A Comparative Approach to Resurrecting the Market of MOD Vehicular Crowdsensing

Chaocan Xiang[*], Yaoyu Li[*], Yanlin Zhou[*], Suining He[†], Yuben Qu[‡], Zhenhua Li[¶], Liangyi Gong[§] and Chao Chen[*]

[*]College of Computer Science, Chongqing University, China
[†]Department of Computer Science and Engineering, The University of Connecticut, U.S.A.
[‡]College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, China
[¶]School of Software, Tsinghua University, China
[§]Computer Network Information Center, Chinese Academy of Sciences, China

Abstract—With the popularity of Mobility-on-Demand (MOD) vehicles, a new market called MOD-Vehicular-Crowdsensing (MOVE-CS) was introduced for drivers to earn more by collecting road data. Unfortunately, MOVE-CS failed after two years of operation. To identify the root cause, we survey 581 drivers and reveal its simple operation model based on blindly competitive rewards. This model brings most drivers few yields, resulting in their withdrawals. In contrast, a similar market termed MOD-Human-Crowdsensing (MOMAN-CS) remains successful thanks to a complex model based on exclusively customized rewards. Hence, we wonder whether MOVE-CS can be resurrected by learning from MOMAN-CS. Despite considerable similarity, we can hardly apply the operation model of MOMAN-CS to MOVE-CS, since drivers are also concerned with passenger missions that dominate their earnings. To this end, we analyze a large-scale dataset of 12,493 MOD vehicles, finding that drivers have explicit preference for short-term, immediate gains as well as implicit rationality in pursuit of long-term, stable profits. Therefore, we design a novel operation model for MOVE-CS, at the heart of which lies a spatial-temporal differentiation-aware task recommendation scheme empowered by submodular optimization. Applied to the dataset, our design would essentially benefit both the drivers and platform, thus possessing the potential to resurrect MOVE-CS.

I. Introduction

Recent years have witnessed the prosperity of the Mobility-on-Demand (MOD) vehicle market, led by Uber, Lyft, DiDi, and so forth [1]. As of December 2020, Uber and Lyft had each incorporated over one million drivers in the U.S. [2], and the global market size is predicted to reach $228 billion by 2022 [3]. Meanwhile, however, we notice that many MOD drivers have suffered from shrinking earnings year by year from 2013 to 2020 [4], probably owing to more competition among them; the circumstance is further aggravated in the past nearly two years due to the recent COVID-19 pandemic [5]. As a result, a new market termed MOD-Vehicular-Crowdsensing (MOVE-CS) was introduced in 2017, pioneered by the Payver platform [6]. Payver pays the drivers to collect road data on the move, mainly according to the road length and the specific segments, typically at $0.01-0.05 per mile.

After receiving the collected road data from the drivers, Payver usually sold them to demanding companies such as digital map construction corporations (*e.g.,* Google Maps [7] and lvl5 [8]). Thereby, the platform and the drivers seemed to have achieved a win-win situation. After only three months of operation, Payver had taken in nearly 2000 Uber and Lyft drivers, collecting the data of more than 500K-mile roads and boosting their earnings by 5% to 15% [9]. Unfortunately, after two years of operation, there remained few participant drivers, and thus Payver had to bankrupt itself in April 2019 [10].

To figure out the root cause of the aforementioned adversity, we surveyed 581 MOD drivers (clarified in Sec. II-A) via Amazon Mechanical Turk, a well-known crowdsourcing platform [11]. They comprise 41.2% of women and 58.8% of men, aging from 20 to 60; 43.6%, 77.3%, and 90.2% of them drive at least once every day, week, and month, respectively. The survey results unveil that MOVE-CS drivers' withdrawals are highly related to the simple operation model adopted by Payver based on blindly competitive rewards. Specifically, because each driver collects data for certain road segments without knowledge of others, they often end up with low-value collected data for repetitive road segments. Hence, this model leads most drivers into few or even negative yields (*e.g.,* when the sensing task is performed while the vehicle is vacant), triggering their opt-outs from the MOVE-CS market.

Contrary to MOVE-CS, we spot that a similar market named MOD-Human-Crowdsensing (MOMAN-CS), led by Gigwalk [12], preserves its success since 2010. It hires people to collect merchandise data (*e.g.,* the location, price, and sales) for specific vendors, and has incorporated 1.7 million participants by 2021 [12]. Behind the success of Gigwalk, we find there is a complex operation model with exclusively customized rewards. Specifically, for a task, Gigwalk posts an initial reward and only allows one person to accept it; if no one takes it on for a long time, the reward will be increased. Now that the operation model of MOMAN-CS is capable of incentivizing humans effectively, and humans steer the vehicles, we wonder if this model can be applied to vehicle incentivization in hopes of resurrecting the MOVE-CS market.

Since there is considerable similarity between the two markets, most mechanisms in MOMAN-CS can be borrowed

to improve MOVE-CS. For example, road data collection in MOVE-CS can be divided into exclusive sensing tasks for drivers to choose. Also, those unpopular road segments can be assigned with more rewards. Nevertheless, we find a key obstacle during the applying process, *i.e.,* the drivers are also concerned with passenger missions which typically dominate their earnings. Therefore, the task selection strategy in MOVE-CS should differ significantly from that in MOMAN-CS.

To address this obstacle, we analyze a large-scale dataset[1] of 12,493 MOD vehicles' service records for one month (03/01/2017–03/31/2017) in a 4,400 km^2 metropolitan area with 10.3 million residents, including each passenger mission's pick-up/drop-off locations, time-variant occupied/vacant statuses, and fine-grained vehicle trajectories (explicated in Sec. II-B). The results reveal that

(1) On a daily basis, we observe that the majority (88.2%) of drivers move from low-yield zones to high-yield zones for picking up passengers, showcasing their *explicit preference* for *short-term, immediate gains*.

(2) On a monthly basis, however, we note that a considerable portion (30%) of drivers still drive from high-yield zones to low-yield zones for picking up passengers with a high occurrence of 21.1%. Surprisingly perhaps, we find their hourly earnings to be 17.5% more than the average level ($126.6 monthly raise), uncovering their *implicit rationality* in pursuit of *long-term, stable profits*.

Motivated by these findings, we present Long-Short-Term Profit-combined Task Recommendation (*LSTRec*), a novel operation model for MOVE-CS, whose primary goal is to satisfy both drivers' explicit preference for short-term gains and their implicit need of long-term profits. To this end, *LSTRec* actively recommends tasks to the drivers with balanced intelligence, in order to not only attract more participants, but also bring sufficient profits to regular drivers. Meanwhile, *LSTRec* should also take the platform's profit into account. In practice, in some cases the interests of drivers and the platform are in correspondence, *e.g.,* when the platform recommends a task enabling a driver to go from a low-yield zone to a high-yield zone, this driver is very likely to accept it even with a relatively low reward. In other cases, their interests might be in conflict, *e.g.,* for an unpopular road segment whose information is however valuable to the platform, the platform has to offer a relatively high reward to motivate drivers.

To address the challenges mentioned above, we design a spatial-temporal differentiation-aware task recommendation scheme empowered by submodular optimization. In specific, based on the historical MOD vehicle dataset, we first construct a two-dimensional pick-up profit heatmap. Then, we predict the evolution of the profit heatmap by exploiting Recurrent Neural Networks (RNN). With the above information, we formulate a task recommendation problem considering both new and regular drivers' concerns, as well as the platform's profit. Unfortunately, it is NP-hard to find the optimal solution

---

[1]We collected all the data (excluding user-sensitive information) under a well-organized IRB with informed consent of involved drivers and passengers.
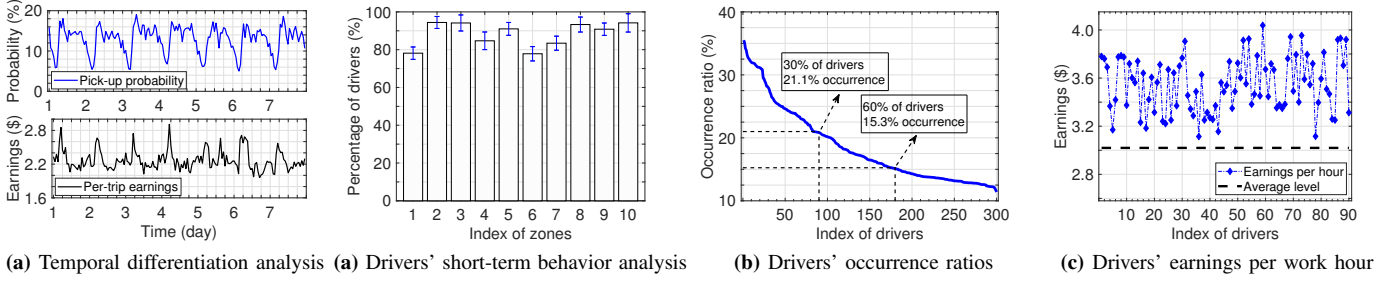
to the problem (the computation cost increases exponentially with the number of drivers). To resolve this, following the methodology of submodular optimization, we devise a near-optimal algorithm, leveraging greedy local-search to achieve an acceptable approximation ratio $(1-e^{-2})/2$ with polynomial time complexity.

Using the aforementioned large-scale MOD vehicle dataset, we emulate the operation process of the original MOVE-CS model and *LSTRec* respectively on a common commodity server. Results show that with *LSTRec*, 87.3% of the recommended tasks cater for drivers' explicit preference of short-term gains; meanwhile, all the drivers are expected to make positive earnings and 50% of them make $3.2\times$ more earnings (than with the original MOVE-CS model), serving their implicit need of long-term profits. Besides benefiting the drivers, *LSTRec* brings 34.3% more profit to the platform. Thus, we feel that our proposed novel model has the potential to resurrect the MOVE-CS market.

II. MOTIVATION

In this section, we investigate the reasons behind the downturn of the MOVE-CS market via user studies and large-scale data analysis, and explore potential methods to resurrect it referring to the thriving MOMAN-CS market.

A. Crowdsourcing-based User Studies

Methodology. To investigate why the above two markets faced completely different fates, we conduct user studies [13] with 581 MOD drivers via Amazon Mechanical Turk. The respondent pool is restricted to qualified drivers. The participants comprise 41.2% of women and 58.8% of men, including North Americans (34.4%), Europeans (12.7%), Asians (38.6%), and others (14.3%, such as Australians and Africans), aging from 20 to 60; 43.6%, 77.3%, and 90.2% of them drive at least once every day, every week, and every month, respectively.

We adopt the USE questionnaire methodology [14] and use a 5-point Likert scale (ranging from Strongly Disagree to Strongly Agree) to assess the participants' perceptions. Results are classified into two groups, *i.e.,* 4 and 5 for agreement; 1, 2, and 3 for disagreement. The queries are designed to get to the bottom of two key questions, *i.e.,* why does the MOVE-CS model fail to encourage MOD drivers? why is the MOMAN-CS model effective to incentivize users?

Results. We first investigate whether the MOD drivers are willing to perform sensing tasks. Survey results indicate *92.6% of participants are willing to perform sensing tasks on the move*. Digging deeper, it seems related to the fact that the majority (63.8%) of drivers take on sensing tasks with expectations of extra earnings, which conforms to common sense.

Further survey on the two models shows that the blind competition model adopted by Payver is not welcomed by 63.3% of participants; 94.3% regard the repeated data collection — which may cause a lower or even negative profit — as a major drawback. Therefore, it is reasonable to deduce that the blind competition model introduces uncertainty in drivers' profits, which severely impacts their enthusiasm for task participation.

(a) Temporal differentiation analysis   **(a)** Drivers' short-term behavior analysis   **(b)** Drivers' occurrence ratios   **(c)** Drivers' earnings per work hour

Fig. 1: Variations of pick-up probability (top) and per-trip earnings (bottom) in different time periods for the same zone.

Fig. 2: In-depth analysis of drivers' behavior patterns via 2D slicing. (a) Drivers' short-term preference on a daily basis, *i.e.,* the percentage of drivers moving from low-yield zone to high-yield zone. (b-c) Drivers' long-term pursuit on an individual basis, including their occurrence ratios of driving from high-yield zone to low-yield zone, and their average earnings in a month.

Contrarily, 95.2% of participants prefer MOMAN-CS, because it not only has transparent rewards (70.2% agreement), but also gives them more choices of tasks (81.3% agreement). To sum up, *MOVE-CS' downfall was likely a result of the employed blind competition model failing to offer drivers stable profits*, while *MOMAN-CS motivates participants successfully with its exclusive task selection and transparent reward*.

B. Large-scale Dataset Collection and Analysis

Dataset collection. Cooperating with an MOD company, we acquire a large-scale MOD driver dataset; all the user-sensitive information is removed according to the local IRB protocols. This dataset comprises 92 GB service records of 12,493 MOD vehicles for one month (03/01/2017–03/31/2017) in a 4,400 km^2 metropolitan area with 10.3 million residents. Each record contains an anonymized vehicle ID, the trajectory time series with an interval of 15 seconds, and an occupied/vacant state indicator. Moreover, with the trajectory series, we calculate the pick-up profits according to the existing policies on MOD vehicle fares [15].

Pick-up profit analysis. The pick-up profit denotes the average profit of MOD drivers from picking up passengers in a zone during a time period (*e.g.,* 1 hour). It is highly dependent on the pick-up probability and the per-trip earnings in this zone. Hence, in the following, we randomly select an area (about 256 km^2) of the city, and divide it into 14×18 uniform zones. Then, we analyze the spatial-temporal differences of pick-up profits in each zone and time period in terms of the pick-up probability and the per-trip earnings.

We initially analyze the temporary diversity of pick-up probability and per-trip earnings during different time periods in a randomly selected zone. As demonstrated in Fig. 1a, both the pick-up probability (top) and the per-trip earnings (bottom) vary significantly with time. In particular, the pick-up probability and per-trip earnings is distributed between 5.0% and 19.1%, $1.96 and $2.91, respectively. Moreover, both pick-up probability and per-trip earnings roughly follow periodic patterns, *e.g.,* Fig. 1a illustrates that the pick-up probability from midnight to 6 a.m. is always smaller than that of the other periods in a day, as most citizens are in sleep. Next, we analyze their spatial diversities in different zones during a time period (*e.g.,* 6 p.m. to 7 p.m.). As demonstrated in Figs. 3a and 3b, similarly, both the pick-up probability and

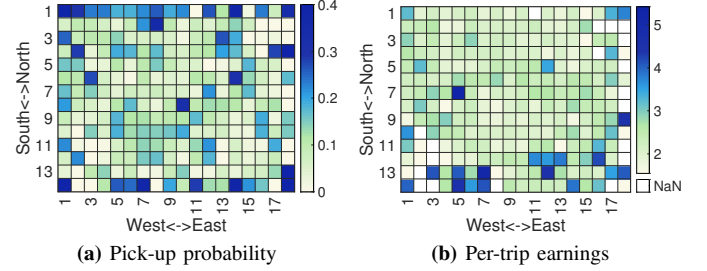**(a)** Pick-up probability   **(b)** Per-trip earnings

Fig. 3: Pick-up profits in different zones during the same time period, in terms of pick-up probability and per-trip earnings ($).

the per-trip earnings are found to vary with zones in the same period. In particular, the pick-up probability and per-trip earnings widely fluctuate between 0 and 39%, $1.55 and $5.41, respectively. In summary, *the pick-up profits of MOD drivers have huge spatial-temporal differences in different zones and time periods*.

MOD drivers' behavior analysis. MOD drivers have behavior patterns with a common goal (making money) but diverse individual preferences (such as how to make more money) based on driving experience. To fully grasp their behavior patterns, we conduct a comprehensive analysis of the large-scale dataset by slicing it — in the aspects of both short-term and long-term profits — on a daily basis and an individual basis, respectively.

First, we slice the dataset on a daily basis to study drivers' short-term preference in each day. Then, we arbitrarily select ten low-yield zones. Targeting each zone, we calculate the corresponding percentage of drivers, who move directly (from this low-yield zone) into a high-yield zone for picking up passengers. As shown in Fig. 2a, the average percentage of one month in all selected zones is 88.2%. It indicates *most drivers in low-yield zones have a tendency of moving out (towards higher-yield zones), which is compelling evidence of drivers' explicit preference for immediate gains*.

Second, to understand drivers' long-term pursuit, we slice the dataset on an individual basis, each slice with the entire one-month driving records of a driver. Then, we randomly select 300 drivers. Focusing on each driver's behavior pattern, we calculate the occurrence of her/his moving, in the entire month, from a high-yield zone into a low-yield zone for passenger pick-ups. After ranking drivers, as shown in Fig. 2b,

we find that 30% of them (90 drivers) have more than 21.1% occurrence, which appears to be weird at first glance. By comparing the hourly pick-up profits of these 90 drivers against the average level of all 12,493 drivers, surprisingly, we find that these drivers make 17.5% more pick-up profits per work hour than the average level (about $126.6 monthly raise considering the 8-hour work day), as shown in Fig. 2c. After a thorough analysis of these findings, the mystery finally uncovers its veil: *regular drivers possess the ability of dynamic profit prediction to some degree, and rationally choose where to go based on this knowledge in pursuit of long-term, stable profits, rather than blindly seek the immediate gains.*

III. *LSTRec* DESIGN FOR MOVE-CS

Motivated by the findings in Sec. II, we design a new Long-Short-Term Profit-combined Task Recommendation model (*LSTRec*) for resurrecting the MOVE-CS market, and advance the crucial research problem.

A. Model Design

Logic behind the design. *LSTRec* leverages the *active task recommendation of the platform* to simultaneously satisfy drivers' explicit and implicit needs for short-term and long-term profits. The logic behind this model design is as follows:

(1) The dataset analysis in Sec. II-B shows that MOD drivers have an explicit preference for short-term, immediate gains as well as implicit rationality in pursuit of long-term, stable profits. Therefore, the model design should respond to drivers' demands to encourage their participation.

(2) However, the short-term and long-term profits can be only predicted with global knowledge of the pick-up profits at any time and place, which is barely possible even for regular drivers, let alone those new registrants. Therefore, it is unrealistic to let drivers actively select from all the tasks in a short response time, while gaining acceptable profits.

(3) Hence, instead of task selection by drivers like the MOMAN-CS model, we deploy the task recommendation scheme, *i.e.,* the professional platform with enough spatial-temporal knowledge, predicts the pick-up profits on drivers' behalf, and actively recommends tasks comprehensively considering their short-term and long-term profits.

To sum up, *the LSTRec model saves drivers from the extremely complicated computation of profit prediction, reducing the response time of each driver, while they are still left enough wiggle room for options. Simultaneously, the platform can also pursue its own interest in this process.*

LSTRec model design. To begin with, our *LSTRec* model consists of three major steps as follows:

(1) *Task publishing*: The new MOVE-CS platform discretizes the required road data collection into S exclusive sensing tasks, according to the topology and length of the roads as well as the specified applications. We denote the set of these published tasks by $\mathcal{S}$, *i.e.,* $\mathcal{S} = \{1, \ldots, S\}$. The zone of each task j ($j \in \mathcal{S}$) and the platform profit from it are represented by z_j and u_j, respectively.

(2) *Task requesting and recommendation*: There are large numbers of MOD drivers delivering passengers in the city, willing to opt in MOVE-CS. Let $\mathcal{M}$ denote the set of these MOD drivers, *i.e.,* $\mathcal{M} = \{1, \ldots, M\}$. Each driver k ($k \in \mathcal{M}$) reports her/his current zone z_k. Then, the platform recommends one task for each driver, along with its location and reward, and the driver's expected profit. Let x_{kj} denote whether task j is recommended to driver k, *i.e.,* $x_{kj} = 1$ if yes, and $x_{kj} = 0$ otherwise. The recommendation set is then denoted by $\mathbf{x} = \{x_{kj}\}$.

(3) *Task acceptance and execution*: Once a driver k is recommended task j, s/he has a probability ρ_{kj} of accepting and performing it. Each driver's acceptance probability depends on her/his preference and reliability, as well as the recommended task, which can be learned from large amounts of MOD vehicle data [16]. To assure a high task execution probability, a task may be recommended to multiple drivers. Hence, the execution probability of each task j can be calculated by $1 - \prod_{k=1}^{M}(1 - \rho_{kj}x_{kj})$. Similar to [17], [18], the total platform profit, *i.e.,* the expected profits of all the performed sensing tasks for the platform, is denoted as

$$U(\mathbf{x}) = \sum_{j=1}^{S} u_j\big(1 - \prod_{k=1}^{M}(1 - \rho_{kj}x_{kj})\big). \tag{1}$$

Finally, after the road data of task j is uploaded to the MOVE-CS platform, driver k gets a reward c_{kj}; the rewards of all drivers then form a set $\mathbf{c} = \{c_{kj}\}$. Hence, the driver's total earnings equal to the sum of the reward by performing a sensing task and the pick-up profit by transporting passengers.

B. Research Problem and Challenge Analysis

The crucial problem in the *LSTRec* model design is how to recommend tasks to drivers alongside proper rewards, by study and prediction of the spatial-temporal differences, which is presented in the following.

Long-short-term profit-aware optimal task recommendation problem (*LSTO*): *Given the historical MOD vehicle dataset, how to recommend each task j to an MOD driver k with the sensing reward $\{c_{kj}\}$, so as to maximize the total platform profit $U(\mathbf{x})$ under the constraint of budget B, while satisfying both drivers' explicit preference for short-term gains and their implicit needs of long-term profits.*

In addressing the above-mentioned problem, there exist three main challenges as follows:

(1) *It is difficult to predict the global distribution of the pick-up profits, due to their spatial-temporal dynamics*. The pick-up profits exhibit spatial-temporal dynamics, as demonstrated in Figs. 1a, 3a, and 3b. Furthermore, the highly complicated movement of both passengers and MOD vehicles between different zones and time complicates such dynamics, hence rendering the accurate prediction on the global distribution of pick-up profits particularly difficult.

(2) *It is challenging to satisfy the demands of both drivers and the platform, which are aligned in some cases but conflicted in others*. In some circumstances, a task may require sensing in a high-yield zone where drivers are eager to move
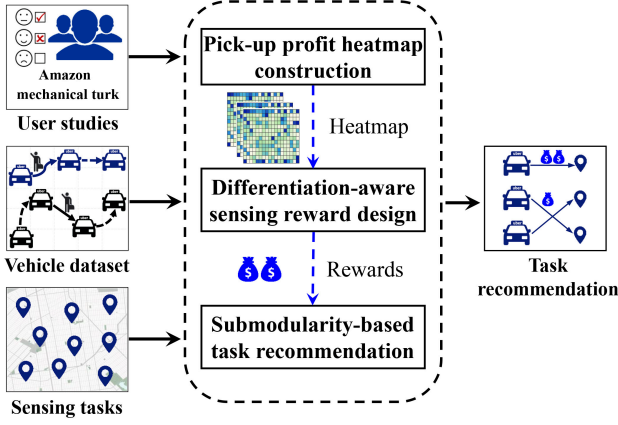
Fig. 4: Overview of our algorithm.

Fig. 5: Illustration of pick-up profit heatmap construction based on the MOD vehicle dataset using dual-attention-based RNN.

A. Pick-up Profit Heatmap Construction

Two-dimensional pick-up profit heatmap model. The pick-up profit is highly dependent on the pick-up probability and the per-trip earnings in each zone. Furthermore, the dataset analysis in Sec. II-B indicates that both the pick-up probability and the per-trip earnings have the spatial-temporal dynamics in different zones and time periods. As a result, we use the two-dimensional heatmaps, called *pick-up profit heatmaps*, to represent the dynamic spatial-temporal pick-up profits.

In particular, we divide the map of an entire city into Z non-overlapping zones, according to the shape of the area and the specified spatial granularity. Let z_i and $\mathcal{Z}$ denote each zone i and the set of zones, respectively, such that $z_i \in \mathcal{Z}$. Similarly, the time is evenly divided into T time slots, and the set of time slots is denoted as $\mathcal{T}$. Each time slot t is also named period t. Let p_i^t and r_i^t denote the pick-up probability and the per-trip earnings in zone i at period t. Hence, the pick-up profit heatmaps $\mathbf{H}^T$ during the periods $[1, T]$ are represented as

$$\mathbf{H}^T = \left\{ \mathbf{h}^t | 1 \leq t \leq T \right\}, \tag{2}$$
$$\mathbf{h}^t = \left\{ (p_i^t, r_i^t) | \forall i \in \mathcal{Z} \right\}, \tag{3}$$

where $\mathbf{h}^t$ denotes the t-th frame of the heatmaps, representing the pick-up profits of all the zones at period t. Moreover, each pixel of the heatmap frame (*i.e.,* $(p_i^t, r_i^t) \in \mathbf{h}^t$) represents the pick-up probability and the per-trip earnings in zone i at period t. Intuitively, in the heatmaps, the warmer the color, the more the pick-up profits that drivers are expected to get (with higher pick-up probability and more per-trip earnings), as illustrated in Figs. 3a and 3b.

Heatmap construction based on MOD vehicle dataset. We construct the pick-up profit heatmap by using the historical MOD vehicle dataset, including their trajectories, occupied/vacant statuses, and pick-up earnings of each trip, as illustrated in Fig. 5a. Specifically, as illustrated in Fig. 5b, based on the trajectories of MOD vehicles and their occupied/vacant status information, we compute the ratio of the vehicles picking up passengers to all the vacant vehicles in zone i within period t as the pick-up probability p_i^t. Moreover, we compute the per-trip earnings r_i^t by using the average income of all the vehicles that pick up passengers from zone i at period t. Thus, we use the historical MOD vehicle dataset during the periods $[1, T]$ to construct the corresponding pick-up profit heatmaps, *i.e.,* $\mathbf{H}^T = \left\{ (p_i^t, r_i^t) | \forall i \in \mathcal{Z}, 1 \leq t \leq T \right\}$.

Heatmap prediction based on RNN. As shown in Fig. 5c, we utilize the pick-up profit heatmaps $\mathbf{H}^T$ of periods $[1, T]$ to accurately predict the future L periods. Note that, the prediction length L is dependent on the time interval of each task recommendation in Sec. IV-B. In particular, the dual-attention based

towards, so that they will probably accept it with a relatively low reward, in alignment with the platform's interest. In other conditions, a task valuable to the platform is perhaps related to an unpopular zone, where drivers are reluctant to go. The two sides do not share mutual benefits so that drivers will only undertake the task if the reward is high enough to reach their expectations, thereby increasing the platform's cost.

(3) *The optimal task recommendation subproblem of LSTO is NP-hard*. This problem can be reduced from the classical *0-1 knapsack problem* [19]: Given a capacity B and a group of items $\{(k, j) | \forall k \in \mathcal{M}, \forall j \in \mathcal{S}\}$, each of which has a value c_{kj} and a weight ρ_{kj}, select a collection of items to maximize the total value $U(\mathbf{x})$ under the capacity constraint of weights (the detailed proofs are omitted owing to the page limit). As a result, it is extremely challenging to achieve the optimal recommendation with computational efficiency, especially for the large-scale MOVE-CS market with massive drivers (such as 12,493 drivers in our dataset).

IV. KEY ALGORITHM DESIGN FOR *LSTRec*

To address the above three challenges, we propose a spatial-temporal differentiation-aware task recommendation scheme empowered by submodular optimization. As illustrated in Fig. 4, it mainly consists of three components:

(1) *Pick-up profit heatmap construction* (Sec. IV-A): Utilizing the historical MOD vehicle dataset, we first construct the two-dimensional pick-up profit heatmaps, which are then used to predict the future heatmaps by exploiting dual-attention-based RNN.

(2) *Differentiation-aware sensing reward design* (Sec. IV-B): Based on the global knowledge of the pick-up profit heatmaps, we learn the spatial-temporal dynamics of pick-up profits, which is fed back to devise the sensing rewards for satisfying the driver's explicit and implicit needs of the short-term and long-term profits, respectively.

(3) *Submodularity-based task recommendation* (Sec. IV-C): Given the sensing reward design, we first analyze the properties of the optimal task recommendation problem by reformulation. Guided by the analysis results, we present an approximation algorithm to address this NP-hard problem, following the methodology of submodular optimization.
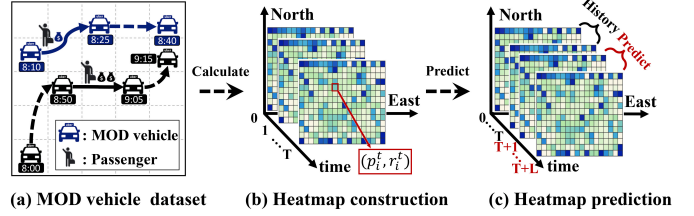
RNN [20], [21] is exploited to accurately predict the next few profit heatmaps based on the historical ones. Its key idea is to leverage an LSTM-based encoder-decoder architecture with dual-attention mechanisms, including spatial attention and temporal attention. The spatial attention is exploited to capture the complex spatial correlations across different zones, while the temporal attention is used to learn the time-varying correlations between different time periods. In summary, we use the MOD vehicle dataset to construct and predict the global pick-up profit heatmaps of the periods $[1, T + L]$, *i.e.,* $\mathbf{H}^{T+L} = \{(p_i^t, r_i^t) | \forall i \in \mathcal{Z}, 1 \le t \le T + L\}$.

B. Differentiation-aware Sensing Reward Design

Based on the global pick-up profit heatmaps, we first compute the spatial-temporal differences of pick-up profits, which are used to design the sensing rewards of drivers.

The pick-up profit difference represents the expected pick-up profit increase for an unoccupied vehicle to drive from one zone to another for performing the sensing task. Formally, we let $I_{kj}(t_0)$ represent the pick-up profit difference for driver k, when moving from her/his original zone z_k to the task's zone z_j at t_0. Thus, $I_{kj}(t_0)$ is given by

$$I_{kj}(t_0) = \mathbb{E}_{t_0}[p_j^t \cdot r_j^t] - \mathbb{E}_{t_0}[p_k^t \cdot r_k^t], \qquad (4)$$

where $\mathbb{E}_{t_0}[\cdot]$ denotes the mathematical expectation with respect to t_0. $\mathbb{E}_{t_0}[p_k^t \cdot r_k^t]$ denotes the expected pick-up profit of driver k at her/his original zone z_k; $\mathbb{E}_{t_0}[p_j^t \cdot r_j^t]$ represents that of driver k at the new zone z_j. Both of them are dependent on the probability p_k^t (p_j^t) that drivers pick up passengers in z_k (z_j) at period t, and the probability $\bar{p}_k^t$ ($\bar{p}_j^t$) that drivers cannot do it in z_k (z_j) before t. Hence, both $\mathbb{E}_{t_0}[p_k^t \cdot r_k^t]$ and $\mathbb{E}_{t_0}[p_j^t \cdot r_j^t]$ can be computed based on the global pick-up profit heatmaps $\mathbf{H}^{T+L} = \{(p_i^t, r_i^t) | \forall i \in \mathcal{Z}, 1 \le t \le T + L\}$, using the expectation model [22].

Learning each driver's pick-up profit difference, we design a sensing reward model, by subtracting it from her/his expected income. In specific, let $b_{kj}(t_0)$ denote driver k's expected income from passenger missions, when driving from z_k to z_j. Thus, according to Eq. (4), the sensing reward $c_{kj}(t_0)$ for executing task j by driver k at t_0 is represented as

$$c_{kj}(t_0) = b_{kj}(t_0) - (\mathbb{E}_{t_0}[p_j^t \cdot r_j^t] - \mathbb{E}_{t_0}[p_k^t \cdot r_k^t]). \quad (5)$$

Note that $b_{kj}(t_0)$ can be computed based on the driving time and distance [23], [24], according to the pricing policies of MOD vehicles [15].

Referring to the actual hourly wage of drivers, the platform offers each of them a reward, called the expected profit, so that drivers are willing to spend time on the sensing tasks. A driver's expected profit is a combination of her/his explicit reward $c_{kj}(t_0)$ directly given by the platform, and implicit reward $\mathbb{E}_{t_0}[p_j^t \cdot r_j^t] - \mathbb{E}_{t_0}[p_k^t \cdot r_k^t]$ subtly obtained by relocation to a higher-yield zone for task j. Thus, as shown in Eq. (5), if the implicit reward is adequate, the platform can lower the explicit reward; if it is insufficient, the platform should offer a higher explicit reward for compensation. As a result, all drivers are granted more profits than the rewards from only passenger missions, named positive profits. In sum, *the sensing reward design based on the pick-up profit differentiation learning can*

ensure positive profits of all the drivers by balancing the explicit and implicit rewards.

C. Submodularity-based Task Recommendation Algorithm

Problem analysis based on equivalent transformation. Firstly, given the sensing reward design $\{c_{kj}\}$, we can equivalently transform the LSTO problem into a set function optimization problem for facilitating the problem analysis. In particular, we define the ground set $\mathcal{V} := \{v = (k, j) | \forall k \in \mathcal{M}, \forall j \in \mathcal{S}\}$. Let $\mathcal{A}$ denote the set of the recommended driver-task pairs, *i.e.,* $\mathcal{A} := \{v = (k, j) | x_{kj} = 1, \forall k \in \mathcal{M}, \forall j \in \mathcal{S}\}$, and $\mathcal{A} \subseteq \mathcal{V}$. Moreover, $\forall \mathcal{A} \subseteq \mathcal{V}$, $U(\mathcal{A}) := \{U(\mathbf{x}) | \forall (k, j) \in \mathcal{A}, x_{kj} = 1; \forall (k, j) \notin \mathcal{A}, x_{kj} = 0\}$. It is noted that $U(\mathcal{A})$ and $U(\mathbf{x})$ are different functions, but we adopt the same symbol $U(\cdot)$ for simplification. Based on the above definitions, the optimal task recommendation subproblem of *LSTO* can be equivalently transformed as

$$\underset{\mathcal{A} \subseteq \mathcal{V}}{\text{Max}} \quad U(\mathcal{A}) = \sum_{j \in \mathcal{S}} u_j \big(1 - \prod_{k \in \mathcal{M}} (1 - \rho_{kj} x_{kj})\big), \quad (6)$$

$$\text{s.t.} \quad \sum_{j:(k,j) \in \mathcal{A}} \mathbb{1}_{(k,j) \in \mathcal{A}} \le 1, \quad (7)$$

$$\sum_{(k,j) \in \mathcal{A}} c_{kj} \rho_{kj} \le B, \quad (8)$$

where $\mathbb{1}$ denotes the indicator function. Eq. (7) restricts each driver is recommended at most one task in each recommendation period, similar to [25], [26]. Eq. (8) indicates that the expected sensing rewards of all the drivers are less than the platform's budget.

Furthermore, based on the set function optimization problem, we theoretically analyze the properties of the problem. At first, we find out the objective function of this problem is *submodular*. In particular, let $\mathcal{A}_1 \subseteq \mathcal{A}_2$; $\mathcal{A}_1^0 := \{(k, j_0) | \forall k, (k, j_0) \in \mathcal{A}_1\}$; $\mathcal{A}_2^0 := \{(k, j_0) | \forall k, (k, j_0) \in \mathcal{A}_2\}$. Hence, $U(\mathcal{A}_1 \cup \{v_0\}) - U(\mathcal{A}_1) = u_{j_0} \rho_{k_0 j_0} \prod_{k:(k,j) \in \mathcal{A}_1^0} (1 - \rho_{kj})$, and $U(\mathcal{A}_2 \cup \{v_0\}) - U(\mathcal{A}_2) = u_{j_0} \rho_{k_0 j_0} \prod_{k:(k,j) \in \mathcal{A}_2^0} (1 - \rho_{kj})$. Since $\prod_{k:(k,j) \in \mathcal{A}_1^0} (1 - \rho_{kj}) \ge \prod_{k:(k,j) \in \mathcal{A}_2^0} (1 - \rho_{kj})$, according to the definition of submodularity [27], the objective function is submodular. Moreover, it is easy to prove that the objective function is also non-negative and monotone [27]. In addition, according to the definitions of the matroid constraint and the knapsack constraint [28], we find constraints (7)(8) to be a matroid constraint and a knapsack constraint, respectively. The detailed proofs are omitted due to the page limit.

In summary, we have the following results: *The optimal task recommendation subproblem of LSTO is maximizing a non-negative, monotone, and submodular objective function with a matroid constraint and a knapsack constraint.*

Algorithm design. Following the methodology of submodular optimization [29], we propose *a greedy local search-based near-optimal task recommendation algorithm*, leveraging two key ideas: (1) Since this problem is maximizing a monotone and submodular objective function, we use the *greedy strategy* to iteratively search the local optimal solution, which has the *largest marginal profit-cost ratio*. (2) In each iteration, we use *local search* to swap a non-recommended driver-task pair with

Algorithm 1: Greedy Local Search-based Near-optimal Task Recommendation Algorithm.

Input: Task set $\mathcal{S}$; MOD driver set $\mathcal{M}$;
Sensing rewards set $\{c_{kj}\}$;
Set of drivers' acceptance probability $\{\rho_{kj}\}$;
Set of tasks' profits to the platform $\{u_j\}$; Budget B;
Output: Recommended task set $\{x_{kj}\}$; Platform profit U.

1   Initialize $\mathcal{A}_0 = \{v_0, v_1\}$, where $v_0 = \arg\max_{v \in \mathcal{V}} U(\{v\})$,
     $v_1 = \arg\max_{v \in \mathcal{V}\setminus\{v_0\}} U(\{v, v_0\}) - U(\{v_0\})$;
2   Initialize $n = 0$, and $swap = true$;
3   **while** $swap$ **do**
4      $swap \leftarrow false$;
5      $\mathcal{V}_s := \{(v_+, v_-)|\forall v_+ \in \mathcal{V}\setminus\mathcal{A}_n, \forall v_- \in \mathcal{A}_n \cup \{\emptyset\}\}$;
6   **while** $(swap \neq true)$ && $(\mathcal{V}_s \neq \emptyset)$ **do**
7          $(v_+^*, v_-^*) = \arg\max_{(v_+, v_-) \in \mathcal{V}_s} \pi(v_+, v_-)$;
8   **if** $\mathcal{A}_n\setminus\{v_-^*\} \cup \{v_+^*\}$ satisfies constraints (7)(8) and $\pi(v_+^*, v_-^*) \geq \frac{\epsilon}{M^2 S^2}$ **then**
9              $\mathcal{A}_{n+1} \leftarrow \mathcal{A}_n\setminus\{v_-^*\} \cup \{v_+^*\}$;
10             $n \leftarrow n + 1$;
11             $swap \leftarrow true$;
12          $\mathcal{V}_s \leftarrow \mathcal{V}_s\setminus\{(v_+, v_-)\}$;
13   Set $\mathbf{x} \leftarrow \{x_{kj} = 1|\forall k, \forall j, (k, j) \in \mathcal{A}_n\}$;
14   Compute $U(\mathbf{x})$ based on $\mathbf{x}$, $\{u_j\}$, and $\{\rho_{kj}\}$, according to Eq. (6);
15   **return** $\mathbf{x}$ and $U(\mathbf{x})$.

Fig. 6: Drivers' profits in the two models (a), and the profit increase ratio of *LSTRec* compared with the original MOVE-CS model (b).

a recommended one, for profit maximization under the matroid constraint (7); a swap is applied if the marginal profit-cost ratio of this swap is more than the lower bound $\frac{\epsilon}{M^2 S^2}$.

In specific, the detailed algorithm design is presented in Alg. 1. Let $\mathcal{A}_n$ denote the set of recommended driver-task pairs in the n-th iteration. Let (v_+, v_-) and $\mathcal{V}_s$ denotes a swap and the swap set, respectively, where v_+ represents a non-recommended pair, *i.e.,* $v_+ = (k_+, j_+) \in \mathcal{V}\setminus\mathcal{A}_n$; v_- denotes a recommended one, *i.e.,* $v_- = (k_-, j_-) \in \mathcal{A}_n \cup \{\emptyset\}$; $\mathcal{V}_s = \{(v_+, v_-)\}$. Note that $\emptyset$ represents a dummy element; swapping v_+ with $\emptyset$ is equivalent to directly adding v_+ into $\mathcal{A}_n$. As a result, the *marginal profit-cost ratio* of this swap (v_+, v_-), *i.e.,* the ratio of the platform profit increase via swap to the sensing reward of v_+, is represented by

$$\pi(v_+, v_-) = \frac{U(\mathcal{A}\setminus\{v_-\} \cup \{v_+\}) - U(\mathcal{A})}{c_{k_+ j_+}}. \quad (9)$$

Based on Alg. 1, we analyze the theoretical performance of this algorithm design. In specific, the number of its iterations is no more than $\frac{M^2 S^2}{\epsilon} \log(MS)$, since the profit increase by swap in each iteration should be at least $\frac{\epsilon}{M^2 S^2}$. Moreover, the time complexity of each iteration in lines 3-12 is $O(M^2 S^2)$. Since ϵ is a constant, the time complexity of Alg. 1 is $O(M^4 S^4 \log(MS))$. Furthermore, according to the aforementioned analysis, this problem is a monotone, submodular maximization problem with a matroid constraint and a knapsack constraint. As a result, referring to [29], Alg. 1 can achieve a $(1 - e^{-2})/2$-approximation ratio.

In conclusion, *Alg. 1 can achieve a near-optimal solution of $(1-e^{-2})/2$-approximation with the polynomial time complexity $O(M^4 S^4 \log(MS))$, where M and S denote the numbers of drivers and tasks, respectively.*
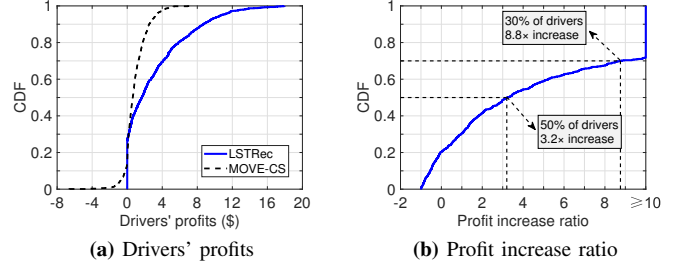
V. EVALUATION

We use the large-scale MOD vehicle dataset to emulate the operation process of the original MOVE-CS model and the *LSTRec* model, respectively. Furthermore, we comprehensively compare the performance of the proposed algorithm with five baseline algorithms.

A. Emulation Methodology and Settings

The emulation of the proposed *LSTRec* model and the original MOVE-CS model is based on the large-scale MOD vehicle dataset (specified in Sec. II-B) as follows. First, the MOVE-CS platform requires road data of 878 road segments with a total length of 191.1 miles in a 32 km^2 area. Then, for higher accuracy, each road is required to be sensed k times with decreasing profit u to the platform ($k = 3$, $u = \$2.5, 1.5$, and 0.5 per mile for the three times respectively). There are M MOD drivers ($M = 1000$), randomly selected as participants willing to collect the road data for the MOVE-CS market. Next, we run the emulation for five days, which may end in advance if the budget is exhausted.

For the MOVE-CS model, drivers collect road data on the move at any time s/he wants during her/his work hours. Drivers on average spend $0.06 per mile on fuels [30]; their data collection costs are only induced in the unoccupied state during extra trips for tasks. After the data are uploaded, each driver gets a reward. The original settings Payver adopted ($0.01-0.05 per mile) are so unreasonable that most participants can only get few or even negative profits. To make a fair comparison, in the emulation, in contrast, we let a portion (1/a) of the platform profit u be the reward, *e.g.,* k=3, $1/a$=0.2, and the rewards are $0.5, 0.3, 0.1$ per mile respectively, according to the economic theory [31]. For the *LSTRec* model, there are N rounds of task recommendations. In each round, as explicated in Sec. III-A, the platform publishes sensing tasks; each corresponds to one collection of a road segment. The platform then predicts drivers' pick-up profits and recommends them sensing tasks. Each driver accepts and accomplishes the task at a probability following the stochastic uniform distribution $\mathcal{U}(0, 1)$. Once the task is accomplished, s/he receives the reward given by the proposed algorithm. We implement the emulation on a commodity server with 3.00GHz dual-core Intel Core Xeon Gold 6561 CPU and 192GB RAM.
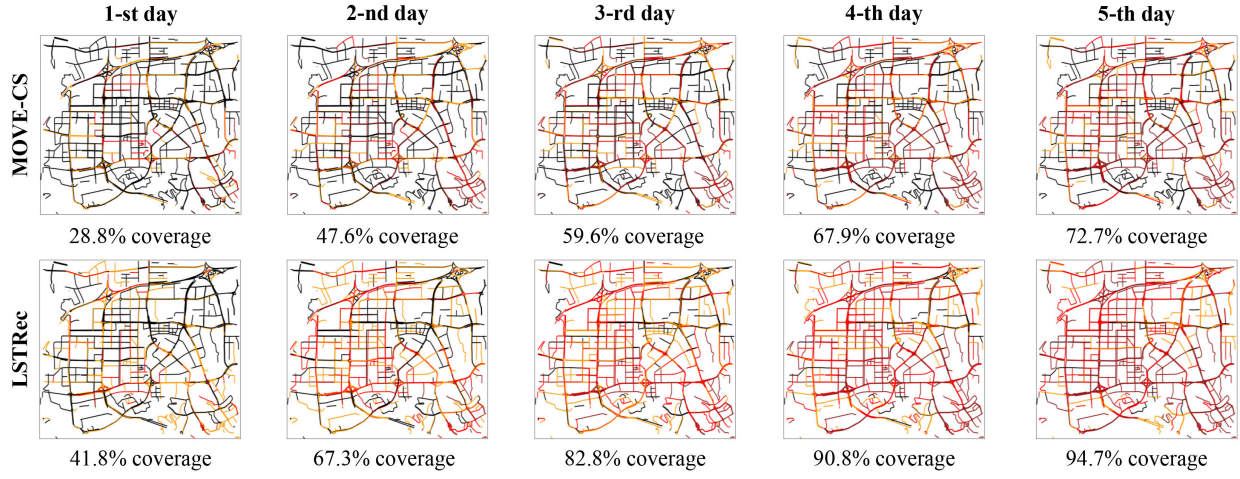
Fig. 7: Road coverage heatmaps in the two models; black, orange, red, and crimson represent collected times of 0, 1, 2, and 3, respectively.

B. Results of Model Evaluation

Drivers' profits. We first evaluate the two models on drivers' profits. As illustrated in Fig. 6a, for the MOVE-CS model, 14.5% of drivers have negative profits from sensing tasks, because they might spend a lot on collecting repeated road data, resulting in rewards far less than the driving cost. In contrast, all the drivers in the *LSTRec* model make positive profits, thanks to the sensing reward design based on the spatial-temporal differentiation of pick-up profits. Moreover, we analyze all the task recommendation results in *LSTRec*. Results show that, 87.3% of the recommended tasks enable drivers from low-yield zones to high-yield zones, consistent with their desires for immediate gains.

Furthermore, we compare the drivers' profits in the two models by calculating the drivers' profit increase ratios in *LSTRec* to those in MOVE-CS. As demonstrated in Fig. 6b. we find in our model, 50% of drivers increase profits by 320%, and 30% have an increase ratio of 880%, compared with MOVE-CS. Further analysis indicates its effectiveness is anchored in the active task recommendation scheme, *i.e.,* incentivizing drivers to complete the tasks suitable for them. Besides, Fig. 6b shows that 20% of drivers suffer decreased profits (than in the MOVE-CS model), due to no tasks recommended to them, which can be solved by prior recommendation to these drivers in the next round.

Platform's profit. We evaluate the platform profit in the two models. We first visualize the coverage heatmap of collected road segments throughout the five days. As shown in Fig. 7, our coverage ratio of collected roads in each day is consistently higher than that in MOVE-CS. Our coverage ratio increases day by day, and ends up 94.7% in the last day, 22.0% higher than that in MOVE-CS. Meanwhile, our platform profit increases by 34.3%, also attributed to the active task recommendation scheme, *i.e.,* encouraging drivers to unpopular roads, increasing the road coverage ratio as well as the platform profit.

Impacts of parameters. We evaluate the impacts of the number of drivers on the model performance, in the aspects of drivers' profits and the platform profit. As shown in Fig. 8a, we illustrate the box-plot of the drivers' profits in the two models.

The results show that *LSTRec* can guarantee all the drivers' positive profits, while 14.6% of drivers have negative profits in MOVE-CS model. Moreover, the drivers' profits of the two models decrease with the number of drivers, since more opt-in drivers lead to more fierce competition for earnings. However, the decrease ratio in MOVE-CS is averagely 32.2% higher than that in *LSTRec*. Moreover, Fig. 8b demonstrates that the platform profits of the two models improve with the number of drivers. The platform profit of *LSTRec* outperforms that of MOVE-CS by 45.8% on average. Other parameters (*e.g.,* budget) show similar effects on results, so we do not show them due to the page limit.

C. Results of Algorithm Evaluation

Baseline algorithms. To comprehensively evaluate the performance of the key algorithm of *LSTRec*, we exploit five baselines as follows: (1) *Hector* [24] greedily recommends sensing tasks with maximal marginal profit-cost efficiency to the drivers, while using their basic driving costs as the rewards. (2) *GA* [32] exploits the Genetic algorithm to maximize the platform profit with the assumption that the sensing rewards are already given. (3) *iLOCuS* [23] recommends the sensing tasks greedily to minimize the task distribution divergence, while utilizing the high pick-up probability of the task's zone as the hidden incentives. (4) *RAD* randomly recommends tasks with a uniform distribution of pick-up profits. (5) *OPT* uses the brutal-force search method to achieve the optimal solution with exponential time cost. In the remaining, we call the proposed algorithm *LSTRec* as well for simplification.

Comparison of algorithms. We first evaluate the platform profit of the *LSTRec* algorithm in different numbers of drivers and tasks, compared to five baselines. As demonstrated in Figs. 9a and 9b, the platform profit of *LSTRec* exceed those of RAD, GA, Hector, and iLOCus by 466.8%, 103.2%, 61.7%, and 257.1% in different numbers of drivers, respectively, and by 516.5%, 132.7%, 44.4%, and 237.8% in different numbers of tasks, respectively. Also, we evaluate the near-optimality of *LSTRec* by comparing it with OPT in a small-scale scenario (*i.e.,* $M = 10, S = 6$). As illustrated in Fig. 10a, *LSTRec* can averagely achieve 97.2% of the optimal platform profit
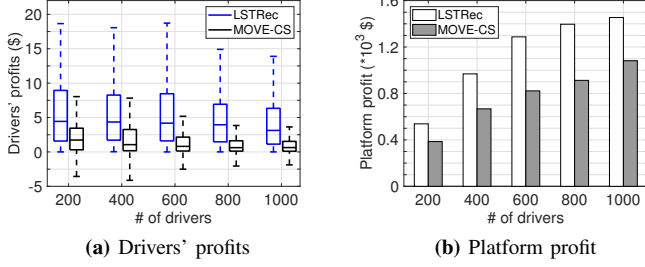
(a) Drivers' profits   **(b)** Platform profit

Fig. 8: Impacts of different numbers of drivers on the performance of two models, in terms of drivers' profits and platform profit.
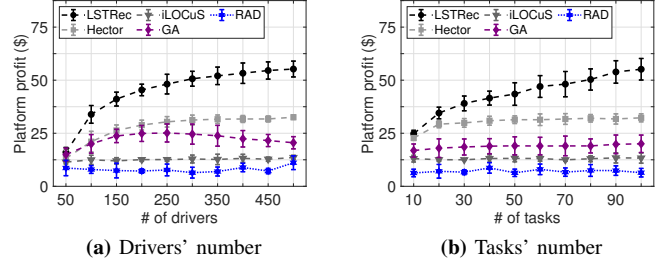

(a) Drivers' number   **(b)** Tasks' number

Fig. 9: Comparison of the *LSTRec* algorithm and four baselines in terms of the platform profit in different numbers of drivers and tasks.

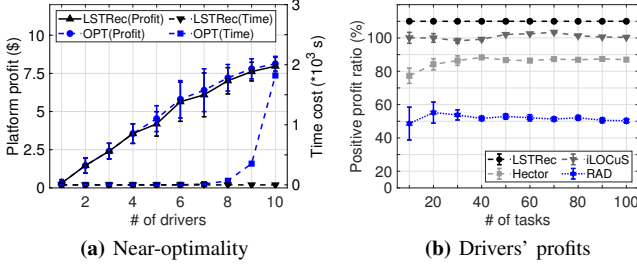**(a)** Near-optimality   **(b)** Drivers' profits

Fig. 10: Comparison of the *LSTRec* algorithm and four baselines in terms of the near-optimality and the drivers' profits.

with only 0.004% of OPT's time cost in different numbers of drivers. As for drivers' profits, there exists a large gap between *LSTRec* and baselines. As shown in Fig. 10b, *LSTRec* guarantees a 100% positive profit ratio (*i.e.,* the percentage of MOD drivers with positive profits after opting in MOVE-CS), outperforming Hector, iLOCus, and RAD by 24.2%, 9.2%, and 58.1%. Hence, differing from the three baselines, *LSTRec* ensures a positive profit for every driver, thanks to consideration of long-term and short-term profits of drivers.

VI. RELATED WORK

Recently, there have been considerable studies of incentivized mobile crowdsensing [33]–[37], most of which focus on human mobility [34]–[36] without considering the special impacts of vehicle mobility. Since this paper belongs to the category of incentivizing vehicular crowdsensing [16], [38], we focus on reviewing its related studies; other orthogonal studies can be referred to [38]–[40].

Specifically, as one of the earliest works, He *et al.* [32] design a participant recruitment strategy, jointly leveraging both the current location and the predictable mobility pattern of vehicles. Then, Wang *et al.* [41] study both the deterministic and probabilistic trajectory models and propose two efficient vehicle recruitment algorithms. Zhu *et al.* [42] use RNN to predict the future vehicle mobility, which is used to select vehicles to maximize their coverage with limited budget. Moreover, Fan *et al.* [24] propose Hector, a novel joint scheduling and incentive mechanism of vehicular crowdsensing. The above works focus on common vehicles without concerning the special MOD vehicles. In contrast, a recent work called iLOCuS [23], highly related to this paper, takes MOD vehicles into account and proposes a hybrid incentive, which combines the monetary rewards and the non-monetary hidden incentives (*i.e.,* the passenger's requests at the task's zone). Nevertheless, iLOCuS

neglects the in-depth demands of MOD drivers for short-term and long-term profits, inefficient to encourage them. Xiang *et al.* [22] propose a sensing task allocation scheme based on the deep reinforcement learning, achieving a near optimal solution with a factor which depends on the maximal and minimal costs of all the sensing tasks. Contrarily, this work uses the greedy local search to achieve a $(1 - e^{-2})/2$-approximation ratio, thereby having more robustness in the real applications with different settings. Distinguished from existing works, we conduct user studies and dataset-based in-depth analysis, uncovering both the explicit and implicit needs of MOD drivers. The results are then fed back to design a novel *LSTRec* model to benefit both drivers and the platform.

In addition, many efficient recommendation systems [43], [44] are proposed, such as web service recommendation [45] and social network recommendation [46], [47]. Also, there are numbers of good works about allocation/recommendation of passenger missions for MOD drivers. For example, Xu *et al.* [48] design an effective order dispatching algorithm, which considers both the immediate passenger satisfaction and the expected future income of drivers. Nevertheless, all of the works are merely focused on either passenger missions or item recommendation, which are orthogonal to our work.

VII. CONCLUSION

In this paper, motivated by findings in user studies and the large-scale vehicle dataset analysis, we propose *LSTRec*, a new Long-Short-Term Profit-combined Task Recommendation model, in an attempt to resurrect the MOVE-CS market. Behind it lies a spatial-temporal differentiation-aware task recommendation scheme empowered by submodular optimization. It involves pick-up heatmap prediction based on RNN, the differentiation-aware sensing reward design, and the submodularity-based task recommendation algorithm. The emulation reveals that *LSTRec* guarantees not only positive profits for drivers, but also a near-optimal profit for the platform, hence having the potential to resurrect MOVE-CS. In the future, we will explore the possible deployment of our *LSTRec* model by collaborating with MOD companies.

REFERENCES

[1] X. Tang, Z. Qin, F. Zhang, Z. Wang, Z. Xu, Y. Ma, H. Zhu, and J. Ye, "A deep value-network based approach for multi-driver order dispatching," in *Proc. ACM KDD*, Anchorage, AK, USA, 2019, pp. 1780–1790.

[2] "How many uber drivers are there?" 2021. [Online]. Available: https://therideshareguy.com/how-many-uber-drivers-are-there/.

[3] "Global mobility on demand market forecast & opportunities by 2022," 2017. [Online]. Available: https://www.techsciresearch.com/report/global-mobility-on-demand-market/1254.html.

[4] "Drivers for uber, lyft are earning less than half of what they did four years ago, study finds," 2019. [Online]. Available: https://www.marketwatch.com/story/drivers-for-uber-lyft-are-earning-less-than-half-of-what-they-/did-four-years-ago-study-finds-2018-09-24.

[5] "Uber and lyft have a driver shortage problem, and it's costing them a lot of money," 2021. [Online]. Available: https://www.theverge.com/2021/4/7/22371850/uber-lyft-driver-shortage-covid-bonus-stimulus.

[6] "Payver," 2019. [Online]. Available: https://www.lazymoneyguy.com/payver/.

[7] "Google maps," 2021. [Online]. Available: https://www.google.com/maps.

[8] "Lvl5," 2020. [Online]. Available: https://www.crunchbase.com/organization/lvl5.

[9] "Thanks to a dashcam, crafty uber drivers are boosting their pay," 2017. [Online]. Available: http://money.cnn.com/2017/07/19/technology/business/rideshare-drivers-camera/index.html.

[10] "Doordash acquires autonomous driving startup scotty labs," 2019. [Online]. Available: https://techcrunch.com/2019/08/20/doordash-acquires-autonomous-driving-startup-scotty-labs/.

[11] "Amazon mechanical turk," 2021. [Online]. Available: https://www.mturk.com/.

[12] "Gigwalk," 2021. [Online]. Available: https://www.gigwalk.com/.

[13] "User study form," 2021. [Online]. Available: https://forms.gle/f6rSLextsfaXy8JS9.

[14] A. M. Lund, "Measuring usability with the USE questionnaire," *Usability interface*, vol. 8, no. 2, pp. 3–6, 2001.

[15] "How much does a ride with the uber app cost?" 2020. [Online]. Available: https://www.uber.com/global/en/price-estimate/.

[16] G. Gao, M. Xiao, J. Wu, L. Huang, and C. Hu, "Truthful incentive mechanism for nondeterministic crowdsensing with vehicles," *IEEE Transactions on Mobile Computing*, vol. 17, no. 12, pp. 2982–2997, 2018.

[17] X. Zhang, G. Xue, R. Yu, D. Yang, and J. Tang, "Truthful incentive mechanisms for crowdsourcing," in *Proc. IEEE INFOCOM*, HongKong, China, 2015, pp. 2830–2838.

[18] D. Zhang, H. Xiong, L. Wang, and G. Chen, "Crowdrecruiter: selecting participants for piggyback crowdsensing under probabilistic coverage constraint," in *Proc. ACM Ubicomp*, Seattle, Washington, USA, 2014, pp. 703–714.

[19] S. Martello, *Knapsack problems: algorithms and computer implementations*. Wiley-Interscience series in discrete mathematics and optimization, 1990.

[20] X. Fan, C. Xiang, C. Chen, P. Yang, L. Gong, X. Song, P. Nanda, and X. He, "Buildsensys: reusing building sensing data for traffic prediction with cross-domain learning," *IEEE Transactions on Mobile Computing*, vol. 20, no. 6, pp. 2154–2171, 2021.

[21] J. Zhang, S. Liang, Z. Deng, and J. Shao, "Spatial-temporal attention network for temporal knowledge graph completion," in *Proc. DASFAA*, 2021, pp. 207–223.

[22] C. Xiang, Y. Li, L. Feng, C. Chen, S. Guo, and P. Yang, "Near-optimal vehicular crowdsensing task allocation empowered by deep reinforcement learning," *Chinese Journal of Computers*, pp. 1–17, 2021 (Early Access).

[23] S. Xu, X. Chen, X. Pi, C. Joe-Wong, P. Zhang, and H. Y. Noh, "iLOCuS: incentivizing vehicle mobility to optimize sensing distribution in crowd sensing," *IEEE Transactions on Mobile Computing*, vol. 19, no. 8, pp. 1831–1847, 2020.

[24] G. Fan, H. Jin, Q. Liu, W. Qin, X. Gan, H. Long, L. Fu, and X. Wang, "Joint scheduling and incentive mechanism for spatio-temporal vehicular crowd sensing," *IEEE Transactions on Mobile Computing*, vol. 20, no. 4, pp. 1449–1464, 2021.

[25] Z. Zhou, J. Feng, B. Gu, B. Ai, S. Mumtaz, J. Rodriguez, and M. Guizani, "When mobile crowd sensing meets UAV: energy-efficient task assignment and route planning," *IEEE Transactions on Communications*, vol. 66, no. 11, pp. 5526–5538, 2018.

[26] K. Dorling, J. Heinrichs, G. G. Messier, and S. Magierowski, "Vehicle routing problems for drone delivery," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 70–85, 2017.

[27] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions-I," *Mathematical Programming*, vol. 14, no. 1, pp. 265–294, 1978.

[28] J. G. Oxley, *Matroid theory*. Oxford University Press, 1992.

[29] K. Sarpatwar, B. Schieber, and H. Shachnai, "Constrained submodular maximization via greedy local search," *Operations Research Letters*, vol. 47, no. 1, pp. 1–6, 2019.

[30] "Driving cost of mod vehicles," 2019. [Online]. Available: https://www.greencarreports.com/news/1123364_nyc-taxi-mpg-requirements-also-cut-pollution-study-confirms.

[31] G. S. Becker, M. Grossman, and R. T. Michael, *Economic theory*. Routledge, 2017.

[32] Z. He, W. Wu, and D. Qi, "High quality participant recruitment in vehicle-based crowdsourcing using predictable mobility," in *Proc. IEEE INFOCOM*, Hong Kong, China, 2015, pp. 2542–2550.

[33] Y. Chen, B. Li, and Q. Zhang, "Incentivizing crowdsourcing systems with network effects," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, 2016, pp. 1–9.

[34] H. Jin, H. Guo, L. Su, K. Nahrstedt, and X. Wang, "Dynamic task pricing in multi-requester mobile crowd sensing with markov correlated equilibrium," in *Proc. IEEE INFOCOM*, Paris, France, 2019, pp. 1063–1071.

[35] J. Lin, M. Li, D. Yang, and G. Xue, "Sybil-proof online incentive mechanisms for crowdsensing," in *Proc. IEEE INFOCOM*, Honolulu, HI, USA, 2018, pp. 2438–2446.

[36] H. Zhang, B. Liu, H. Susanto, G. Xue, and T. Sun, "Incentive mechanism for proximity-based mobile crowd service systems," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, 2016, pp. 1–9.

[37] Z. Cai, Z. Duan, and W. Li, "Exploiting multi-dimensional task diversity in distributed auctions for mobile crowdsensing," *IEEE Transactions on Mobile Computing*, vol. 20, no. 8, pp. 2576–2591, 2021.

[38] M. Xiao, G. Gao, J. Wu, S. Zhang, and L. Huang, "Privacy-preserving user recruitment protocol for mobile crowdsensing," *IEEE/ACM Transactions on Networking*, vol. 28, no. 2, pp. 519–532, 2020.

[39] X. Zhang, Z. Yang, W. Sun, Y. Liu, S. Tang, K. Xing, and X. Mao, "Incentives for mobile crowd sensing: a survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 54–67, 2015.

[40] X. Ji, Y. He, J. Wang, W. Dong, X. Wu, and Y. Liu, "Walking down the stairs: Efficient collision resolution for wireless sensor networks," in *Proc. IEEE INFOCOM*, Toronto, Canada, 2014, pp. 961–969.

[41] X. Wang, W. Wu, and D. Qi, "Mobility-aware participant recruitment for vehicle-based mobile crowdsensing," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 5, pp. 4415–4426, 2018.

[42] X. Zhu, Y. Luo, A. Liu, W. Tang, and M. Z. A. Bhuiyan, "A deep learning-based mobile crowdsensing scheme by predicting vehicle mobility," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4648–4659, 2021.

[43] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "LightGCN: simplifying and powering graph convolution network for recommendation," in *Proc. ACM SIGIR*, Xi'an, China, 2020, pp. 639–648.

[44] Z. Duan, W. Li, X. Zheng, and Z. Cai, "Mutual-preference driven truthful auction mechanism in mobile crowdsensing," in *Proc. IEEE ICDCS*, Dallas, Texas, USA, 2019, pp. 1233–1242.

[45] D. Liu, Z. Cao, M. Hou, H. Rong, and H. Jiang, "Pushing the limits of transmission concurrency for low power wireless networks," *ACM Transactions on Sensor Networks*, vol. 16, no. 4, pp. 1–29, 2020.

[46] Z. Wang, Y. Zhang, H. Chen, Z. Li, and F. Xia, "Deep user modeling for content-based event recommendation in event-based social networks," in *Proc. IEEE INFOCOM*, Honolulu, HI, USA, 2018, pp. 1304–1312.

[47] J. L. Z. Cai, M. Yan, and Y. Li, "Using crowdsourced data in location-based social networks to explore influence maximization," in *Proc. IEEE INFOCOM*, San Francisco, CA, USA, 2016, pp. 1–9.

[48] Z. Xu, Z. Li, Q. Guan, D. Zhang, Q. Li, J. Nan, C. Liu, W. Bian, and J. Ye, "Large-scale order dispatch in on-demand ride-hailing platforms: a learning and planning approach," in *Proc. ACM KDD*, London, UK, 2018, pp. 905–913.