

Article

# Probabilistic Coverage Constraint Task Assignment with Privacy Protection in Vehicular Crowdsensing

Zhe Li, Xiaolong Liu, Yang Huang and Honglong Chen \*

College of Control Science and Engineering, China University of Petroleum (East China), Qingdao 266580, China; lizhe@upc.edu.cn (Z.L.); lx18970691736@163.com (X.L.); yanghuang97@outlook.com (Y.H.)

\* Correspondence: chenhl@upc.edu.cn

**Abstract:** The increasing popularity of portable smart devices has led to the emergence of vehicular crowdsensing as a novel approach for real-time sensing and environmental data collection, garnering significant attention across various domains. Within vehicular crowdsensing, task assignment stands as a fundamental research challenge. As the number of vehicle users and perceived tasks grows, the design of efficient task assignment schemes becomes crucial. However, existing research solely focuses on task deadlines, neglecting the importance of task duration. Additionally, the majority of privacy protection mechanisms in the current task assignment process emphasize safeguarding user location information but overlook the protection of user-perceived duration. This lack of protection exposes users to potential time-aware inference attacks, enabling attackers to deduce user schedules and device information. To address these issues in opportunistic task assignment for vehicular crowdsensing, this paper presents the minimum number of participants required under the constraint of probability coverage and proposes the User-Based Task Assignment (UBTA) mechanism, which selects the smallest set of participants to minimize the payment cost while measuring the probability of accomplishing perceived tasks by user combinations. To ensure privacy protection during opportunistic task assignment, a privacy protection method based on differential privacy is introduced. This method fuzzifies the sensing duration of vehicle users and calculates the probability of vehicle users completing sensing tasks, thus avoiding the exposure of users' sensitive data while effectively assigning tasks. The efficacy of the proposed algorithm is demonstrated through theoretical analysis and a comprehensive set of simulation experiments.

**Keywords:** vehicular crowdsensing; spatio-temporal correlation; task assignment; differential privacy



**Citation:** Li, Z.; Liu, X.; Huang, Y.; Chen, H. Probabilistic Coverage Constraint Task Assignment with Privacy Protection in Vehicular Crowdsensing. *Sensors* **2023**, *23*, 7798. <https://doi.org/10.3390/s23187798>

Academic Editors: Gianluigi Ferrari, Luca Davoli, Laura Belli and Marco Martalò

Received: 7 August 2023

Revised: 5 September 2023

Accepted: 8 September 2023

Published: 11 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In recent years, the widespread adoption of portable smart IoT devices has become an integral part of people's lives [1,2]. These devices have evolved with the rapid advancement of science and technology, now equipped with a diverse array of powerful embedded sensors, including cameras, microphones, GPS, gyroscopes, accelerometers, and compasses. Together, these sensors collaboratively gather extensive information about human activities and the surrounding environment. The convergence of mobile smart terminals, featuring high-performance built-in sensors, and the rapid growth of the Internet have given rise to vehicular crowdsensing (MCS) [3,4]. This paradigm offers a novel approach for real-time sensing and environmental data collection. Vehicular crowdsensing effectively utilizes the mobile population to enable low-cost, real-time, large-scale, and fine-grained data collection. It transforms vehicle users from passive data consumers into active data providers, thereby introducing a new service mode for vast multi-source heterogeneous data awareness in the realm of the large-scale Internet of Things [5–7]. The great potential of vehicular crowdsensing has led to its widespread application across various fields, including environmental monitoring, smart cities, industrial sensors, city awareness, and social networks [8–10]. Its ability to harness the collective power of mobile users to contribute

to data collection has made it instrumental in addressing diverse challenges and needs in these domains.

Task assignment stands as a critical research aspect in the realm of vehicular crowdsensing [11–15]. The success of vehicular crowdsensing largely hinges on the design of effective task assignment schemes. As the number of vehicle users and perceived tasks in the Cluster Awareness System continues to grow rapidly, an efficient task allocation mechanism becomes crucial to enhance user participation and achieve substantial benefits. Recently, numerous researchers have devoted efforts to developing task allocation mechanisms. Some of these works [16,17] focus on maximizing perceived quality or minimizing incentive costs under various constraints, while some other works aim to reduce energy consumption [18]. Additionally, there has been significant research attention directed toward the assignment of location-dependent and time-sensitive tasks [19]. These endeavors seek to address the challenges posed by task complexity and dynamic environmental conditions, further advancing the capabilities of vehicular crowdsensing systems.

Most of the existing research [20–22] takes into account the effective time of perceived tasks, which can be categorized as either time-sensitive tasks or delay-tolerant tasks. Time-sensitive tasks [23] require immediate completion and demand high time-effectiveness, typically due to emergency situations. On the other hand, delay-tolerant tasks [24] target relatively stable perceptual activities and can be completed over an extended period. Some delay-tolerant task assignment schemes focus on recommending appropriate routes for vehicle users or scheduling tasks based on user interests and habits. However, the aforementioned studies primarily focus on task deadlines while disregarding the time required to perform each task. This oversight is particularly critical when the available time of each vehicle user is limited, as it significantly impacts the global optimization of task assignment. Thus, there is an urgent need to design an effective task assignment scheme that considers the sensing duration of vehicle users in a spatio-temporal cluster-aware system.

Privacy protection concerns in task assignment also hold significant importance [25–27]. Vehicular crowdsensing faces potential security threats due to the large volume of data streams and the absence of robust security mechanisms. The interactions between vehicle users, platforms, and third parties create a scenario wherein privacy breaches become a serious issue. In vehicular crowdsensing, users are required to upload their perceived data to the platform, which can contain sensitive information such as user identities, locations, and other private attributes. If the platform experiences a security breach or loses trustworthiness (e.g., by selling data to third parties for profit), users' privacy will be exposed to unauthorized entities.

Recently, researchers have proposed several methods to address user privacy protection in vehicular crowdsensing. To et al. [28] introduced a mechanism based on differential privacy and geo-broadcasting to safeguard users' location privacy. The approach confuses the number of users in a particular area using a trusted third-party platform, thus preventing attackers from inferring users' precise locations. Similarly, Wang et al. [29] explored the issue of publishing statistical information for location-based datasets. They employed differential privacy technology to ensure that users' personal location information remains uncompromised. However, the privacy protection schemes currently available are not applicable to vehicular crowdsensing scenarios that prioritize perception duration. Currently, there is no research focusing on privacy protection concerning perception duration in task assignment. Consequently, there is an urgent need to develop effective solutions that cater to users' privacy protection requirements while optimizing task assignment in space–time-related vehicular crowdsensing. These solutions aim to prevent users' sensitive information from being disclosed to unauthorized entities.

To sum up, this paper focuses on the space–time characteristics of vehicular crowdsensing and conducts research on task allocation in this context. Additionally, it designs a privacy protection mechanism to safeguard vehicle users' sensitive data while ensuring efficient task allocation. The research encompasses two main areas: probabilistic task

assignment with a probabilistic coverage constraint in vehicular crowdsensing, and privacy protection in the task assignment process. By addressing these aspects, this study successfully resolves the issues of ineffective task assignment caused by neglecting the spatio-temporal characteristics of users and tasks. Moreover, it tackles the problem of user-sensitive data leakage, resulting from the oversight of privacy protection concerning users' perceptual time in the task assignment process of spatio-temporal-related vehicular crowdsensing.

## 2. Related Work

### 2.1. Task Assignment without Privacy Protection

Task assignment stands as a crucial research aspect in vehicular crowdsensing. It involves the rational allocation of perceived tasks to vehicle users by servers, considering system optimization objectives and real-world constraints. And the research delves into the trade-off between cost and benefit, representing two opposing factors. Li et al. [30] considered the problem of dynamic participant selection with heterogeneous sensing tasks, that is, the sensing tasks come in real time and have different requirements for space–time coverage. Its goal is to minimize the sensing cost while maintaining a certain probability coverage level. Wang et al. [31] proposed a two-stage hybrid task assignment framework that can be used to recruit both opportunistic and participatory users, with the goal of maximizing the number of tasks completed under the constraints of the overall budget. Wang et al. [32] researched the robust task allocation problem in MCS systems, aiming to enhance the robustness of the allocation scheme and minimize the users' detour costs.

However, the majority of existing research only focuses on the deadline of the perceptual task, disregarding the working time of the vehicle user during task execution. Furthermore, few studies take into account both the space–time constraints of vehicle users and perceived tasks. In a cluster-aware system, vehicle users typically possess different route budgets and time budgets. Additionally, sensory tasks are linked to specific locations and sensing duration. As the complexity and scale of perception systems increase, the corresponding optimization problems for task assignment is NP-hard, making it challenging to find effective optimal algorithms. In such cases, suboptimal algorithms need to be considered. Hence, given the unique characteristics of task perception duration and the space–time constraints of users and tasks in spatio-temporal correlated vehicular crowdsensing, there is a critical need to thoroughly investigate the spatio-temporal-related task assignment problem. The objective is to achieve optimal task assignment while considering the intricate real-world space–time constraints. In addition, none of the above works take into account the protection of user privacy, which will affect the effectiveness of task assignment.

### 2.2. Task Assignment with Privacy Protection

In the task assignment process of vehicular crowdsensing, vehicle users often share data with the platform through wireless access points or cellular infrastructure, which involves many sensitive data (such as location information, device ID, etc.). These sensitive data may leak users' daily travel or real identity, resulting in a significant reduction in users' motivation to participate in perceived tasks. In recent years, some research work has focused on the issue of privacy protection in the process of task assignment for vehicular crowdsensing. Zhang et al. [33] assumed that, in the absence of any trusted entity, users were allowed to locally confuse their location using a method based on differential privacy to reduce the risk of location privacy disclosure. This method about trusted entity [34,35] could ensure that opponents with prior knowledge could only obtain little additional information from the ambiguous location. Wang et al. [36] proposed a task assignment framework for location privacy protection with geographic location confusion to minimize the expected travel distance of selected vehicle users under the constraints of differential privacy and distorted privacy. Xu et al. [26] proposed a privacy-preserving fine-grained task assignment scheme for the crowdsensing system and designed a novel incentive mechanism

based on the capabilities of workers. To be specific, a ciphertext-policy attribute-based encryption (CP-ABE) scheme with the characteristics of hidden policy is adopted to select workers and protect the privacy. Qian et al. [37] proposed an optimal location privacy preserving and service quality guaranteed task allocation in vehicle-based crowdsensing networks and utilized differential privacy to preserve participants' location privacy, where every participant can submit the obfuscated location to the platform instead of the real one.

The current research on privacy protection during the task assignment process for vehicular crowdsensing is still in its early stages. While some mechanisms have been proposed, they suffer from issues like high computational costs, vulnerability, or reliance on trusted third-party platforms. Moreover, the existing privacy mechanisms in task assignment predominantly focus on safeguarding user location information, overlooking the protection of user perceived time. However, it is crucial to consider the privacy of user-perceived time in space-time-related cluster awareness systems since attackers can potentially infer user schedules and device information through these data. Therefore, there is a pressing need to delve deeper into privacy protection issues in the task assignment process for space-time-related vehicular crowdsensing. This research aims to achieve optimal task assignment while ensuring the fulfillment of users' privacy protection requirements.

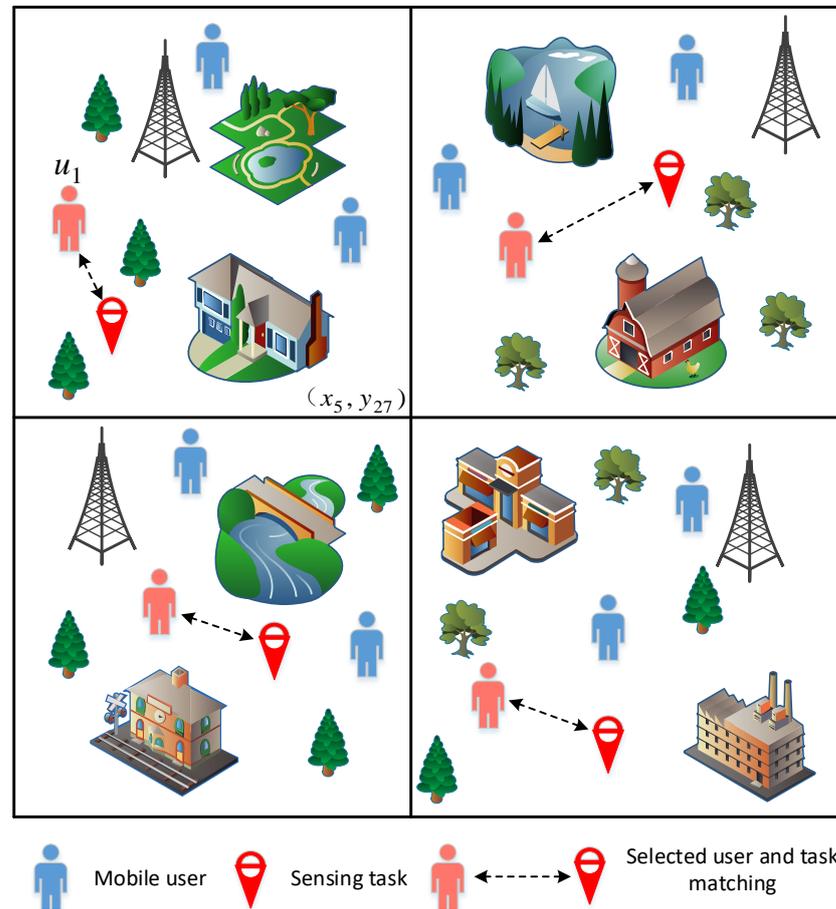
### 3. Problem Description

In this paper, we mainly focus on the spatio-temporal-related vehicular crowdsensing, in which the perception task is not only located in a specific spatial region, but also needs a certain amount of sensing time to complete. In this model, vehicular crowdsensing aims to recruit vehicle users who can dedicate a specific duration to perform tasks effectively. For instance, in application scenarios like crowd monitoring, air quality assessment, and noise level detection, the perception tasks necessitate vehicle users to collect data (such as congestion levels, air pollution indicators, and noise levels) within a certain time frame to obtain sufficient and meaningful information.

The environmental department of the municipal government has planned an air quality monitoring project to update the air quality index every hour throughout the day. However, due to the high costs associated with deploying, maintaining, and consuming energy for basic monitoring equipment, the municipal government decided to leverage vehicular crowdsensing to accomplish the task. The project aims to recruit vehicle users equipped with smartphones, who will install a specific application to perceive air quality. To execute the project effectively, the sensing area is divided into 45 regions based on signal tower coverage. Each day is further divided into 10 sensing cycles, spanning from 8:00 to 18:00, with each cycle lasting for 1 h. About 1000 smartphone users have agreed to participate in the two-week air quality sensing project. Each selected vehicle user is responsible for sensing air quality in a specific area and transmitting the collected data through the designated signal tower. The project's objective is to ensure that the tasks performed by the recruited users cover at least 90% of the 45 areas in all sensing cycles. Consequently, in each sensing cycle, no fewer than 41 air quality monitoring tasks should be completed. To minimize the total project cost while ensuring the perceived quality, the municipal government intends to select the minimum number of users from the pool of 1000 candidates. For this purpose, the municipal government obtained the two-week call records of the 1000 candidates (with time stamps of the calls and communication tower IDs, after desensitizing other information) with the assistance of telecom operators, after obtaining the consent of users.

As shown in Figure 1, for example, vehicle user  $u_1$  is assigned a sensing task associated with the space-time unit  $(x_5, y_{27})$  (that is, the sensing task is in the fifth sensing cycle and located in the 27th perception area), user  $u_1$  air quality monitoring needs to be carried out in the space-time unit  $(x_5, y_{27})$ . Air quality monitoring is a time-dependent task in vehicular crowdsensing, requiring vehicle users to spend a specific duration for obtaining effective data. To ensure the successful completion of the sensing task, users are required to upload their stay time in the sensing area during the designated sensing cycle. To protect

users' privacy, the stay time data will be subject to privacy protection measures. Upon selecting the smallest vehicle user set based on users' historical calls and mobile records, each chosen participant will receive a fixed reward and activate the application on their mobile phones. Throughout the entire project, when vehicle users reach a specific sensing area within the sensing cycle, they can utilize the application on their mobile phones to perceive and upload air quality data.



**Figure 1.** An example of task assignment.

#### 4. System Model

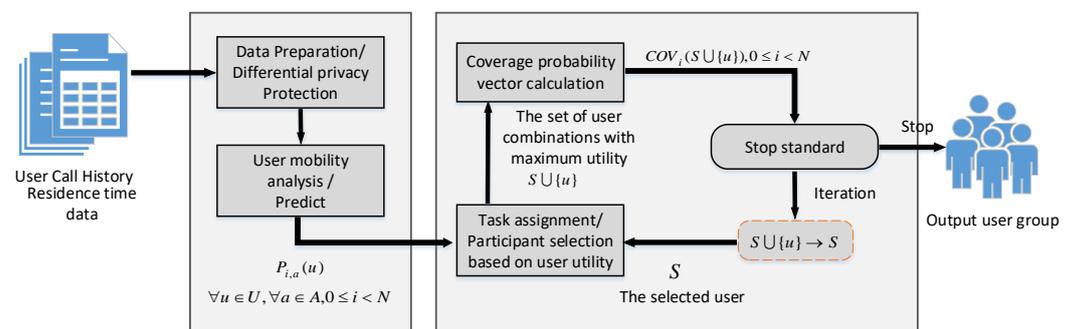
As can be seen from the above examples, the goal of this research work is to select the smallest set of participants for the perception task, and ensure that in each sensing cycle, there is a sensing area that is not lower than the predefined coverage. We propose a vehicular crowdsensing opportunistic task allocation problem. The main goal is to minimize the total incentive cost while satisfying the predefined coverage constraints. In the described model, the project can exist for a period of time (for example, two weeks), including multiple sensing cycles, such as 10 cycles per day, from 8:00 to 18:00, each cycle of 1 h. We set that, if a vehicle user completes the sensing task in the sensing period  $i$  and the sensing area  $a$ , it is said that the area  $a$  is covered in the cycle  $i$ . If a vehicle user goes to multiple sensing areas and completes the sensing tasks in these areas in the sensing cycle  $i$ , it is said that these sensing areas are covered in the sensing cycle  $i$ . Therefore, the goal of this study is to give a group of call records and confused residence time data of vehicle users, select the minimum number of users, and meet the space–time coverage constraint of the specified proportion, that is, in all sensing cycles, the proportion of the covered sensing area should be equal to or greater than the set coverage threshold.

According to the above definitions, we model the problem as follows. Given a group of vehicle users  $\mathcal{U}$  willing to participate in the sensing project, the divided sensing area set

$A$ , the call records (including call timestamp and communication tower ID), and dwell time data of all vehicle users, we define  $\mathbb{S}$  as the participants ( $\mathbb{S} \subset \mathcal{U}$ ) selected from the vehicle user set  $\mathcal{U}$ ,  $a_i(\mathbb{S})$  is the set of sensing areas covered by the participant set  $\mathbb{S}$  in the  $i$  sensing cycle, then the whole problem is to find a subset  $\mathbb{S}$  of the vehicle user set  $\mathcal{U}$ , which can be expressed as follows:

$$\begin{aligned} & \text{minimize } |\mathbb{S}|, \\ & \text{s.t. } \begin{cases} \frac{|a_i(\mathbb{S})|}{|A|} \geq R \\ 0 \leq i \leq N \end{cases} \end{aligned} \quad (1)$$

where  $R$  is the preset coverage threshold and  $N$  is the number of sensing cycles. Due to the fact that we cannot know in advance when vehicle users will appear and in which sensing area to perform sensing tasks (that is, when selecting the set of participants,  $a_i(\mathbb{S})$  is uncertain), the whole task allocation problem can be divided into two sub problems: vehicle user mobility prediction problem and prediction-based participant selection problem. The swarm intelligence sensing project uses a platform-centric task assignment method. The platform collects and stores the historical call records and residence time data of vehicle users. Before the sensing task is executed, participants are selected from all candidate vehicle users. Only the selected participants are required to perform the task and upload the perceived results in each sensing cycle. Considering the two sub problems of task assignment, the mechanism we designed is divided into two stages. In the first stage, we predict the flow of users between perceived areas based on the historical call records and residence time data of vehicle users, and in the second stage, we select participants according to the prediction results. The overall system framework is shown in Figure 2.



**Figure 2.** The task assignment framework.

#### 4.1. The First Stage—Data Preparation and User Mobility Analysis

After obtaining the call records and residence time data of all candidate vehicle users, the vehicular crowdsensing platform proceeds to calculate the mobility of each user. This mobility metric represents the probability of each user successfully completing the sensing task in a specific sensing area during a designated sensing cycle. The calculation process involves two steps as follows:

(1) Calculate the probability that each user makes at least one call in a specific sensing cycle and sensing area. First, map the historical call record of each user to  $n$  sensing cycles, so that the parameter  $\lambda_{u,I,a}$  can be calculated, that is, the average number of calls made by vehicle user  $u$  ( $u \in \mathcal{U}$ ) in sensing area  $a$  ( $a \in A$ ) in sensing cycle  $i$  ( $0 \leq i < N$ ). Get  $\lambda_{u,I,a}$ , the probability  $P'_{i,a}(u)$  that the vehicle user  $u$  makes at least one call in the sensing period  $i$  and the sensing area  $a$  can be calculated.

(2) Calculate the probability that each user's residence time in a specific sensing cycle and sensing area is equal to or greater than the sensing duration required by the task. What the vehicle user uploads is the residence time after confusion, and the sensing tasks concerned in this paper are all time-dependent tasks (that is, the task needs a certain perception time to be completed, and the user can obtain valid data). Therefore, it is

necessary to calculate the probability  $P''_{i,a}(u)$  that the residence time of the vehicle user  $u$  in the sensing cycle  $i$  and sensing area  $a$  is not less than the sensing duration required by the task. Finally, the probability  $P_{i,a}(u)$  of each user completing the task in a specific sensing cycle and area is the product of  $P'_{i,a}(u)$  and  $P''_{i,a}(u)$ .

#### 4.2. The Second Stage—Participant Selection

According to the user mobility data obtained in the first stage, we propose a user utility-based task assignment mechanism (UBTA), which iteratively selects participants by calculating the utility of the user portfolio, which is divided into three steps:

(1) Select the most effective user among all candidate vehicle users, and add the user to the recruited user set.

(2) Among the remaining non-recruited vehicle users, select the user who has the greatest utility after combining with the recruited vehicle users, and add this user to the recruited user set.

(3) Continue to select and add new vehicle users until the recruited user set can cover the sensing area of the preset threshold in each sensing cycle.

### 5. Task Assignment Algorithm Based on User Utility

In this section, users utilize the Laplacian mechanism to add noise to their own residence time, and then upload the confusing residence time instead of the real residence time to the platform to protect their privacy. In this process, only the platform and the users are aware of the privacy protection mechanism, so the attackers cannot infer from the residence time which user completed the task. At the same time, based on the fact that the platform is not interfered with by confusing data, the User6Based Task Allocation (UBTA) algorithm that we propose selects the smallest set of participants to minimize the payoff cost while measuring the probability of performing the tasks perceived by the user combinations. Moreover, the Poisson distribution, commonly used in queuing theory, is employed in this research. In real-life scenarios, numerous situations adhere to the Poisson distribution, such as passengers arriving at an airport or calls received by a customer service desk. Similarly, the Poisson distribution finds relevance in the domain of cellular networks. For this study, we assume that the call sequence adheres to a non-homogeneous Poisson process. Consequently, we model the probability of each vehicle user completing each sensing task based on the Poisson distribution. In the sensing period  $i$  ( $0 \leq i < N$ ), the probability that the vehicle user  $u$  makes  $n$  calls in the sensing area  $a$  ( $a \in A$ ) can be expressed as follows:

$$p_{i,a}(u, n) = \lambda_{u,i,a}^n * e^{-\lambda_{u,i,a}} / n!, \quad (2)$$

where  $\lambda_{u,i,a}$  represents Poisson strength, and its estimated value is the average number of calls made by vehicle user  $u$  in sensing area  $a$  in sensing period  $i$ . For example, in order to estimate  $\lambda_{u,i,a}$  of sensing period  $i$  from 08:00 to 09:00, we will calculate the average number of calls made by vehicle user  $u$  in sensing area  $a$  from 08:00 to 09:00 in historical call records. Therefore, the probability that the vehicle user  $u$  makes at least one call in the sensing area  $a$  during the sensing period  $i$  can be calculated as follows:

$$P'_{i,a}(u) = \sum_{n=1}^{+\infty} p_{i,a}(u, n) = 1 - e^{-\lambda_{u,i,a}}, \quad (3)$$

In this paper, the sensing time of vehicle users plays a crucial role in determining whether a task can be successfully completed. The sensing task is considered effectively completed only when the vehicle users reach a specific perception area and remain there for a duration equal to or exceeding the required perception time. To facilitate accurate and efficient task assignment, vehicle users can submit their stay time in the sensing period and area to the platform. However, privacy concerns arise due to potential malicious attackers who may infer daily habits, tracks, and other sensitive information from the real stay time, leading to privacy risks. To address this issue, this paper employs the Laplacian mechanism,

which locally confuses the residence time. Subsequently, vehicle users upload the fuzzy residence time instead of the actual residence time to the platform, thereby safeguarding their privacy. We use  $l_{u,i,a}$  to represent the residence time of the vehicle user  $u$  in the perception area  $a$  during the sensing period  $i$ , and  $\tilde{l}_{u,i,a}$  to represent the confusing residence time generated by adding Laplace noise to the real residence time. When  $l_{u,i,a}$  is known, it is easy to know whether the vehicle users have the ability to complete the sensing task or not. However, the existing methods cannot predict the user mobility problem because the users only upload the confused staying time, so a new solution is urgently needed.

Considering the noise added to the residence time, we transform the problem into the probability that the fuzzy residence time is greater than the perceived duration. When the sensing duration required to define the perception task is  $\delta$ , the goal of the platform is to obtain the probability that  $l_{u,i,a}$  is not less than  $\delta$ , which is expressed as  $P(l_{u,i,a} \geq \delta)$ . Assume that vehicle user  $U$  adds Laplace noise to the real residence time  $l_{u,i,a}$  by using privacy budget parameter  $\epsilon$ , and obtains the fuzzy residence time  $\tilde{l}_{u,i,a}$  and uploads it to the platform. Therefore, we can obtain:

$$l_{u,i,a} = \tilde{l}_{u,i,a} - \eta, \eta \sim \text{Laplace}(0, 1/\epsilon), \quad (4)$$

where  $\eta$  is a variable that obeys Laplace distribution, and the smaller the privacy budget parameter  $\epsilon$ , the greater the added noise and the higher the level of privacy protection. For  $P(l_{u,i,a} \geq \delta)$ , there is the following equation:

$$P(l_{u,i,a} \geq \delta) = P(\tilde{l}_{u,i,a} - \eta \geq \delta) \quad (5)$$

$$= P(\tilde{l}_{u,i,a} - \delta \geq \eta) \quad (6)$$

$$= P(\eta \leq \tilde{l}_{u,i,a} - \delta), \quad (7)$$

where  $\tilde{l}_{u,i,a} - \delta$  platform can be calculated, the above equation can be regarded as the probability problem of continuous random variables; then, we can obtain the following equation:

$$P(\eta \leq \tilde{l}_{u,i,a} - \delta) = \int_{-\infty}^{\tilde{l}_{u,i,a} - \delta} f(\eta) d\eta, \quad (8)$$

where  $f(\eta)$  represents the probability density function of variable  $\eta$ :

$$f(\eta) = \frac{\epsilon}{2} e^{-\epsilon|\eta|}. \quad (9)$$

As a result, the platform can calculate the probability  $P''_{i,a}(u)$  of vehicle user  $u$  in perception cycle  $i$  and perception region  $a$  not less than the sensing duration required by the task according to the fuzzy residence time  $\tilde{l}_{u,i,a}$  and privacy budget parameter  $\epsilon$ .

$$P''_{i,a}(u) = P(\eta \leq \tilde{l}_{u,i,a} - \delta) = \int_{-\infty}^{\tilde{l}_{u,i,a} - \delta} \frac{\epsilon}{2} e^{-\epsilon|\eta|} d\eta. \quad (10)$$

Based on the above calculation process, we can obtain the probability  $P_{i,a}(u)$  of each user completing the sensing task in specific sensing cycle  $i$  and sensing area  $a$ , that is, the probability that sensing area  $a$  is covered by vehicle user  $u$  in sensing cycle  $i$ . In addition,  $P_{i,a}(u)$  is the product of  $P'_{i,a}(u)$  and  $P''_{i,a}(u)$ .

Next, we calculate the utility of users. Given all candidate mobile user set  $U$  and selected participant set  $S$ , first combine each unselected user ( $\forall U \in U \setminus S$ ) with the selected participant to generate a combinatorial set  $S \cup \{u\}$ . Then, calculate the utility of each combinatorial set  $S \cup \{u\}$  and select the combinatorial set with the maximum utility. Finally, utilize the result as the newly selected participant set for the next iteration. In this

paper, the utility of user combinatorial set is defined as the expectation of sensing area covered by  $\mathbb{S} \cup \{u\}$  in all sensing cycles, and its calculation formula is as follows:

$$Utility(\mathbb{S} \cup \{u\}) = \sum_{0 \leq i < N} \sum_{a \in A} Q_{i,a}(\mathbb{S} \cup \{u\}), \quad (11)$$

The task assignment mechanism based on user utility calculates the utility of each combination set, selects the combination set with the maximum utility, and continues the next iteration until the stop criterion is met. After obtaining the set of combinations with maximum utility ( $\mathbb{S} \cup \{u\}$ ), the algorithm computes a covering probability vector, where each element of the vector is the probability of being covered by the combinatorial set ( $\mathbb{S} \cup \{u\}$ ) with a preset percentage ( $R$ ) of the perceived region in the corresponding perceptual cycle. The following formula gives the probability calculation formula in the  $i$ -th perception cycle.

$$COV_i(\mathbb{S} \cup \{u\}) = \sum_{A_c \in A}^{|A_c| \geq \tau} \prod_{\forall a \in A_c} Q_{i,a}(\mathbb{S} \cup \{u\}) \quad (12)$$

$$* \prod_{\forall a' \in A \setminus A_c} (1 - Q_{i,a'}(\mathbb{S} \cup \{u\})),$$

where  $\tau = |A| * R$  indicates the number of the minimum sensing areas that should be covered in each sensing cycle, and  $A_c(|A_c| \geq \tau)$  is a subset of the sensing area set  $A$ , indicating the combination of sensing areas that should be covered.

After obtaining the probability vector calculated based on the aggregation of the maximum utility user group, the algorithm identifies the minimum probability value within the vector. It then compares this minimum probability value with the given stop threshold. If the minimum probability is greater than or equal to the stop threshold, the combination set returned will be considered as the final user set selected by the algorithm. On the other hand, if the minimum probability falls below the stop threshold, the algorithm will use the combination set as the selected participant and initiate the next iteration. To summarize, the algorithm iteratively repeats the process until it finds a combination set with a minimum probability value greater than or equal to the stop threshold. This set will then be used as the final selected user set.

Given the number of users  $|\mathcal{U}|$ , in the worst case, the proposed algorithm UBTA runs  $|\mathcal{U}| * (|\mathcal{U}| + 1)/2$  iterations (i.e., all users are selected) to obtain the result, and in the best case, the algorithm runs only  $|\mathcal{U}|$  iterations (i.e., one user is selected) to obtain the result.

## 6. Performance Evaluation

### 6.1. Experiment Settings

In this section, we evaluate the proposed UBTA algorithm, first design the greedy algorithm for comparison, then introduce the experimental settings, and finally give the detailed experimental results between the UBTA algorithm and the comparison method.

For the task assignment problem in this chapter, we designed a greedy algorithm (GA), whose basic idea is to select from all unselected vehicle users in each iteration the user that covers the most perceived area in the history of call records until the set probability threshold is reached and the cycle stops. We will compare the UBTA algorithm with the GA algorithm in subsequent experiments.

This paper conducts an analysis and study of passenger flow patterns at various passenger hotspots during different time periods by mining GPS historical data from taxis. The dataset used comprises GPS historical data from 4000 taxis in Shanghai, collected over the course of one month. To account for the distinct travel patterns of residents on working days and weekends, the dataset is divided into two subsets: one for working days and the other for weekends. With the massive GPS historical data available in these subsets, this study quantifies the number of passenger-carrying events occurring at different time periods within the passenger hot spot areas. To predict passenger flow in various regions,

the Gaussian process regression algorithm is employed. Ultimately, these predictions aim to enhance the efficiency of taxi passenger transportation.

The setup for our study is as follows: we set the number of perception areas, denoted as  $|A|$ , to 45. The project involves 10 cycles, each spanning from 8:00 to 18:00, with each cycle lasting for 1 h. The sensing duration required for each task is  $\delta = 3$  min. Additionally, we assume that the noise added to the user dwell time data follows the standard Laplacian distribution, with a privacy budget parameter  $\epsilon = 1$ . In this study, we conducted a two-week experiment using a computer-generated simulation dataset. The data from the previous week served as the training dataset, used for task assignment. On the other hand, the data from the following week acted as the test dataset, simulating the spatio-temporal coverage of the selected participants. The experiment dataset comprises call records and residence time data of 1000 vehicle users in each space-time unit. It includes user ID, signal tower ID, the number of calls in each space-time unit, the real residence time of users, and the residence time after adding noise. The completion of the task by vehicle users depends on two factors: whether they make a call in a specific space-time unit and whether their stay time is longer than the sensing duration required for the task.

## 6.2. Experimental Results

Table 1 shows the comparison of the number of vehicle users who are selected by the UBTA algorithm and the GA algorithm. Evidently, the performance of the UBTA algorithm is better than the GA algorithm. When the preset coverage value is 95%, the number of vehicle users recruited by UBTA is 10.11–54.80% less than that recruited by GA. When the coverage default is 85%, the number of vehicle users recruited by UBTA is 7.94–56.89% less than the number of vehicle users recruited by GA.

**Table 1.** The number of vehicle users selected by UBTA and GA.

Sensing Program	R = 95%		R = 85%		R = 50%	
	UBTA	GA	UBTA	GA	UBTA	GA
1	113	250	58	63	20	21
2	471	644	228	321	71	72
3	456	675	222	398	69	72
4	502	620	235	298	70	72
5	449	696	216	501	68	73
6	711	791	316	458	93	94
Average	450.33	612.67	212.5	339.83	65.17	67.33

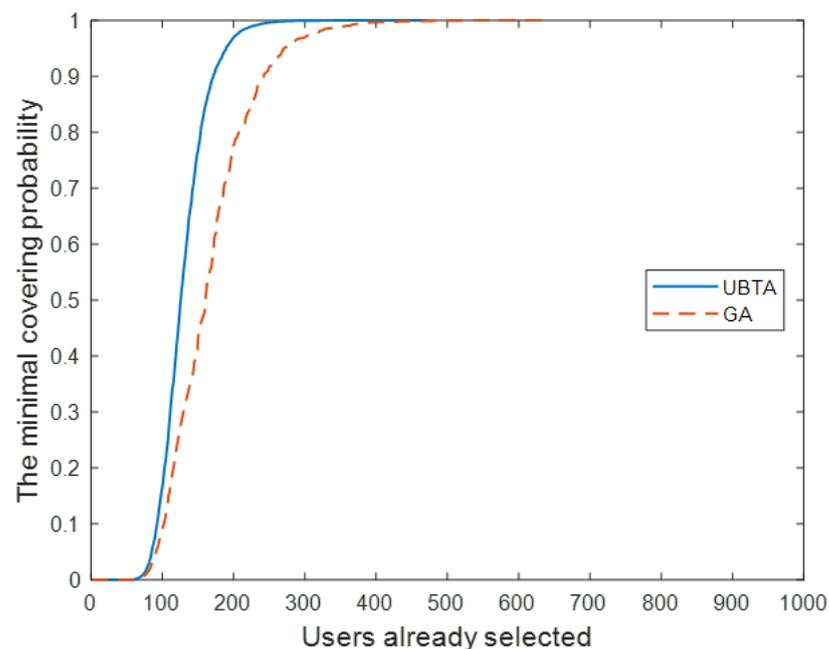
It is worth noting that, when the coverage default is 50%, there is little difference between the number of vehicle users recruited by the UBTA algorithm and the GA algorithm. Therefore, we compared the vehicle users selected by the two algorithms when the coverage preset was 95%, 85%, and 50%, respectively, and calculated the proportion of common users recruited by the two algorithms (i.e., the proportion of vehicle users in the intersection of recruitment results). As shown in Table 2, when the preset coverage value is 50%, the average proportion of common users recruited by the algorithm is 77.43%, much higher than 54.61% and 40.82% when the preset coverage value is 95% and 85%, which means that most users recruited by the GA algorithm in this case are also in the set of users recruited by the UBTA algorithm. We call the users recruited by the UBTA algorithm high-quality users (that is, those users can reach the preset coverage rate in a small number);

therefore, the GA algorithm recruits high-quality users in the majority of selected users when the preset coverage rate is 50%, which is also the reason why the number of users recruited by the GA algorithm is close to the UBTA algorithm at this time.

**Table 2.** Percentage of shared vehicle users among UBTA and GA.

Sensing Program	$R = 95\%$	$R = 85\%$	$R = 50\%$
1	44.65%	32.00%	79.85%
2	58.85%	47.04%	77.78%
3	55.41%	39.45%	75.42%
4	66.61%	48.99%	73.61%
5	52.30%	30.74%	80.23%
6	49.86%	46.72%	77.66%
Average	54.61%	40.82%	77.43%

First, Figure 3 shows how the minimum coverage probability in all sensing cycles varies with the number of mobile users selected. Clearly, the minimum coverage probability of UBTA grows faster, and its curve converges to a pre-set threshold with a smaller number of mobile users.



**Figure 3.** Participant selection process.

Next, we set the coverage default to 95% and compared the process of selecting vehicle users by UBTA and GA. Figure 4 shows how the minimum coverage probability varies with the number of selected vehicle users in all sensing cycles. Clearly, the minimum coverage probability of UBTA grows faster and its curve converges to a pre-set threshold with a smaller number of vehicle users.

Finally, we test the effectiveness of the two algorithms, that is, test data are used to test whether the participants recruited by the algorithm can reach the pre-set coverage. Figure 5 shows the average percentage of the sensing area covered by the recruited participants in

each sensing cycle. As can be seen from Figures 3 and 4, for the UBTA algorithm and the GA algorithm, the coverage of perception area in each sensing cycle has reached the pre-set threshold of 95% and 85%, thus demonstrating the effectiveness of the two algorithms. Overall, the UBTA algorithm recruited fewer vehicle users to meet pre-set coverage criteria, thereby reducing perceived costs.

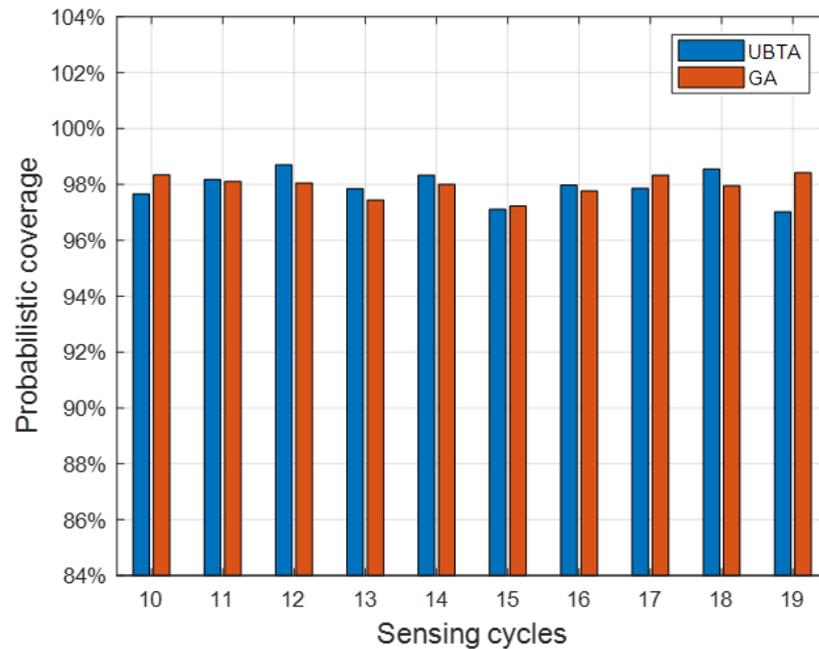


Figure 4. Average coverage of sensing areas:  $R = 95\%$ .

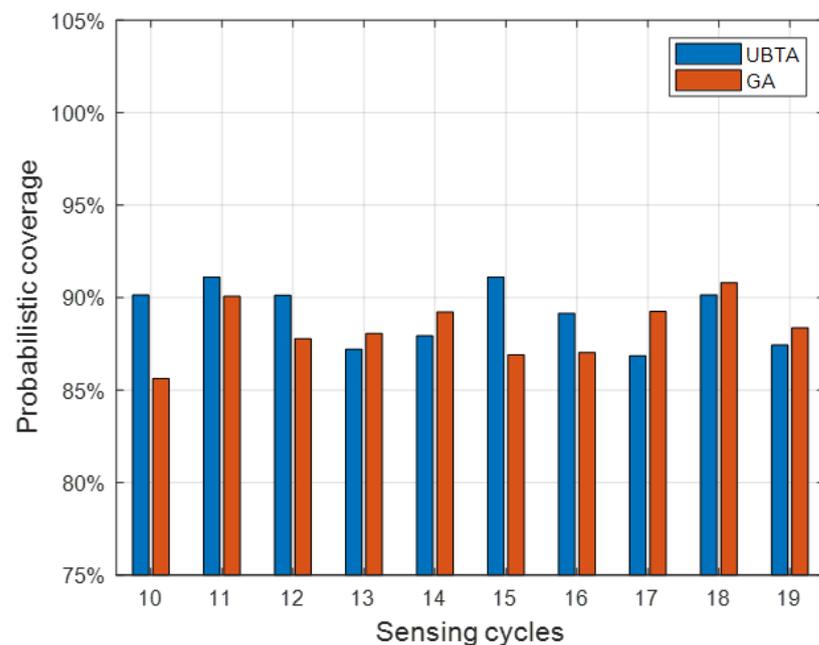


Figure 5. Average coverage of sensing areas:  $R = 85\%$

## 7. Conclusions

In this paper, we address the task allocation challenge in vehicular crowdsensing through an opportunistic approach based on the real-world scenario. We propose a user utility-based task allocation algorithm that involves several key steps. Firstly, we predict the mobility of users to understand their movement patterns. Next, we calculate the utility

of various user combinations to assess their effectiveness in completing tasks. Additionally, we measure the probability of sensing task coverage for each user combination. Based on these calculations, we determine the set of participants with the minimum number of users required for optimal task allocation. To protect user privacy, we employ differential privacy technology. Users upload their residence time data with Laplace noise, ensuring that their sensitive information remains secure. The platform then analyzes user mobility by solving the probability problem of continuous random variables, utilizing the fuzzy residence time. The proposed algorithm is further analyzed theoretically, including time complexity and efforts to reduce the computational complexity of the coverage probability vector. Finally, we conduct experimental evaluations to demonstrate the effectiveness of the proposed algorithm.

**Author Contributions:** Conceptualization, Z.L. and Y.H.; methodology, Y.H.; software, X.L.; validation, Z.L. and H.C.; formal analysis, X.L.; investigation, H.C.; resources, Z.L.; data curation, Y.H.; writing—original draft preparation, Y.H.; writing—review and editing, X.L.; visualization, Z.L.; supervision, H.C.; project administration, H.C.; funding acquisition, Z.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded in part by the Major Scientific and Technological Projects of CNPC under Grant ZD2019-183-003.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article. No public involvement in any aspect of this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Chi, X.; Chen, H.; Li, G.; Ni, Z.; Jiang, N.; Xia, F. EDSP-Edge: Efficient Dynamic Edge Service Entity Placement for Mobile Virtual Reality Systems. *IEEE Trans. Wirel. Commun.* **2023**, *Early Access*. [[CrossRef](#)]
2. Lin, K.; Chen, H.; Yan, N.; Ni, Z.; Wang, Z.; Yu, J. Double Polling-Based Tag Information Collection for Sensor-Augmented RFID Systems. *IEEE Trans. Mob. Comput.* **2023**, *Early Access*. [[CrossRef](#)]
3. Ganti, R.K.; Ye, F.; Lei, H. Mobile crowdsensing: Current state and future challenges. *IEEE Commun. Mag.* **2011**, *49*, 32–39. [[CrossRef](#)]
4. Guo, B.; Wang, Z.; Yu, Z.; Wang, Y.; Yen, N.Y.; Huang, R.; Zhou, X. Mobile Crowd Sensing and Computing: The Review of An Emerging Human-powered Sensing Paradigm. *ACM Comput. Surv. (CSUR)* **2015**, *48*, 7. [[CrossRef](#)]
5. Zhang, C.; Zhu, L.; Xu, C.; Sharif, K. PRVB: Achieving Privacy-Preserving and Reliable Vehicular Crowdsensing via Blockchain Oracle. *IEEE Trans. Veh. Technol.* **2021**, *70*, 831–843. [[CrossRef](#)]
6. Liu, L.; Wang, L.; Lu, Z.; Liu, Y.; Jing, W.; Wen, X. Cost-and-Quality Aware Data Collection for Edge-Assisted Vehicular Crowdsensing. *IEEE Trans. Veh. Technol.* **2022**, *71*, 5371–5386. [[CrossRef](#)]
7. Xue, H.; Chen, H.; Dai, Q.; Lin, K.; Li, J.; Li, Z. CSCT: Charging Scheduling for Maximizing Coverage of Targets in WRSNs. *IEEE Trans. Comput. Soc. Syst.* **2022**, *Early Access*. [[CrossRef](#)]
8. Fascista, A.; Coluccia, A.; Ravazzi, C. A Unified Bayesian Framework for Joint Estimation and Anomaly Detection in Environmental Sensor Networks. *IEEE Access* **2022**, *11*, 227–248. [[CrossRef](#)]
9. Diviaco, P.; Iurcev, M.; Carbajales, R.J.; Potleca, N.; Viola, A.; Burca, M.; Busato, A. Monitoring Air Quality in Urban Areas Using A Vehicle Sensor Network (VSN) Crowdsensing Paradigm. *Remote Sens.* **2022**, *14*, 5576. [[CrossRef](#)]
10. Zhang, D.; Wang, L.; Xiong, H.; Guo, B. 4W1H in Mobile Crowd Sensing. *IEEE Commun. Mag.* **2014**, *52*, 42–48. [[CrossRef](#)]
11. Yin, B.; Li, J.; Wei, X. Rational Task Assignment and Path Planning Based on Location and Task Characteristics in Mobile Crowdsensing. *IEEE Trans. Comput. Soc. Syst.* **2022**, *9*, 781–793. [[CrossRef](#)]
12. Yucel, F.; Bulut, E. Online Stable Task Assignment in Opportunistic Mobile Crowdsensing with Uncertain Trajectories. *IEEE Internet Things J.* **2022**, *9*, 9086–9101. [[CrossRef](#)]
13. Song, S.; Liu, Z.; Li, Z.; Xing, T.; Fang, D. Coverage-Oriented Task Assignment for Mobile Crowdsensing. *IEEE Internet Things J.* **2020**, *7*, 7407–7418. [[CrossRef](#)]
14. Huang, Y.; Chen, H.; Ma, G.; Lin, K.; Ni, Z.; Yan, N.; Wang, Z. OPAT: Optimized Allocation of Time-Dependent Tasks for Mobile Crowdsensing. *IEEE Trans. Ind. Inform.* **2022**, *18*, 2476–2485. [[CrossRef](#)]
15. Ma, G.; Chen, H.; Huang, Y.; Wei, W.; Liu, X.; Wang, Z. Utility-Based Heterogeneous User Recruitment of Multitask in Mobile Crowdsensing. *IEEE Internet Things J.* **2023**, *10*, 9796–9808. [[CrossRef](#)]

16. Liu, Y.; Guo, B.; Wang, Y.; Wu, W.; Yu, Z.; Zhang, D. TaskMe: Multitask Allocation in Mobile Crowd Sensing. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Heidelberg, Germany, 12–16 September 2016; pp. 403–414.
17. Cheng, P.; Lian, X.; Chen, L.; Shahabi, C. Prediction-Based Task Assignment in Spatial Crowdsourcing. In Proceedings of the IEEE 33rd International Conference on Data Engineering (ICDE), San Diego, CA, USA, 19–22 April 2017; pp. 997–1008.
18. Liu, C.; Zhang, B.; Su, X.; Ma, J.; Wang, W.; Leung, K.K. Energy-Aware Participant Selection for Smartphone-Enabled Mobile Crowd Sensing. *IEEE Syst. J.* **2017**, *11*, 1435–1446. [[CrossRef](#)]
19. He, S.; Shin, D.; Zhang, J.; Chen, J. Toward Optimal Allocation of Location Dependent Tasks in Crowdsensing. In Proceedings of IEEE INFOCOM 2014–IEEE Conference on Computer Communications, Toronto, ON, Canada, 27 April–2 May 2014; pp. 745–753.
20. Cheung, M.; Hou, F.; Huang, J.; Southwell, R. Distributed Time-Sensitive Task Selection in Mobile Crowdsensing. *IEEE Trans. Mob. Comput.* **2021**, *20*, 2172–2185. [[CrossRef](#)]
21. Wang, H.; Zhao, D.; Ma, H.; Ding, L. MB-GVNS: Memetic Based Bidirectional General Variable Neighborhood Search for Time-Sensitive Task Allocation in Mobile Crowd Sensing. *IEEE Trans. Veh. Technol.* **2020**, *69*, 2219–2229. [[CrossRef](#)]
22. Jiang, L.; Niu, X.; Xu, J.; Wang, Y.; Wu, Y.; Xu, L. Time-Sensitive and Sybil-Proof Incentive Mechanisms for Mobile Crowdsensing via Social Network. *IEEE Access* **2018**, *6*, 48156–48168. [[CrossRef](#)]
23. Zhan, Y.; Xia, Y.; Zhang, J.; Wang, Y. Incentive Mechanism Design in Mobile Opportunistic Data Collection with Time Sensitivity. *IEEE Internet Things J.* **2018**, *5*, 246–256. [[CrossRef](#)]
24. Xiao, M.; Wu, J.; Huang, L.; Cheng, R.; Wang, Y. Online Task Assignment for Crowdsensing in Predictable Mobile Social Networks. *IEEE Trans. Mob. Comput.* **2017**, *16*, 2306–2320. [[CrossRef](#)]
25. Cheng, Y.; Ma, J.; Liu, Z.; Wu, Y.; Wei, K.; Dong, C. A Lightweight Privacy Preservation Scheme with Efficient Reputation Management for Mobile Crowdsensing in Vehicular Networks. *IEEE Trans. Dependable Secur. Comput.* **2023**, *20*, 1771–1788. [[CrossRef](#)]
26. Xu, X.; Yang, Z.; Xian, Y. ATM: Attribute-Based Privacy-Preserving Task Assignment and Incentive Mechanism for Crowdsensing. *IEEE Access* **2021**, *9*, 60923–60933. [[CrossRef](#)]
27. Zhao, B.; Tang, S.; Liu, X.; Zhang, X.; Chen, W. iTAM: Bilateral Privacy-Preserving Task Assignment for Mobile Crowdsensing. *IEEE Trans. Mob. Comput.* **2021**, *20*, 3351–3366. [[CrossRef](#)]
28. To, H.; Ghinita, G.; Shahabi, C. A Framework for Protecting Worker Location Privacy in Spatial Crowdsourcing. *Proc. VLDB Endowment.* **2014**, *7*, 919–930. [[CrossRef](#)]
29. Wang, Q.; Zhang, Y.; Lu, X.; Wang, Z.; Qin, Z.; Ren, K. Real-Time and Spatio-Temporal Crowd-Sourced Social Network Data Publishing with Differential Privacy. *IEEE Trans. Dependable Secur. Comput.* **2018**, *15*, 591–606. [[CrossRef](#)]
30. Li, H.; Li, T.; Wang, W.; Wang, Y. Dynamic Participant Selection for Large-Scale Mobile Crowd Sensing. *IEEE Trans. Mob. Comput.* **2017**, *18*, 2842–2855. [[CrossRef](#)]
31. Wang, J.; Wang, F.; Wang, Y.; Wang, L.; Qiu, Z.; Zhang, D.; Guo, B.; Lv, Q. HyTasker: Hybrid Task Allocation in Mobile Crowd Sensing. *IEEE Trans. Mob. Comput.* **2020**, *19*, 598–611. [[CrossRef](#)]
32. Wang, L.; Yu, Z.; Wu, K.; Yang, D.; Wang, E.; Wang, T.; Mei, Y.; Guo, B. Towards Robust Task Assignment in Mobile Crowdsensing Systems. *IEEE Trans. Mob. Comput.* **2023**, *22*, 4297–4313. [[CrossRef](#)]
33. Zhang, X.; Ding, J.; Li, X.; Yang, T.; Wang, J.; Pan, M. Mobile Crowdsensing Task Allocation optimization with Differentially Private Location Privacy. In Proceedings of the ICC 2020–2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6.
34. Antonio, M.; Fernández-Gago, C.; López-Villa, R. A Test Environment for Wireless Hacking in Domestic IoT Scenarios. *Mob. Netw. Appl.* **2022**, 1–10. [[CrossRef](#)]
35. Chen, D.; Chang, G.; Sun, D.; Li, J.; Jia, J.; Wang, X. TRM-IoT: A Trust Management Model Based on Fuzzy Reputation for Internet of Things. *Comput. Sci. Inf. Syst.* **2011**, *8*, 1207–1228. [[CrossRef](#)]
36. Wang, L.; Yang, D.; Han, X.; Zhang, D.; Ma, X. Mobile Crowdsourcing Task Allocation with Differential-and-Distortion Geo-Obfuscation. *IEEE Trans. Dependable Secur. Comput.* **2021**, *18*, 967–981. [[CrossRef](#)]
37. Qian, Y.; Ma, Y.; Chen, J.; Wu, D.; Tian, D.; Hwang, K. Optimal Location Privacy Preserving and Service Quality Guaranteed Task Allocation in Vehicle-Based Crowdsensing Networks. *IEEE Trans. Intell. Transp. Syst.* **2021**, *22*, 4367–4375. [[CrossRef](#)]

**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.