Incentivizing Vehicular Crowdsensing System for Large Scale Smart City Applications

Susu Xu[a], Xinlei Chen[b], Xidong Pi[a], Calee Joe-Wong[b], Pei Zhang[b], and Hae Young Noh[a]

[a]Carnegie Mellon University, Pittsburgh, PA, 15213
[b]Carnegie Mellon University, Moffett Field, CA, 94035

ABSTRACT

Mobile crowd sensing (MCS) enables many smart city applications (e.g., transportation monitoring/management, environmental monitoring, etc.). Recently, MCS systems built on non-dedicated vehicular platforms like taxis have become popular due to their large-scale coverage and low-cost deployment and maintenance. However, the goal of MCS may be inconsistent with the goal of vehicles. For example, MCS expects to get large and balanced sensing coverage over the city, while the taxis gather in busy areas to search for new ride requests. This inconsistency between the goals of MCS and vehicles results in a low sensing coverage and decreases the quality of the collected information.

To address this inconsistency and optimize the sensing coverage, this paper presents an incentivizing system to optimize the sensing coverage of the sampled data. Key challenges to resolving this inconsistency include limited budget constraining the ability to incentivize more vehicles and complicate vehicle and trajectory selection problem making it difficult to obtain the incentivizing strategy. To address these challenges, we design a customized incentive by combining monetary incentives and potential ride request at the destination to reduce the cost of incentivizing vehicles and utilize the budget efficiently. Meanwhile, we formulate the problem of incentivizing trajectory planning as a non-linear multiple-choice knapsack problem and propose a heuristic algorithm to approximate the optimal incentivizing strategy. The experiments based on the real-world data show that our system achieves up to 26.99% improvement in the sensing coverage compared to benchmark methods.

Keywords: Incentive Mechanism, Cyber Physical Systems, Smart City

1. INTRODUCTION

Vehicular mobile crowdsensing systems utilize sensors mounted on non-dedicated individual mobile vehicle agents (e.g. taxis, drones) to collect city-scale spatio-temporal information for smart city applications, including infrastructure,[1,2] environment,[3–6] and social applications.[7] Compared to conventional mobile sensing systems, vehicular mobile crowd sensing systems have the advantage of lower operation and maintenance cost and lower energy consumption. The rapid growth of smart city applications has brought diverse demand of various types of city-wide spatio-temporal data, such as air quality and traffic conditions, which also raise higher requirements on the quality of data collected by the vehicular mobile crowdsensing system.

One of the main goals of the crowdsensing system is to provide high-quality sensing data for further data analysis. The high-quality sensing data refers to the data with sufficient sensing coverage and high resolution and precision. High sensing coverage means that the collected sensing data distributes in a uniform/balanced manner across the time and spatial domain. Therefore, the crowdsensing system needs to optimize the sensing coverage balance level such that sufficient spatio-temporal information is contained in the collected sensing data.

However, there may exist inconsistencies between the goals of the crowdsensing system and the non-dedicated vehicles, which result in low sensing coverage. The vehicles, e.g., taxis, are non-dedicated sensing platforms. Their main goal is to find more ride requests to make profits instead of sensing data. Therefore, the vehicles may not distribute themselves as uniformly as what the crowdsensing system prefers. This inconsistency between the

Further author information: (Send correspondence to Susu Xu and Xinlei Chen)
Susu Xu: E-mail: susux@andrew.cmu.edu
Xinlei Chen: E-mail: xinlei.chen@west.cmu.edu

goals of vehicle and crowdsensing system may result in a lack of data collected in some areas of the city. In this way, the vehicular crowdsensing system systems cannot provide sufficient spatio-temporal information for further data analysis, which will significantly impact the performance of data-driven city monitoring and management. To address the inconsistency, people use either monetary or non-monetary rewards to incentivize vehicles to new trajectories[8–11] that yield better sensing coverage. But when there is a huge number of vehicles, the budget needs to be carefully allocated to improve coverage as much as possible. In addition, to achieve a good quality of sensing coverage with a limited budget, most previous works select a subset of vehicles to collect long-term sensory data in their current locations,[12–14] or select dispatching destinations for vehicles.[13,15] These methods mainly focus on the static distribution of sensors at a fixed time point but ignore the influences of the vehicles' dynamic mobility on the quality of sensing coverage.

To address the above inconsistency, our main goal is to efficiently incentivize vehicles to effectively utilize the limited budget to optimize the sensing coverage. Key challenges to achieving this objective include: 1) limited budget to incentivize more vehicles and 2) selection of vehicles and trajectories is a highly complex problem with multiple constraints. To overcome these challenges, our work has 2 main contributions: 1) we design a new incentive, which combines increasing potential ride requests along with the incentivized trajectories as non-monetary rewards for vehicles, to efficiently reduce the monetary cost of incentivizing, and 2) we formulate the vehicle and trajectory selection into an optimization problem combining the predictions of vehicle mobility and ride requests to optimize the incentivizing strategy.

2. SYSTEM OVERVIEW

A typical vehicular crowd sensing system has 2 key components: vehicle agents and a crowdsourcer. Vehicle agents refer to the non-dedicated vehicle platforms, e.g., taxis platform, which are mounted with various sensors and collect data along with their trajectories. Crowdsourcer is the system which collects data from vehicle agents. The collected spatio-temporal city-scale data may be further used for air quality monitor, traffic monitoring and other smart city applications. The crowdsourcer often has a number of budget for the data collection. Based on the budget and vehicle agents' status (e.g., location, occupancy status), the crowdsourcer incentivizes vehicle agents to different trajectories and organizes the data collected by vehicle agents along with their trajectories.

The incentivizing process includes 3 key modules: 1) vehicle mobility prediction, 2) customized incentive based on ride request prediction, and 3) multi-incentive optimization algorithm. Given a target sensing area, we divide the map of a city into $a \times b$ grids and the sensing time period into T time points, where T is the incentivizing period. The grid locations are denoted as (i, j), where $1 \leq i \leq a$, $1 \leq j \leq b$, and the time point is $1 \leq t \leq T$. Assume we have C vehicle agents in total, and that each vehicle agent $c \in C$ is pre-mounted with required sensors and assumed to run inside the $a \times b$ area. As Figure 1 shows, we firstly collect the location and occupancy status of each vehicle agent. Given the location information, we then predict the mobility probability of agent c appearing in location (i, j) at the time point t. Based on current timestamp, the ride request prediction model predicts the probability that vehicle c can obtain a ride requestin location (i, j) at the time point t, which is used to calculate its incentive later.

With the vehicle mobility probability and ride request probability in spatio-temporal domain, our multi-incentive algorithm is proposed to combines multiple types of incentives and obtain an optimal vehicle incentivizing strategy, including which non-occupied vehicle agents to select, which trajectories to assign to the selected vehicles, and customized incentive to pay to the selected vehicles.

The obtained incentivizing strategy is then delivered to each vehicle agent. The selected vehicle agents choose to accept the incentivizing strategy or not. If accepted, the selected vehicle agents have to run along with the assigned trajectory during the incentivizing period. If they refuse to accept, the vehicle agents run as usual. While vehicle agents are running inside the target area, the pre-mounted sensors automatically collect and upload the data to the crowdsourcer. Finally, the crowdsourcer organizes and delivers the collected data for further data analysis and start a new round of task.

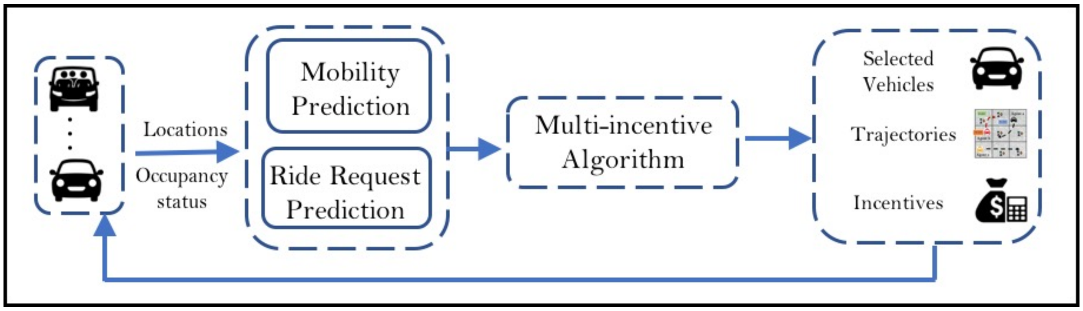


Figure 1. The overview of vehicle incentivizing system

3. OPTIMIZATION PROBLEM FORMULATION

In this section, we formulate the problem of incentivizing strategy planning to optimize the sensing coverage into an optimization problem and introduce a heuristic algorithm to solve the problem. We first model the sensing coverage using entropy of the spatio-temporal sensing data distribution. Then we design a new incentive customized to each vehicle's mobility and ride request probability, such that the monetary cost of incentivizing vehicle agents is reduced. Finally, we formulate the problem into a multiple-choice knapsack problem.

We first introduce a new problem formulation with a fine-grained objective function defining the quality of data coverage in a more precise way. Given i, j, t referring to longitude, latitude, and time point, with the collected sensing data distribution P and target uniform distribution, the quality of sensing distribution is defined as the entropy of the collected sensing data distribution:

$$H(P) = -\sum_{i,j,t} P(i,j,t) \log P(i,j,t).$$

Therefore, optimizing the sensing distribution is equivalent to maximizing the entropy $H(P)$. From the perspective of information theory, $H(P)$ measures the change of information when one revises beliefs from the prior uniform distribution to the posterior probability distribution P.[16]

To efficiently utilize the budget, we design a new customized incentive $B(c)$ that utilizes the increasing of ride request probability that vehicle c obtains by following our assigned trajectory D_c instead of refusing our assignment:

$$B(c) = \max(r_{max} - r_u \cdot Request(D_c^T), r_{min}) \tag{1}$$

where r_{min} and r_{max} are the minimum and maximum monetary incentive to actuate one taxi respectively. The $Request(D_c^T)$ represents the predicted ride request distribution at the location of the trajectory D_c at the ending time T of the incentivizing period. r_u is the unit monetary incentive for one ride request. Moreover, even with a high possibility to get new ride request, each taxi still needs a minimum monetary incentive to be motivated, which is r_{min}.

With the new defined objective function and incentive, we can formulate the problem. Given the target area with longitude of a, latitude of b, and incentivizing period T, I_c is a binary indicator of whether the vehicle c is incentivized, and k_c specifies the trajectory assigned to vehicle c, our system aims at optimizing the sensing coverage within budget and physical constraints as follows:

$$\min_{\substack{I_1,\cdots,I_C \\ k_1,\cdots,k_C}} \; -H(P) \tag{2}$$

$$\text{subject to} \; \sum_{c=1}^{C} B(c) \cdot I_c \leq R \tag{3}$$

$$D_c(i,j,t) \cdot I_c \in \{0,1\} \;. \tag{4}$$

$$D_c = D_c^{k_c} \text{ where } k_c \in \{0,1,\cdots,K_c\} \tag{5}$$

Note that $B(c)$ is determined by I_c and k_c. $k_c = 0$ is the probabilistic trajectory when c cruises without incentivizing or passengers, and $k_c > 0$ represents a deterministic incentivized trajectory. The Inequality 3 represents the budget constraint. The constraints 4 and 5 represent the set of vehicle potential trajectories, which has has a finite size due to the vehicle mobility's physical continuity.

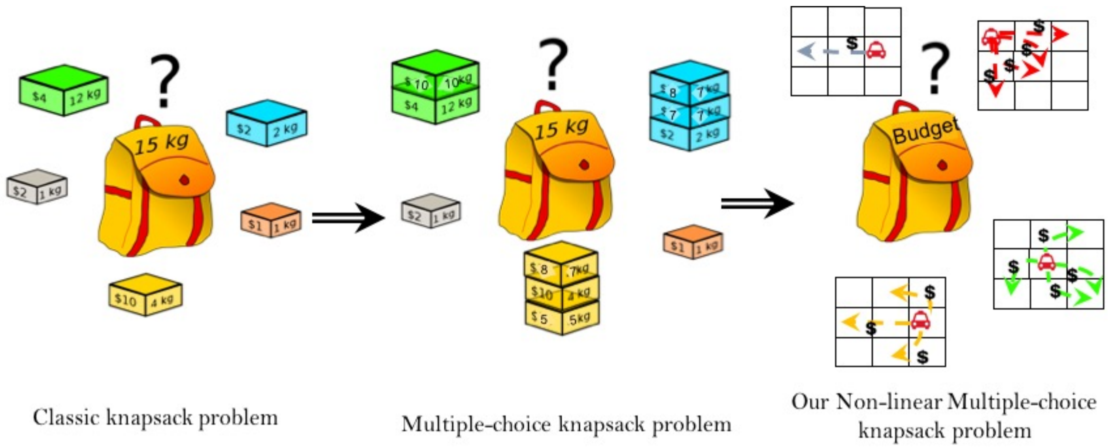


Figure 2. Formulated the optimization problem into multiple-choice knapsack problem.

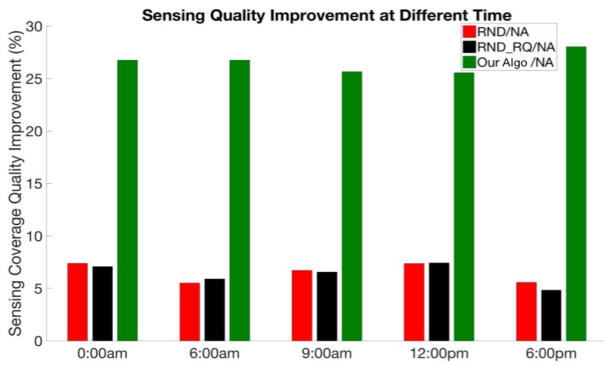
After simplification, the problem can be represented as a non-linear multiple choice knapsack problem.[17–19] As Figure 2 shows, in the formal definition of multiple choice knapsack problem, there are m classes $N_l, \cdots, N_m$ of items to be packed into a knapsack of capacity c. Each item $j \in N_i$ has a profit and a size W_{ij}, and the problem is to choose exactly one item from each class such that the objective is minimized/maximized without exceeding the capacity. In our problem, as Figure 2 shows, m classes of items represent C cars. Each item $j \in N_i$ represents a potential trajectory, which refers to 1) potential incentivizing trajectories or 2) run as usual. When an unoccupied vehicle agent is not actuated and runs randomly, it keeps sensing data, so "run as usual" is also one of the items for each vehicle agent in the knapsack formulation. To clarify, when one free vehicle agent is not actuated, it is not that there is no trajectory actuated for that vehicle agent or class. In contrast, the actuated trajectory is a "run as usual" trajectory. The capacity refers to the total budget. The minimized objective total profit represents the entropy. The knapsack problem in our scenario is to choose exactly one trajectory or random run for each unoccupied car, such that the budget does not exceed the total budget and the entropy is minimized.

To solve the formulated problem, we propose a heuristic algorithm combining vehicle mobility prediction and ride request prediction to approximate the optimal vehicle incentivizing strategy in polynomial time. The proposed algorithm is based on the idea of Complementary Constructive Procedure (CCP).[20] We first initialize a feasible solution S within the constraints until the budget is satisfied. Then we keep updating the solution to optimize the sensing coverage. With the initialized feasible solution, we firstly decompose the objective function in grid level by finding the time and location pair with a maximum number of vehicles passing through. Then we decompose the time-location pair with respect to each passing vehicle's trajectory. For vehicles belonging to the passing vehicle set, we calculate the contribution of incentivizing each passing vehicle to a new candidate trajectory on optimizing the objective function. Then we select the new candidate trajectory with the maximum contribution to optimize the objective function. Finally, the solution is updated based on the selected vehicles and trajectories. With multiple iterations, the solution keeps updating until the estimate of sensing coverage quality converges.

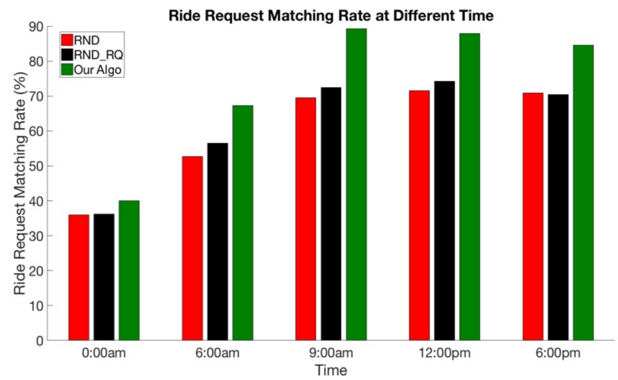
4. EXPERIMENTS AND RESULTS

To evaluate our algorithm, we use real-world taxi trajectories, including trips of $20,067$ taxis in one month in the city of Beijing, China. The evaluation area occupies a size of $15km$ by $15km$ and is discretized into a

10 × 10 map grid, in which each grid has the size of 1 × 1. To show the sensing performance improvement due to our algorithm, we used three other incentivizing methods as benchmark methods: *No Actuation (NA)* which does not incentivize any vehicle agents and just lets all agents run as usual, *Random Actuation (RND)*, which randomly selects vehicle agents and respective trajectories within the given budget using all monetary incentive, and *Random Actuation with Ride Request Prediction (RND_RQ)*, which uses our customized incentive combining non-monetary reward to randomly select vehicle agents and respective trajectories within the given budget.



(a)



(b)

Figure 3. (a) The sensing coverage improvement percentage of our algorithm, *RND* and *RND_RQ* with respect to *NA* at different times of the day; (b) The sride request matching rate of our algorithm, *RND* and *RND_RQ* with respect to *NA* at different times of the day

We first compare the sensing coverage improvement percentage of our algorithm, *RND*, and *RND_RQ* with respect to *NA* at different times of the day. As Figure 3(a) shows, our algorithm always outperforms the benchmark methods and achieves up to 26.99% improvement compared to benchmark methods. We can also find at around 12:00pm, the sensing coverage improvement is the least. This may be because most taxis tend to gather in the central business area, which makes it difficult to incentivize taxis to the sparse area. In addition, our system shows consistent improvement (25% - 27%) at different times of the day. This proves the robustness of our system on improving sensing coverage quality. Our system considered the stochastic mobility of all vehicle agents during planning, which makes it more robust to the unexpected trajectory changes of vehicle agents. Meanwhile, we also find that the performance of all methods is related to the initial vehicle distribution without any incentivizing. in our experiment, there are around 200 taxis occupied among 500 total taxis during incentivizing. These occupied taxis are not able to be incentivized, which means that the improvement is significantly limited by these occupied taxis, especially when they gather in the center area. Meanwhile, if the traffic condition and weather condition influence the sensing distribution under no incentivizing, our incentivizing performance will also be impacted in a similar way.

Figure 3(b) shows how ride request matching rate changes at different times of the day. The results show that our algorithm always achieves the highest ride request matching rate compared to the benchmark methods. This is because our customized incentive design considers the ride request probability increase at the destination as an incentive, which forces the optimization algorithm to select the destination with a higher ride request probability to reduce the monetary incentivizing cost. Meanwhile, the ride request matching rate is also decided by the spatial distribution of taxi and ride request at that time. Our system incentivizes more taxis than the benchmark methods and consistently keeps a higher ride request matching rate. The advantage comes from the robustness of our mobility prediction model and ride request prediction model.

5. CONCLUSION

This paper presents a system that incentivizes non-dedicated vehicular platforms like taxis to optimize the sensing coverage for mobile crowdsensing in large-scale smart city applications. The main challenge in achieving this goal is the limited budget for incentivizing. We design a new customized incentive by considering the ride request improvement as the non-monetary reward to utilize the budget in an efficient way. Then we formulate the problem as a multiple-choice knapsack problem and propose a multi-incentive algorithm to approximate the optimal incentivizing strategy. This algorithm integrates 1) a mobility prediction model that guides the selection of taxis to incentivize and 2) a ride request prediction model to help match ride requests with taxis, lower incentive cost, and improve taxi drivers' motivation. Extensive simulation and field experiments on taxi testbed show that our system can achieve up to 26.99% better sensing coverage quality and up to 20% more ride request matching rate than baselines.

REFERENCES

[1] Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., and Weaver, J., "Google street view: Capturing the world at street level," *Computer* **43**(6), 32–38 (2010).

[2] Li, X., Shu, W., Li, M., Huang, H.-Y., Luo, P.-E., and Wu, M.-Y., "Performance evaluation of vehicle-based mobile sensor networks for traffic monitoring," *IEEE transactions on vehicular technology* **58**(4), 1647–1653 (2009).

[3] Devarakonda, S., Sevusu, P., Liu, H., Liu, R., Iftode, L., and Nath, B., "Real-time air quality monitoring through mobile sensing in metropolitan areas," in [*Proceedings of the 2nd ACM SIGKDD international workshop on urban computing*], 15, ACM (2013).

[4] Zheng, Y., Liu, F., and Hsieh, H.-P., "U-air: When urban air quality inference meets big data," in [*Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*], 1436–1444, ACM (2013).

[5] Chen, X., Xu, X., Liu, X., Noh, H. Y., Zhang, L., and Zhang, P., "Hap: Fine-grained dynamic air pollution map reconstruction by hybrid adaptive particle filter," in [*Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*], 336–337, ACM (2016).

[6] Wang, Y., Liu, X., Wei, H., Forman, G., Chen, C., and Zhu, Y., "Crowdatlas: Self-updating maps for cloud and personal use," in [*Proceeding of the 11th annual international conference on Mobile systems, applications, and services*], 27–40, ACM (2013).

[7] Eisenman, S. B., Miluzzo, E., Lane, N. D., Peterson, R. A., Ahn, G.-S., and Campbell, A. T., "Bikenet: A mobile sensing system for cyclist experience mapping," *ACM Transactions on Sensor Networks (TOSN)* **6**(1), 6 (2009).

[8] Amanatidis, G., Birmpas, G., and Markakis, E., [*Coverage, Matching, and Beyond: New Results on Budgeted Mechanism Design*], Springer Berlin Heidelberg (2016).

[9] Zheng, Z., Wu, F., Gao, X., Zhu, H., Chen, G., and Tang, S., "A budget feasible incentive mechanism for weighted coverage maximization in mobile crowdsensing," *IEEE Transactions on Mobile Computing* **PP**(99), 1–1 (2017).

[10] Jaimes, L. G., Vergara-Laurens, I. J., and Raij, A., "A survey of incentive techniques for mobile crowd sensing," *IEEE Internet of Things Journal* **2**(5), 370–380 (2015).

[11] Ogie, R. I., "Adopting incentive mechanisms for large-scale participation in mobile crowdsensing: from literature review to a conceptual framework," *Human-centric Computing and Information Sciences* **6**(1), 24 (2016).

[12] Jaimes, L. G., Vergara-Laurens, I., and Labrador, M. A., "A location-based incentive mechanism for participatory sensing systems with budget constraints," in [*Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*], 103–108, IEEE (2012).

[13] Younis, M. and Akkaya, K., "Strategies and techniques for node placement in wireless sensor networks: A survey," *Ad Hoc Networks* **6**(4), 621–655 (2008).

[14] Liu, Y., Niu, J., and Liu, X., "Comprehensive tempo-spatial data collection in crowd sensing using a heterogeneous sensing vehicle selection method," *Personal and Ubiquitous Computing* **20**(3), 397–411 (2016).

[15] Song, Z., Liu, C. H., Wu, J., Ma, J., and Wang, W., "Qoi-aware multitask-oriented dynamic participant selection with budget constraints," *IEEE Transactions on Vehicular Technology* **63**(9), 4618–4632 (2014).

[16] Kullback, S. and Leibler, R. A., "On information and sufficiency," *The annals of mathematical statistics* **22**(1), 79–86 (1951).

[17] Dudziński, K. and Walukiewicz, S., "Exact methods for the knapsack problem and its generalizations," *European Journal of Operational Research* **28**(1), 3–21 (1987).

[18] Aho, I., "Interactive knapsacks," *Fundamenta Informaticae* **44**(1-2), 1–23 (2000).

[19] Bretthauer, K. M. and Shetty, B., "The nonlinear knapsack problem–algorithms and applications," *European Journal of Operational Research* **138**(3), 459–472 (2002).

[20] Hifi, M., Michrafy, M., and Sbihi, A., "A reactive local search-based algorithm for the multiple-choice multi-dimensional knapsack problem," *Computational Optimization and Applications* **33**(2-3), 271–285 (2006).