Conference on Pervasive Computing and Communications. IEEE, 2012, pp. 103–108.
16. Yang, D.; Xue, G.; Fang, X.; Tang, J. Crowdsourcing to smartphones: Incentive mechanism design for mobile phone sensing. Proceedings of the 18th annual international conference on Mobile computing and networking. ACM, 2012, pp. 173–184.
17. Chakeri, A.; Jaimes, L.G. An iterative incentive mechanism design for crowd sensing using best response dynamics. 2017 IEEE International Conference on Communications (ICC). IEEE, 2017, pp. 1–7.
18. Chakeri, A.; Jaimes, L.G. An incentive mechanism for crowdsensing markets with multiple crowdsourcers. *IEEE Internet of Things Journal* **2017**, *5*, 708–715.
19. Chintakunta, H.; Wang, X.; Jaimes, L.G. Improving sensing coverage in vehicular crowdsensing using location diversity. 2022 International Conference on Connected Vehicle and Expo (ICCVE). IEEE, 2022, pp. 1–6.
20. Chakeri, A.; Wang, X.; Jaimes, L.G. A Vehicular Crowdsensing Market for AVs. *IEEE Open Journal of Intelligent Transportation Systems* **2022**, *3*, 278–287.

21. Chakeri, A.; Wang, X.; Goss, Q.; Akbas, M.I.; Jaimes, L.G. A Platform-Based Incentive Mechanism for Autonomous Vehicle Crowdsensing. *IEEE Open Journal of Intelligent Transportation Systems* **2021**, *2*, 13–23.
22. Nistorica, V.I.; Chilipirea, C.; Dobre, C. How many people are needed for a crowdsensing campaign? 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP). IEEE, 2016, pp. 353–358.
23. Zheng, Z.; Wu, F.; Gao, X.; Zhu, H.; Tang, S.; Chen, G. A budget feasible incentive mechanism for weighted coverage maximization in mobile crowdsensing. *IEEE Transactions on Mobile Computing* **2016**, *16*, 2392–2407.
24. Zhang, M.; Yang, P.; Tian, C.; Tang, S.; Wang, B. Toward optimum crowdsensing coverage with guaranteed performance. *IEEE Sensors Journal* **2015**, *16*, 1471–1480.
25. Xu, J.; Xiang, J.; Li, Y. Incentivize maximum continuous time interval coverage under budget constraint in mobile crowd sensing. *Wireless Networks* **2017**, *23*, 1549–1562.
26. Lee, J.S.; Hoh, B. Sell your experiences: a market mechanism based incentive for participatory sensing. 2010 IEEE International Conference on Pervasive Computing and Communications (PerCom). IEEE, 2010, pp. 60–68.
27. Liu, Y.; Li, H.; Zhao, G.; Duan, J. Reverse auction based incentive mechanism for location-aware sensing in mobile crowd sensing. 2018 IEEE International Conference on Communications (ICC). IEEE, 2018, pp. 1–6.
28. Haklay, M.; Weber, P. Openstreetmap: User-generated street maps. *IEEE Pervasive computing* **2008**, *7*, 12–18.
29. Krajzewicz, D.; Erdmann, J.; Behrisch, M.; Bieker, L. Recent Development and Applications of SUMO - Simulation of Urban MObility. *International Journal On Advances in Systems and Measurements* **2012**, *5*, 128–138.
30. Goss, Q.; Akbaş, M.İ.; Jaimes, L.G.; Sanchez-Arias, R. Street Network Generation with Adjustable Complexity Using k-Means Clustering. 2019 SoutheastCon. IEEE, 2019, pp. 1–6.
31. Goss, Q.; Akbaş, M.İ.; Chakeri, A.; Jaimes, L.G. An Association-Rules Learning Approach to Unsupervised Classification of Street Networks. 2020 SoutheastCon. IEEE, 2020, pp. 1–7.
32. Uber Movement. 2017. Available online. <https://movement.uber.com/>. accessed on Jan 2020.
33. Sun, Y.; Ren, Y.; Sun, X. Uber Movement Data: A Proxy for Average One-way Commuting Times by Car. *ISPRS International Journal of Geo-Information* **2020**, *9*, 184.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.