Taxi Dispatch with Real-Time Sensing Data in Metropolitan Areas — a Receding Horizon Control Approach*

Fei Miao
University of Pennsylvania
miaofei@seas.upenn.edu

Shan Lin
Stony Brook University
shan.x.lin@stonybrook.edu

Sirajum Munir
University of Virginia
sm7hr@virginia.edu

John A. Stankovic
University of Virginia
stankovic@cs.virginia.edu

Hua Huang
Stony Brook University
huanghua.yh@gmail.com

Desheng Zhang
University of Minnesota
zhang@cs.umn.edu

Tian He
University of Minnesota
tianhe@cs.umn.edu

George J. Pappas
University of Pennsylvania
pappasg@seas.upenn.edu

ABSTRACT

Traditional transportation systems in metropolitan areas often suffer from inefficiencies due to uncoordinated actions as system capacity and traffic demand change. With the pervasive deployment of networked sensors in modern vehicles, large amounts of information regarding traffic demand and system status can be collected in real-time. This information provides opportunities to perform various types of control and coordination for large scale intelligent transportation systems. In this paper, we present a novel receding horizon control (RHC) framework to dispatch taxis, which combines highly spatiotemporally correlated demand/supply models and real-time GPS location and occupancy information. The objectives include reducing taxi idle driving distance and matching spatiotemporal ratio between demand and supply for service quality. Moreover, our RHC framework is compatible with different predictive models and optimization problem formulations. This compatibility property allows us to model disruptive passenger demands and traffic conditions into a robust optimization problem. Extensive trace driven analysis with a real taxi data set from San Francisco shows that our solution reduces the average total idle distance by 52%, and reduces the total supply demand ratio error across the city by up to 45%.

Categories and Subject Descriptors

H.4 [**Information System Application**]: Miscellaneous; I.2.8 [**Problem Solving, Control Methods, and Search**]: Control Application, Transportation

*This work was supported by NSF grant numbers CNS-1239483, CNS-1239108, CNS-1239226.

General Terms

Algorithms, Design, Experimentation, Performance

Keywords

Intelligent Transportation System, Real-Time Taxi Dispatch, Receding Horizon Control, Mobility Pattern

1. INTRODUCTION

Sensing and wireless networking technologies are increasingly deployed in transportation systems, such as highway management, traffic light control, supply chains, and autonomous vehicles. These systems demonstrate significant safety and efficiency improvements over traditional systems. In modern taxi networks, real-time occupancy status and Global Positioning System (GPS) location of each taxi can be collected. This data set provides rich spatiotemporal information about passenger demand and their mobility patterns. Hence, real-time information provides opportunities to improve coordination of taxi networks, for higher service quality, and lower idle driving distance.

Traditional metropolitan taxi networks heavily rely on drivers' experience to identify passengers on streets and maximize individual profit. However, self-interested, uncoordinated behaviors of drivers cause spatiotemporal mismatch between supply and passenger demand. There are taxi dispatch services operated by large taxi companies. Most of these services dispatch taxis based on locality or greedy algorithms, such as sending the nearest vacant taxi to pick up a passenger [15], or first-come, first-served. Such approaches prioritize immediate customer satisfaction at the cost of global resource utilization and service fairness, because the potential cost of rebalancing the entire taxi supply network is not considered.

In this paper, we consider the following design challenge: optimizing for anticipated future idle driving cost and globally geographical service fairness, while fulfilling current, local demand. To accomplish such goal, a control framework is needed to incorporate both system models learned from historical data and real-time GPS information. To the best of the authors knowledge, no previous work has considered this problem. Zhang and Pavone design an optimal rebalancing method for autonomous cars, that considers both global service fairness and possible future costs [27], but idle driving distance and GPS information are not considered. Both costs of idle cruising and missing tasks are included to assign trucks in the temporal perspective in [25], but real-time location information is not involved in the assignment.

To utilize large-scale real-time information of the taxi network, we consider a computational efficient, moving time horizon framework. Moreover, the dispatch solutions need to consider future costs of balancing supply demand ratio under practical constraints. Thus, we take a receding horizon control (RHC) approach to dynamically dispatch taxis in large-scale networks. General learning techniques are applied to historical monitoring data sets, to characterize passenger mobility patterns and demand models [26], which then provide demand predicting model besides current bookings in the system. Real-time GPS and occupancy information is collected to update supply and demand information for future dispatch. Our design aims to regulate the mobility of idle taxis for high performance large-scale transportation management.

The contributions of this work are as follows,

- To the best of our knowledge, we are the first to design an RHC framework for large-scale taxi dispatching. Compared with current local greedy algorithms, the novel framework allows us to consider both current and future requests, saving costs under constraints by involving expected future costs for re-balancing supply.
- The framework combines large-scale data in real-time control. Sensing data is used to build predictive passenger demand, taxi mobility models, and serve as real-time feedback for RHC.
- Extensive trace driven analysis based on a San Francisco taxi data set shows that our approach reduces average estimated taxi network idle distance by 52% as in Figure 5, and the total supply demand ratio error of all regions by 45% as in Figure 6, compared to the actual historical taxi system performance.
- Spatiotemporal context information such as disruptive passenger demand is incorporated into our control framework. This allows our control solutions to be more robust and accurate to such disturbances under uncertain contexts as shown in Figure 8.

The rest of the paper is organized as follows. More related work is introduced in Section 2. The structure of taxi monitoring system and control problems are introduced in Section 3. The taxi dispatch problem is formulated, followed by an RHC framework that interacting between historical model and real-time information in Section 4. A case study to evaluation the RHC framework with a real data set is shown in Section 5. Concluding remarks are provided in Section 6.

2. STATE-OF-THE-ART

There are three categories of research topics related to our work: taxi dispatch systems, transportation system modeling, and multi-agent coordination and control.

A number of recent works study taxi dispatching services along with the pervasive deployment of GPS in mordern taxis. Authors of [23] focus on minimizing total customer waiting time by concurrent dispatching multiple taxis and allowing taxis to exchange their booking assignments. In [22, 12, 21], authors aim to maximize drivers' profits by providing routing recommendations. These works give valuable results, but they only consider the current passenger requests and available taxis. Our design uses receding horizon control to consider both current and predicted future requests.

Various mobility and vehicular network modeling techniques have been proposed for transportation systems [5, 4]. Researchers have developed methods to predict travel time [7, 10], traveling speed [3], and charaterize taxi performance features [14]. These works provide insights to transportation system properties and suggest potential enhancement on transportation system performance. Our design takes a step further to develop dispatch methods based on available predictive data analysis.

There is a large number of works on mobility coordination and control. Different from taxi services, these works usually focus on region partition and coverage control so that coordinated agents can perform tasks in their specified regions [6, 1, 11]. Other related works include dynamic vehicle routing problems [2] and robust traffic flow management under uncertainty [24]. Their task models and design objectives are different from taxi dispatching problem. Also, model predictive control has been widely applied for process control, task scheduling, cruise control, and multi-agent transportation networks [16, 17, 13]. These works provide solid results for related mobility scheduling and control problems. However, none of these works incorporates both the real-time sensing data and historical mobility patterns into a receding horizon control design, leveraging the taxi supply based on the spatiotemporal dynamics of passenger demand.

3. TAXI DISPATCH PROBLEM: MOTIVATION AND SYSTEM

Taxi networks provide a primary transportation service in modern cities. Most street taxis respond to passengers' requests on their paths, and take passengers to their specified destinations. Therefore, existing taxi networks rely on drivers to drive around and arbitrarily pick up passengers on streets. This service model has successfully served up to 25% public passengers in metropolitan areas, such as San Francisco and New York [9, 18]. In existing taxi networks, a couple of key dynamics affect their service quality: a) dynamic passenger demand. The spatiotemporal patterns of demand include both regular factors, such as rush hours and busy areas, and irregular ones, such as weather, traffic, holiday schedule, major events, etc. b) dynamic taxi supply. Taxis have different mobility patterns, since each driver has his/her own working schedule and cruising areas. From the perspective of system performance, balancing spatiotemporal taxi supply across the whole city is a design requirement, similar to the idea of balancing server node utilization in [27]. On the other hand, idle driving introduces a direct cost, hence, we also consider another system-level objective — to reduce total idle driving cost instead of individual driver's profit.

Existing infrastructures serve as the basis of our design. Taxi companies in metropolitan areas already monitor the performance of the taxi network in real time. Figure 1 shows a typical monitoring infrastructure, which consists of a large number of geographically distributed sensing and communication components in each

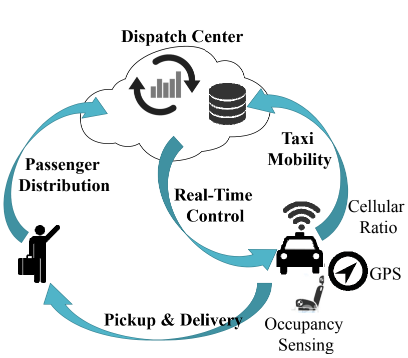


Figure 1: A prototype of the taxi dispatch system

taxi and a data center. In particular, the sensing components include a GPS unit and a trip recorder. Each taxi automatically reports its GPS location and occupancy status to the data center via cellular radio. The data center collects and stores data. Then, the monitoring center sends dispatch commands to taxi drivers also via cellular radio, which notify drivers over the speaker or on a special display. We assume that the geographical coordinates of every taxi are available in real time.

Given both stored historical data and the real-time taxi monitoring information described above, the process of dispatching taxis includes two phases: analysis of historical data and real-time computation of dispatch solution. This paper focuses on phase two; more precisely, we provide a scalable control framework that dispatches vacant taxis to balance current and future demand with small idle driving distance, by utilizing both historical data and real-time monitoring data. It is worth noting that heading to the allocated position is part of idle driving distance for a vacant taxi, which introduces a trade-off between the two design objectives. Hence, one design constraint we consider is to match global supply and demand without introducing large idle driving cost.

4. ALGORITHM DESIGN

4.1 Taxi Dispatch Problem Formulation

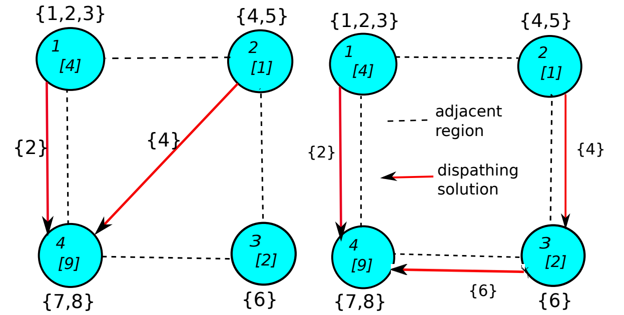
Informally, the goal of our taxi dispatch system is to schedule vacant taxis towards current and future passengers with least total idle mileage. Based on the spatiotemporal patterns of passenger demands in the city, the dispatch center dynamically allocates vacant taxis to different regions in order to match the passenger demands. We use supply demand ratio of different regions as a measure of service quality, since sending more taxis for more requests is a natural system-level requirement, to make customers of different positions equally satisfied. A similar service metric of service node utilization rate has been proposed in [27].

To calculate a dispatch solution, the system is equipped with model learning techniques to predict spatiotemporal patterns of passenger demand, either purely from history data, or combining real-time information of the taxi network. The real-time information includes each taxi's GPS location and occupancy status with a time stamp that periodically reported to the dispatch center.

Besides balancing supply and demand, another design requirement is to include future cost when calculating the current dispatch solution. It is difficult to perfectly predict the future of the large-scale taxi service system in practice, hence, we use a heuristic future idle driving distance to describe anticipated future cost associated to meeting customer requests. Considering control objectives and computational efficiency, we choose a receding horizon approach. We assume that the optimization time horizon is T, indexed by $k = 1, \dots, T$.

4.1.1 Supply and demand in taxi dispatch

With a large amount of historical data on taxi GPS and occupancy status, we extract basic dynamic demand information, such as demand during rush hours and in busy areas. We assume that the entire area of a city is divided into n regions according to some specific method. For example, a city can be divided into administrative sub-districts. We also assume that during a time slot k, the total number of requests we want to serve by current vacant taxis at the j-th region is denoted by r_j^k, the total number of requests in the entire city is denoted by $R^k = \sum_{j=1}^{j=n} r_j^k$, and define a request vector as $r^k \in \mathbb{R}^{1\times n} : [r_1^k, \dots, r_n^k]$. These are the demands we want to meet during time $k = 1, \dots, T$ with minimal idle driving



(a) A dispatch solution – taxi 2 goes to region 4, and taxi 4 goes to region 4.

(b) A dispatch solution – taxi 2 goes to region 4, taxi 4 goes to region 3, and taxi 6 goes to region 4.

Figure 2: Unbalanced supply and demand at different regions before dispatching and possible dispatch solutions. A circle represents a region, with a number of predicted requests ($[\cdot]$ inside the circle) and vacant taxis ($\{$ taxi IDs $\}$ outside the circle) before dispatching. A black dash edge means adjacent regions. A red edge with a taxi ID means sending the corresponding vacant taxi to the pointed region, to balance supply demand ratio.

cost.

Given the latest occupancy status information, we assume there are total N vacant taxis in the entire city that can be dispatched. The initial supply information consists of real-time GPS position for the i-th vacant taxi, denoted by $P_i^0 \in \mathbb{R}^{1\times 2}, i = 1, \dots, N$, and the initial position matrix for all available taxis is denoted by $P^0 \in \mathbb{R}^{N\times 2}$ (How does information of occupied taxis affect the supply/demand model will be discussed in subsection 4.2).

The basic idea of the dispatching problem is illustrated in Figure 2. Specifically, each region has a predicted number of requests the dispatch system needs to meet, and vacant taxis with IDs at different locations given real-time sensing information. The supply demand ratio at each region before dispatching is unbalanced. We want to find a dispatch solution in order to balance the supply demand ratio, while satisfying practical constraints and not introducing large idle driving cost for current and future time. Once the taxi reaches the target location to pick up passengers, the dispatch system will wait until the next dispatch period.

4.1.2 Variables, constraints and a cost function

With the above initial information about supply and demand, to calculate a dispatch decision at the region level for current vacant taxis, we define the dispatch order matrix $X^k \in \{0,1\}^{N\times n}$ as a binary variable matrix, satisfying that $X_{ij}^k = 1$ if and only if the i-th taxi is sent to the j-th region during time k. Then the constraint

$$X^k \mathbf{1}_n = \mathbf{1}_N, \quad k = 1, \dots, T,$$

must be satisfied, since every taxi should be dispatched to one region at time k, where $\mathbf{1}_N$ be a length N column vector of all 1s.

A routing process for a taxi usually needs a starting position and a destination position first, and then the driver follows the path shown on the GPS unit in the taxi. Since the routing process design is not the focus of this work, the dispatch center can simply sends a two dimensional GPS location for the taxi driver as destination. In practice there are taxi stations on the road in a metropolitan area, and each taxi has a preferred station or is randomly assigned one by the monitoring system at every region. Denote the preferred geometry location matrix for the i-th taxi by $W_i \in \mathbb{R}^{n\times 2}$, and $[W_i]_j$ – the j-th row of W_i is the dispatch position sent to the i-th taxi when $X_{ij}^k = 1$. It is equivalent to send the following location vector to the i-th vacant taxi:

$$X_i^k W_i = \sum_{q\neq j} X_{iq}^k [W_i]_q + X_{ij}^k [W_i]_j = [W_i]_j \in \mathbb{R}^{1\times 2},$$

Parameters	Description
N	the total number of vacant taxis
n	the number of regions
$r^k \in \mathbb{R}^{1\times n}$	the total number of predicted requests to be served by current vacant taxis at each region
$C^k \in [0,1]^{n\times n}$	matrix that describes taxi mobility patterns during one time slot
$P^0 \in \mathbb{R}^{N\times 2}$	the initial positions of vacant taxis provided by GPS data
$W_i \in \mathbb{R}^{n\times 2}$	preferred positions of the i-th taxi at n regions
$\alpha \in \mathbb{R}^N$	the upper bound of distance each taxi can drive for balancing the supply
$\beta \in \mathbb{R}_+$	the weight factor of the objective function
$R^k \in \mathbb{R}_+$	total number of predicted requests in the city
Variables	Description
$X^k \in \{0,1\}^{N\times n}$	the dispatch order matrix that represents the region each vacant taxi should go
$d_i^k \in \mathbb{R}_+$	heuristic idle driving distance of the i-th taxi for reaching the suggested location

Table 1: Parameters and variables of the RHC problem (8)**.**

since $X_{iq}^k = 0$, $X_{iq}^k W_i = [0\ 0]$ for $q \neq j$. W_i does not need to change with time k and can be set up by the dispatch system before calculating the dispatch solution.

Cost for violating service fairness: One design requirement is to fairly serve the requests across the entire city. An intuitive measurement of whether demand is matched at every region is: error$=\sum_{j=1}^n |s_j^k - r_j^k|$,
where s_j^k is the total number of vacant taxis sent to the j-th region. However, even supply is already fairly allocated to each region, this error can still be large, since the total number of vacant taxis and requests are different. Therefore, to measure how supply matches demand at different regions, we use the measure metric—supply demand ratio. When the supply demand ratio of every region equals to that of the whole city, we have the following equation for $j = 1, \dots, n$, $k = 1, \dots, T$,

$$\frac{\mathbf{1}_N^T X_{\cdot j}^k}{r_j^k} = \frac{N}{R^k}, \iff \frac{\mathbf{1}_N^T X_{\cdot j}^k}{N} = \frac{r_j^k}{R^k}, \tag{1}$$

where $X_{\cdot j}^k$ is the j-th column of matrix X^k. For expression convenience, equation (1) is equivalent to the following equation about two row vectors

$$\frac{1}{N}\mathbf{1}_N^T X^k = \frac{1}{R^k} r^k, \quad k = 1, \cdots, T. \tag{2}$$

But equation (2) is too strict to be a constraint, since variables are integers, and taxis' driving speed is limited that they may not be able to reach far away for serving some request in time k, and there may be no feasible solutions satisfying (2). Thus, we define the total supply demand error cost J_E to measure how much the global supply demand ratio is violated at all regions

$$J_E = \sum_{k=1}^T \left\| \frac{1}{N}\mathbf{1}_N^T X^k - \frac{1}{R^k} r^k \right\|_1. \tag{3}$$

Estimated idle driving distance associated with meeting the dispatch solutions: Traversing from position P_i^0 to position $X_i^1 W_i$ for service will introduce cost since the taxi drives on the road without picking up a passenger in order to serve a target region. Hence, we consider to minimize this kind of idle driving distance while dispatching taxis. Driving in a city is approximated as traveling on a grid road. To estimate the distance without knowing the exact path, we use the Manhattan norm or one norm between two geometry positions, which is widely applied as a heuristic cost in path planning algorithms [19]. We define $d_i^k \in \mathbb{R}$ as the estimated idle driving distance of the i-th taxi for reaching the dispatched location $X_i^k W_i$, then the distance associated with going to $X_i^1 W_i$ is approximated by

$$d_i^1 = \|P_i^0 - X_i^1 W_i\|_1, \quad i = 1, \dots, N.$$

To consider possible future costs in the current dispatch solution as re-balancing costs, we need to estimate the driving distance to reach $X_i^k W_i$ from the ending position of time $k-1$ for serving requests described by r^k. However, during time $k-1$, taxis mobility pattern is related to the pick-up and drop-off locations of passengers, which is not controlled by the dispatch system. So we assume the estimated ending position during $k-1$ is approximated as a linear function of the starting position $X_i^{k-1} W_i$ provided by the mobility pattern model, denoted by:

$$\mathbb{E}P_i^{k-1} = f(X_i^{k-1} W_i), f: \mathbb{R}^{1\times 2} \to \mathbb{R}^{1\times 2}. \tag{4}$$

For instance, when taxi's mobility pattern during time slot k is described by a matrix $C^k \in \mathbb{R}^{n\times n}$ satisfying $\sum_{j=1}^n C_{ij} = 1$, where C_{ij}^k is the probability that a taxi drops off a passenger at region j near the end of time k when the trace starts from region i. Then the process to get a heuristic d_i^k is illustrated in Figure 3.

Remark 1. *When road congestion information is available to the dispatch system, function in* (4) *can be generalized to include real-time congestion information. For instance, there is a high probability that a taxi stays in the same region for time k under congestion.*

It is worth noting that we do not assume the information of passenger's destination is available to the system for future time slots $k = 1, \dots, T$, since many passengers just wait for service at taxi stations instead of reserving one in advance in metropolitan areas. When destination and travel time of all trips are provided to the dispatch center via additional software or device for time k, the information is considered as a constant matrix P^k in problem (8)*, instead of the function defined in* (4).

Given X_i^{k-1} and the mobility pattern matrix $C^{k-1} \in [0,1]^{n\times n}$, the probability of ending at each region for taxi i is
$p = \sum_{j=1}^n [C^{k-1}]_j I(X_{ij}^{k-1} = 1) = X_i^{k-1} C^{k-1} \in \mathbb{R}^{1\times n}$,
where the indicator function $I(X_{ij}^{k-1} = 1) = 1$ if and only if

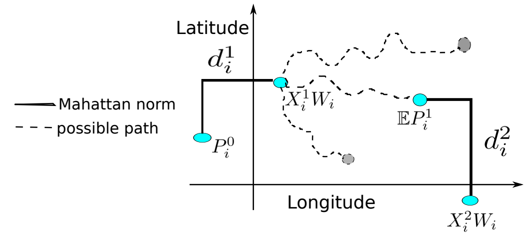


Figure 3: Illustration of a heuristic future driving distance for $k = 2$: predict ending location $\mathbb{E}P_i^1$ in (5)**, get a heuristic distance to location $X_i^2 W_i$ denoted by d_i^2 in** (6)**.**

$X_{ij}^{k-1} = 1$, and $[C^{k-1}]_j$ is the j-th row of C^{k-1}. We take the expected value of the i-th taxi by the end of time $k-1$ at each dimension as a predicted ending location of the i-th taxi

$$\mathbb{E}P_i^{k-1} = \sum_{j=1}^{n} p_j [W_i]_j = pW_i = X_i^{k-1} C^{k-1} W_i \in \mathbb{R}^{1\times 2}. \quad (5)$$

If we send a taxi to a target location $X_i^k W_i$ at the beginning of time k from position $\mathbb{E}P_i^{k-1}$, the approximated driving distance is

$$d_i^k = \|(X_i^{k-1}C^{k-1} - X_i^k)W_i\|_1,\ k = 2, \cdots, T,\ i = 1, \dots, N. \quad (6)$$

Here d_i^k is only a heuristic distance function of X_i^{k-1} and X_i^k, since the estimation accuracy depends on the predicting accuracy of the mobility pattern model. We use the expected ending position (not expected distance, this is a non-convex function of X_i^k) to avoid a random variable distance d_i^k, since a random variable will result in a stochastic programming that computationally expensive for a large-scale optimization problem. The distance d^1 is based on real-time location information P^0, and we use heuristic $d^2, \dots, d^T$ to measure possible future costs for meeting requests. We consider the effect of an inaccurate model when choosing T, which is discussed in the Remark, and experiment result is shown in Section 5.

The distance every taxi can drive should be bounded during limited time, and this distance upper bound for every taxi is denoted by a vector $\alpha \in \mathbb{R}^N$ any d^k should satisfy

$$d^k \leqslant \alpha.$$

Total idle driving distance to satisfy service fairness is denoted by

$$J_D = \sum_{k=1}^{T} \sum_{i=1}^{N} d_i^k \quad (7)$$

4.1.3 An RHC problem formulation

Since there exists a trade-off between two objectives as discussed in Section 3, we define a weight parameter β when summing up the costs related to both objectives. To summarize, we formulate the following problem (8) based on the definitions of variables, parameters, constraints and objective function

$$\begin{aligned}
\underset{X^k, d^k}{\text{minimize}} \quad & J = J_E + \beta J_D \\
& = \sum_{k=1}^{T} \left(\left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{R^k} r^k \right\|_1 + \beta \sum_{i=1}^{N} d_i^k \right) \\
\text{subject to} \quad & d_i^1 = \|P_i^0 - X_i^1 W_i\|_1, \quad i = 1, \dots, N, \\
& d_i^k = \|f^k(X_i^{k-1} W_i) - X_i^k W_i\|_1, \\
& i = 1, \dots, N, \quad k = 2, \dots, T, \\
& d^k \leqslant \alpha, \quad k = 1, 2, \dots, T, \\
& X^k \mathbf{1}_n = \mathbf{1}_N, \quad k = 1, 2, \dots, T, \\
& X_{ij}^k \in \{0, 1\}, i, j \in \{1, 2, \dots, N\}.
\end{aligned} \quad (8)$$

A list of parameters and variables is shown in Table 1. In particular, if transition probability $C^k, k = 1, \dots, T$ is learned by the modeling technique, $d_i^k = \|(X_i^{k-1}C^{k-1} - X_i^k)W_i\|_1$. When mixed integer programming is not efficient enough for a large-scale taxi network regarding to the problem size, one relaxation method is replacing the constraint $X_{ij}^k \in \{0, 1\}, \forall k, i, j$ by

$$0 \leq X_{ij}^k \leq 1,$$

With this approximation, every element of X^k is not restricted to a binary variable. After getting an optimal solution X^1 of relaxed form (8), for the i-th taxi, set the largest value of X_i^1 to 1, and the others to 0. This may violate the constraint of d_i^0, however, we set a conservative upper bound and it does not effect the dispatch effects in the experiment. The computational complexity of the relaxation form of (8) is polynomial of $V_n = nNT$.

Remark 2. *Previous work has developed multiple ways to learn passenger demand and taxi mobility patterns [3, 7, 12], and accuracy of the predicted model will affect the results of dispatch solutions. With perfect knowledge of customer demand and taxi mobility models, we can set a large time horizon to consider future costs in the long run. However, in practice we do not have perfect predictions, thus a large time horizon may amplify the prediction error over time. With an approximated mobility pattern matrix C^k, the dispatch solution with large T is even worse than small T, as shown in Figure 12. Applying real-time information to adjust taxi supply is a remedy to this problem. Formulation* (8) *is one computationally efficient approach to describe the design requirements. Modeling techniques are beyond the scope of this work, and dispatch methods considering effects of different modeling approaches is a future work.*

4.1.4 A robust RHC problem

One advantage of the formulation (8) is its flexibility to adjust the constraints and objective function according to different conditions. With prior knowledge of scheduled events that disturb the demand or mobility pattern of taxis, we take the effects of the events into consideration by setting uncertainty parameters. For instance, when we have basic knowledge that total demand in the city during time k is about $\tilde{R}^k$, but each region r_j^k belongs to some uncertainty set, denoted by an entry wise inequality $R_1^k \preceq r^k \preceq R_2^k$, given $R_1^k \in \mathbb{R}^n, R_2^k \in \mathbb{R}^n$. Then $r_j^k \in [R_{1j}^k, R_{2j}^k]$, $j = 1, \dots, n$, and is an uncertainty parameter instead of a fixed one as in problem (8). If we ignore the change of $\tilde{R}$ for different r^k and fix it on the denominator, by adding an uncertain range space for r^k, we have a robust optimization problem (9)

$$\begin{aligned}
\underset{X^k, d^k}{\text{min.}} \ \max_{R_1^k \preceq r^k \preceq R_2^k} \quad & J = \sum_{k=1}^{T} \left(\left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{\tilde{R}^k} r^k \right\|_1 + \beta \sum_{i=1}^{N} d_i^k \right) \\
\text{subject to} \quad & \text{constraints of problem (8).}
\end{aligned} \quad (9)$$

The robust optimization problem (9) is solvable in real-time, and we have the following Lemma 1

Lemma 1. *The robust RHC problem* (9) *can be solved exactly as a deterministic optimization problem.*

PROOF. In the objective function, only the first term is related to r^k. To avoid the maximize expression over an uncertain r^k, we first optimize the term over r^k for any fixed X^k. Let $X_{\cdot j}^k$ represent the j-th column of X^k, then

$$\begin{aligned}
& \max_{R_1^k \preceq r^k \preceq R_2^k} \left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{\tilde{R}^k} r^k \right\|_1 \\
= & \max_{R_1^k \preceq r^k \preceq R_2^k} \sum_{j=1}^{n} \left| \frac{1}{N} \mathbf{1}_N^T X_{\cdot j}^k - \frac{r_j^k}{\tilde{R}^k} \right| \\
= & \sum_{j=1}^{n} \max_{r_j^k \in [R_{1j}^k, R_{2j}^k]} \left| \frac{1}{N} \mathbf{1}_N^T X_{\cdot j}^k - \frac{r_j^k}{\tilde{R}^k} \right|.
\end{aligned}$$

The second equality is true because we can optimize each r_j^k separately in this equation. For $R_{1j}^k \le r_j^k \le R_{2j}^k$, we have

$$\frac{R_{1j}^k}{\tilde{R}^k} \le \frac{r_j^k}{\tilde{R}^k} \le \frac{R_{2j}^k}{\tilde{R}^k}.$$

Then the problem is to maximize each absolute value for $j = 1, \dots, n$. Consider the following problem for $x, a, b \in \mathbb{R}$ to examine the character of maximization problem over an absolute value:

$$\begin{aligned}\max_{x_0 \in [a,b]} |x - x_0| = & \begin{cases} |x-a|, & \text{if } x > (a+b)/2 \\ |x-b|, & \text{otherwise} \end{cases} \\ =& \max\{|x-a|, |x-b|\} \\ =& \max\{x-a, a-x, x-b, b-x\}.\end{aligned}$$

Similarly, for the problem related to r_j^k, we have

$$\begin{aligned}\max_{r_j^k \in [R_{1j}^k, R_{2j}^k]} \left| \frac{\mathbf{1}_N X_{\cdot j}^k}{N} - \frac{r_j^k}{\tilde{R}^k} \right| =& \max\{ \left| \frac{\mathbf{1}_N X_{\cdot j}^k}{N} - \frac{R_{1j}^k}{\tilde{R}^k} \right|, \\ & \left| \frac{\mathbf{1}_N X_{\cdot j}^k}{N} - \frac{R_{2j}^k}{\tilde{R}^k} \right| \}.\end{aligned} \tag{10}$$

Thus, with slack variables $t^k \in \mathbb{R}^n$, we re-formulate the robust RHC problem as

$$\begin{aligned}\min_{X^k, d^k, t^k} \quad & J' = \sum_{k=1}^{T} (\sum_{j=1}^{n} t_j^k + \beta \sum_{i=1}^{N} d_i^k) \\ \text{subject to} \quad & t_j^k \ge \frac{\mathbf{1}_N X_{\cdot j}^k}{N} - \frac{R_{1j}^k}{\tilde{R}^k}, t_j^k \ge \frac{R_{1j}^k}{\tilde{R}^k} - \frac{\mathbf{1}_N X_{\cdot j}^k}{N}, \\ & t_j^k \ge \frac{\mathbf{1}_N X_{\cdot j}^k}{N} - \frac{R_{2j}^k}{\tilde{R}^k}, t_j^k \ge \frac{R_{2j}^k}{\tilde{R}^k} - \frac{\mathbf{1}_N X_{\cdot j}^k}{N}, \\ & j = 1, \dots, n, \quad , k = 1, \dots, T, \\ & \text{constraint of problem (8).}\end{aligned}$$

Then we reduce the robust RHC problem to a deterministic optimization problem.

Taxi mobility patterns during disruptive events are not easily estimated (in general), while we have knowledge such as a rough number of people are taking part in a conference or competition, or even more customer reservations because of events in the future. By introducing extra knowledge besides historical data model, the dispatching system responds to such disturbances faster than the situation without a robust optimization. Comparison of results of (1) and problem (8) is shown in Section 5.

4.2 RHC Framework Design

Demand and taxi mobility patterns can be learned from historical data, but they are not sufficient to calculate a dispatch solution with dynamic positions of taxis. Hence, we design an RHC framework to adjust dispatch solutions and incorporate historical model with real-time sensing information. Real-time GPS and occupancy information then act as feedback by providing latest taxi locations, and demand-predicting information for an online learning method like [26]. Formulation (8) or (9) are embedded in one iteration of the algorithm, to provide dispatching solutions.

Though we get a sequence of solutions in T steps– $X^1, \dots, X^T$, we only send recommendations to vacant taxis according to X^1.We summarize the complete process of dispatching taxis with both historical and real-time data as Algorithm 1, followed by a detail computational process of each iteration.

Algorithm 1: RHC Algorithm for real-time taxi dispatch

Inputs: Time slot length t_1 minutes, period of sending dispatch solutions t_2 minutes (t_1/t_2 is an integer); a preferred station location table for every taxi in the network; estimated request vectors $\hat{r}(h_1), h_1 = 1, \dots, 1440/t_1$, mobility patterns $\hat{f}(h_2), h_2 = 1, \dots, 1440/t_2$; parameters of problems (8), (9): prediction horizon $T \ge 1$, β, and α.
Initialization: The predicted requests vector $r = \hat{r}(h_1)$ for corresponding algorithm start time h_1.
while *At the beginning of each t_2 time slot* **do**
(1). Update sensor information for initial positions of vacant taxis P^0 and occupied taxis P'^0, total number of vacant taxis N, preferred dispatch location matrices W_i.
if *time is the beginning of an h_1 time slot* **then**
Calculate $\hat{r}(h_1)$ if the system applies an online training method; count total number of occupied taxis $n_o(h_1)$; update vector r.
end
(2). Update the demand vectors and mobility functions r^k, $f^k(\cdot)$(for example, C^k), $k = 1, 2, \dots, T$.
(3). **if** *there is priority knowledge of disruptive events such that r^k is in an uncertainty set* **then**
solve relaxed form of problem (9);
else
solve relaxed form of problem (8);
end
(4). Send dispatch orders to vacant taxis according to the optimal solution of matrix X^1. Let $h_2 = h_2 + 1$.
end
Return: Stored sensor data and dispatch solutions.

Remark 3. *Predicted values of requests $\hat{r}(h_1)$ depend on the modeling method of the dispatch system. For instance, if the system only applies historical data set to learn each $\hat{r}(h_1)$, $\hat{r}(h_1)$ is not updated with real-time sensing data. When the system applies online training method such as [26] to update $\hat{r}(h_1)$ for each h_1, values of r, r^k are calculated based on the real-time value of $\hat{r}(h_1)$.*

4.2.1 Update r

We receive sensing data of both occupied and vacant taxis in real-time. Predicted requests that vacant taxis should serve during h_1 is re-estimated at the beginning of each h_1 time. To approximate the service capability when an occupied taxi turns into vacant during time h_1, we define the total number of drop off events at different regions as a vector $dp(h_1) \in \mathbb{R}^{n \times 1}$. Given $dp(h_1)$, the probability that a drop off event happens at region j is

$$pd_j(h_1) = dp_j(h_1)/\mathbf{1}_n^T dp(h_1), \tag{11}$$

where $dp_j(h_1)$ is the number of drop off events at region j during h_1. We assume that an occupied taxi will pick up at least one passenger around after turning vacant, and approximate future service ability of occupied taxis at region j during time h_1 as

$$r_{oj}(h_1) = \lceil pd_j(h_1) \times n_o(h_1) \rceil, \tag{12}$$

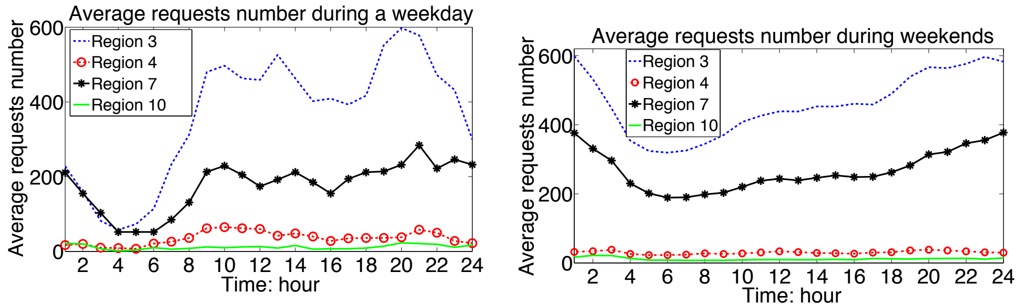
where $\lceil \cdot \rceil$ is a ceil function, $n_o(h_1)$ is the total number of current occupied taxis at the beginning of time slot h_1, provided by real-time sensor information of occupied taxis. Let

$$r = \hat{r}(h_1) - r_o(h_1),$$

then the estimated service capability of occupied taxis is deducted from r for time slot h_1.

Taxicab GPS Data set				Format		
Collection Period	Number of Taxis	Data Size	Record Number	ID	Status	Direction
05/17/08-06/10/08	500	$90MB$	$1,000,000$	Date and Time	Speed	GPS Coordinates

Table 2: San Francisco Data in the Evaluation Section. Giant baseball game in AT&T park on May 31, 2008 is a disruptive event we use for evaluating the robust optimization formulation.



(a) Requests during weekdays (b) Requests during weekends

Figure 4: Requests at different hours during weekdays and weekends, for four selected regions. A given historical data set provides basic spatiotemporal information about customer demands, which we utilize with real-time data to dispatch taxis.

4.2.2 Update r^k

We assume requests are uniformly distributed during h_1. Then for each time k of length t_2, if the corresponding physical time is still in the current h_1 time slot, the request is estimated as an average part of r; else, it is estimated as an average part for time slot $h_1+1, h_1+2, \dots,$ etc. The method is described as following

$$r^k = \begin{cases} \frac{1}{H} r, & \text{if } (k+h_2-1)t_2 \leq h_1 t_1 \\ \frac{1}{H}\hat{r}(\left\lceil \frac{(k+h_2-1)t_2}{t_1} \right\rceil), & \text{otherwise} \end{cases}$$

where $H = t_1/t_2$. When there is disruptive events and estimated request is a range $\hat{r}(h_1) \in [\hat{R}_1(h_1), \hat{R}_2(h_1)]$, similarly we set r^k to an uncertain set of $1/H[\hat{R}_1(h_1), \hat{R}_2(h_1)]$.

The lengths of discrete time slots for learning historical models and updating real-time information do not need to be the same, hence in Algorithm 1 we consider a general case for different t_1, t_2. The main computational cost of each iteration is on step (3), and t_2 should be no shorter than the computational time of the optimization problem. We regulate parameters according to experimental results based on a given data set, since there are no closed form equations to decide optimal design values of these parameters. How to adjust parameters such as objective weight β, time slots t_1, t_2, prediction horizon T are shown in Section 5.

4.2.3 Generalization of Algorithm 1

Distributed RHC algorithm: Since the relaxed form of iteration step (3) is polynomial of the problem size V_n, a centralized framework works well for a certain range of variable numbers based on the computational capability of the monitoring system (see Section 5 for more detail). When centralized computation is not efficient enough for a large-scale problem, we can design a distributed information collecting and iterative computing RHC algorithm. In general a distributed framework introduces a trade-off between dispatch cost and computational complexity, and is one direction of future work.

Real-time information of waiting requests: We do not restrict the data source of waiting requests – it can be either predicted results or customer reservation records. Some companies provide taxi service according to the current requests in the queue. If reservations are received by the dispatch system, we then assign value of the waiting requests vector r^k in (8) according to the reservation, and the solution is subject to customer bookings.

5. CASE STUDY: METHOD EVALUATION

We conduct trace-driven simulations based on a San Francisco taxi data set [20] summarized in Table 2. In this data set, a record for each individual taxi includes four values: the geometric position (latitude and longitude), a binary indication of whether the taxi is vacant or with passengers, and the Unix epoch time. With these records, we learn average requests and mobility patterns of taxis, which serves as the input of Algorithm 1. We note that our learning model is not restricted to the data set used in this simulation, and other models [26] and date sets can also be incorporated.

We implement Algorithm 1 in Matlab using an optimization toolbox called CVX [8]. We assume that all vacant taxis follow the dispatch solution and go to suggested regions. Inside a target region, we assume that a vacant taxi automatically picks up the nearest request, and we calculate the total idle mileage including distance across regions and inside a region by simulation. The mileage between two points is approximated as proportional to their geographical distance on the road map, since a city is a small area on the earth surface. Geometric location of a taxi is directly provided by GPS data. Hence, we calculate geographic distance directly from the data first, and then convert the result to mileage.

Experimental figures shown in Subsection 5.1 and 5.3 are average results of all weekday data from the data set of Table 2. Results shown in Subsection 5.2 are based on weekend data.

Estimate request and drop off vectors: Requests during different times of a day in different regions vary a lot, and Figure 4 compares bootstrap results of requests $\hat{r}(h_1)$ on weekdays and weekends for selected regions. This shows a motivation of this work—necessary to balance the number of taxis according to the demand from the perspective of system-level optimal performance. Drop off vectors $dp(h_1)$ are also calculated via bootstrap method.

5.1 RHC with real-time sensor information

Real-time GPS and occupancy data provides latest position information of all vacant and occupied taxis. When dispatching available taxis with true initial positions, the total idle distance is reduced 52% compared with the result without dispatch methods, as shown in Figure 5. This is because the optimization problem (8) returns a solution with smaller idle distance cost given more accurate initial position information P^0. Figure 5 also shows that even applying dispatch solution calculated without real-time information is better than non dispatched result.

Based on the partition of Figure 7, Figure 6 shows that the supply demand ratio at each region of the dispatch solution with real-time information is closest to the supply demand ratio of the whole city, and the error $\left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{R^k} r^k \right\|_1$ is reduced 45% compared with no dispatch results. Even the supply demand ratio error of dispatching without real-time information is better than no dispatch solutions. We still allocate vacant taxis to reach a nearly balanced supply demand ratio regardless of their initial positions, but idle distance is increased without real-time data, as shown in Figure 5. Based on the costs of two objectives shown in Figures 5 and 6, the total cost is higher without real-time information, mainly results from a higher idle distance.

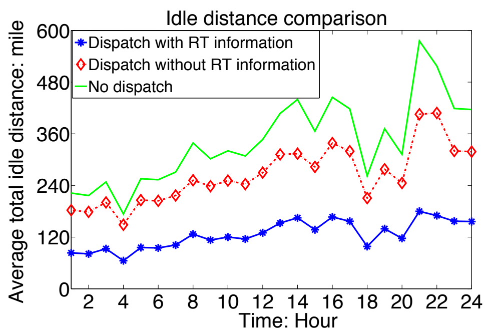


Figure 5: Average idle distance comparison for no dispatch, dispatch without real-time data, and dispatch with real-time GPS and occupancy information. Idle distance is reduced 52% given real-time information, compared with historical data without dispatch solutions.

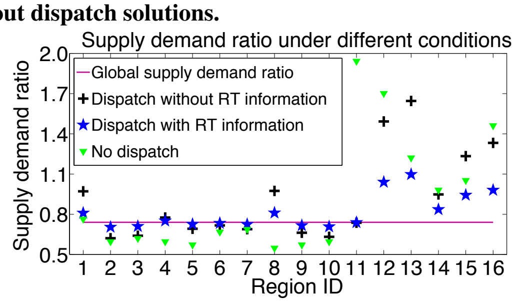


Figure 6: Supply demand ratio of the whole city and each region for different dispatch solutions. With real-time GPS and occupancy data, the supply demand ratio of each region is closest to the global level. The error of supply demand ratio $\left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{R^k} r^k \right\|_1$ is reduced 45% with real-time information, compared with historical data without dispatch solutions.

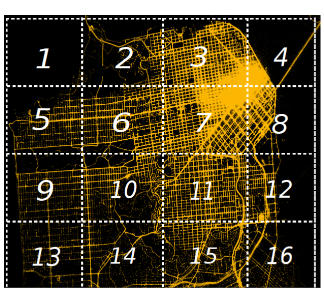


Figure 7: A heatmap about the density of our San Francisco taxicab GPS data set, with a region partition method. The denser the area, the more the GPS data points. Region 3 covers several busy areas, include Financial District, Chinatown, Fisherman Wharf. Region 7 is mainly Mission District, Mission Bay, the downtown area of SF.

To estimate a mobility pattern matrix $\hat{C}(h_2)$, we define a matrix $T(h_2)$, where $T(h_2)_{ij}$ is the total number of passenger trajectories that starting at region i and ending at region j during time slot h_2. We also apply bootstrap process to get $\hat{T}(h_2)$, and $\hat{C}(h_2)_{ij} = \hat{T}(h_2)_{ij} / (\sum_j \hat{T}(h_2)_{ij})$.

For simulation simplicity, we partite the city map to equal-area regions. to get the longitude and latitude position $W_i \in \mathbb{R}^{n \times 2}$ of each vacant taxi, we randomly pick up a station position in the city map by uniform distribution.

5.2 Robust optimization

One disruptive event of the San Francisco data set is Giant baseball at AT&T park, and we choose the historical record on May 31, 2008 as an example to evaluate the robust optimization formulation (9). Customer request number for areas near AT&T park is affected, especially Region 7 around $5:00$pm, which increases about 40% than usual. Figure 8 shows that with a robust optimization formulation (9), the error $\left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{R^k} r^k \right\|_1$ is reduced 25% compared with problem (8), 46% compared with historical supply data without dispatch. Even under solutions of (8), the total supply demand ratio error is reduced 28% compared historical data without dispatch. In general, we consider the factor of disruptive events in a robust RHC iteration, thus the system level supply distribution responses to the demand better under disturbance.

5.3 Design parameters for Algorithm 1

Parameters like the length of time slots, the region division func-

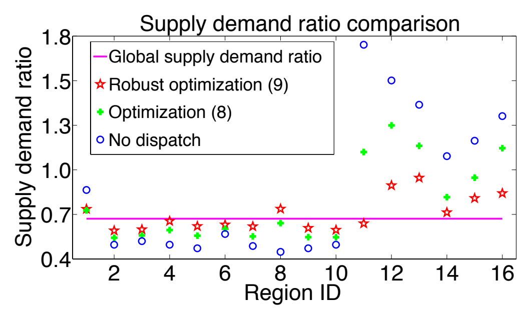


Figure 8: Comparison of supply demand ratio at each region, for solutions of robust optimization (9), problem (8) in the RHC framework, and historical data without dispatch. With the dispatch solution of problem (9), the supply demand ratio of each region is closer to the ratio of the whole city, and the total supply demand ratio error $\left\| \frac{1}{N} \mathbf{1}_N^T X^k - \frac{1}{R^k} r^k \right\|_1$ is reduced 46%.

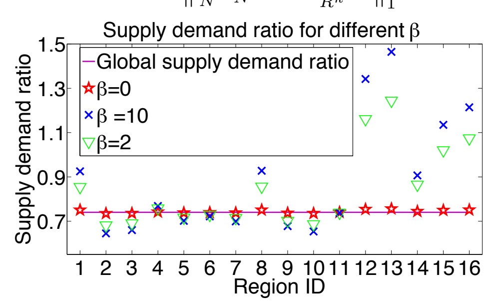


Figure 9: Comparison of supply demand ratios at each region during one time slot for different β values. When β is smaller, we put less cost weight on idle distance that taxis are allowed to run longer to some region, and taxi supply matches with the customer requests better.

β	0	2	10	without dispatch
s/d error	0.645	1.998	2.049	2.664
idle distance	3.056	1.718	1.096	4. 519
total cost	0.645	5.434	13.009	47.854

Table 3: Average cost comparison for different values of β. Idle distance is calculated as the difference between geographical coordinators of two points.

tion, the objective weight parameter and the prediction horizon T of Algorithm 1 affect the results of dispatching cost in practice. Optimal values of parameters for each individual data set can be different. Given a data set, we change one parameter to a larger/smaller value while keep others the same, and compare results to choose a suboptimal value of the varying parameter. We compare the cost of choosing different parameters for Algorithm 1, and explain how to adjust parameters according to experimental results based on a given historical data set (including both GPS and occupancy record) for a city.

How the objective weight of (8) **–** β **affects the cost: How to choose the number of regions**: The cost function includes two parts — idle cruising cost and the supply demand ratio mismatch cost. This trade-off between two parts is addressed by β, and the weight of idle distance increases with β. A larger β returns a solution with smaller total idle geographical distance, while a larger error between supply demand ratio, i.e., a larger $\left\|\frac{1}{N}\mathbf{1}_N^T X^k - \frac{1}{R^k} r^k\right\|_1$ value. The two components of the cost with different β by Algorithm 1, and cost of historical data without dispatch algorithm are shown in Table 3. The supply demand ratio mismatch cost is shown in the s/d error row. The idle distance row shows the distance between geographical coordinators between points, since the idle driving distance in problem (8) is calculated based on GPS sensing data.

We calculate the total cost as (s/d error $+\beta\times$ idle distance) (Use $\beta = 10$ for the without dispatch column). Though with $\beta = 0$ we can dispatch vacant taxis to make the supply demand ratio of each region closest to that of the whole city, a larger idle geographical distance cost is introduced compared with $\beta = 2$ and $\beta = 10$. Compare the idle distance when $\beta = 0$ with the data without dispatch, we get 23% reduction; compare the supply demand ratio error of $\beta = 10$ with the data without dispatch, we get 32%.

Average total idle distance during different hours of one day for a larger β is smaller, as shown in Figure 10. The supply demand ratio error at different regions of one time slot is increased with larger β, as shown in Figure 9.

How to decide the number of regions: In general, the dispatch solution of problem (8) for a vacant taxi is more accurate by dividing a city into regions of smaller area, since the dispatch solution is closer to road-segment level. However, we should consider other factors when deciding the number of regions, like the process of predicting requests and mobility patterns based on historical data. A linear model is not a good prediction for future events when the region area is too small, since pick up and drop off events are more irregular in over partitioned regions. While Increasing n, we also increase the computation complexity (the area of each region does not need to be the same as we divide the city in this experiment).

Figure 11 shows that the idle distance will decrease with a larger region division number, but the decreasing rate slows down; while the region number increases to a large number, the average cost

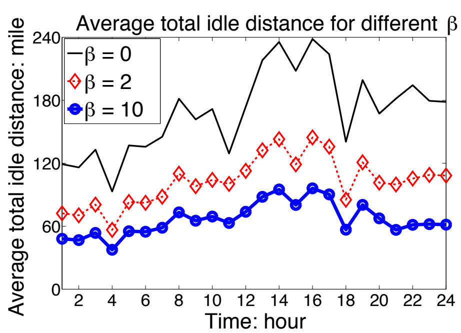


Figure 10: Average total idle distance of taxis at different hours. When β is larger, the idle distance cost weights more in the total cost, and the dispatch solution causes less total idle distance.

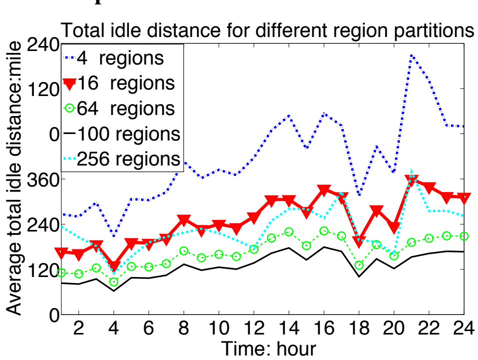


Figure 11: Average total idle distance of all taxis during one day, for different region partitions. Idle distance decreases with a larger region-division number, till the region number increases to a certain level.

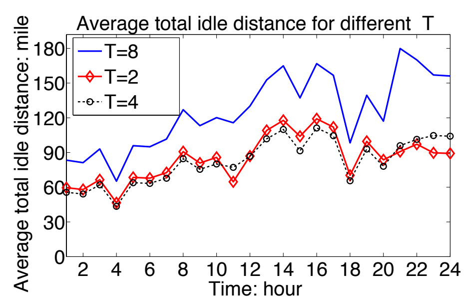


Figure 12: Average total idle distance at different time of one day compared for different prediction horizons.When $T = 4$, idle distance is decreased at most hours compared with $T = 2$. For $T = 8$ the costs are worst.

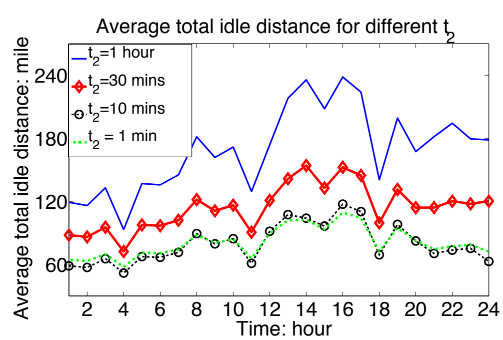


Figure 13: Comparison of average total idle distance for different t_2 – the length of time slot for updating sensor information. With a smaller t_2, the cost is smaller. But when $t_2 = 1$ is too small to complete calculating problem (8)**, the dispatch result is not guaranteed to be better than $t_2 = 10$.**

also increases because the linear model applied in this work does not describe system's behavior well.

How to decide the prediction Horizon T: In general, when T is larger, the total idle distance to get a good supply demand ratio in future time slots should be smaller. However, when T is large enough, increasing T can not reduce the total idle distance any more, since the model prediction error, especially the error of estimating mobility pattern matrix C^k compensates the advantage of considering future costs. For $T = 2$ and $T = 4$, Figure 12 shows that the average total idle distance of vacant taxis at most hours of one day decreases as T increases. For $T = 8$ the driving distance is the largest. Theoretical reasons are discussed in Section 4.2.

Decide the length of time slot t_2: For simplicity, we choose the time slot t_1 as one hour, to estimate requests. A smaller time slot t_2 for updating GPS information can reduce the total idle geographical distance with real-time taxi positions. However, one iteration of Algorithm 1 is required to finish in less than t_2 time, otherwise the dispatch order will not work for the latest positions of vacant taxis, and the cost will increase. Hence t_2 is constrained by the problem size and computation capability.

Figure 13 shows that smaller t_2 returns a smaller idle distance, but when $t_2 = 1$ Algorithm 1 can not finish one step iteration in one minute, and the idle distance is not reduced. The supply demand ratio at each region does not vary much for $t_2 = 30, t_2 = 10$ minutes and $t_2 = 1$ hour. Comparing two parts of costs, we get that t_2 mainly affects the idle driving distance cost in practice.

6. CONCLUSION

In this paper, we propose a novel RHC approach for the taxi dispatch problem. This method utilizes both historical and real-time GPS and occupancy data to build models, and applies predicted models and sensing data to decide locations for vacant taxis considering multiple objectives. From a system level perspective, we compute suboptimal dispatch solutions when reaching a global balanced supply demand ratio with least associated cruising distance, under given constraints. By applying the RHC framework on a San Francisco data set, we show how to regulate parameters in the framework design process according to experiments. Evaluation results support the system level performance improvements of our RHC framework. In the future, we will enhance problem formulation considering information like passenger destination, road congestion, and effects of model uncertainties.

References

[1] I. Amundson and X. D. Koutsoukos. A survey on localization for mobile wireless sensor networks. In *Mobile Entity Localization and Tracking in GPS-less Environnments*, pages 235–254. Springer Berlin Heidelberg, 2009.

[2] A. Arsie, K. Savla, and E. Frazzoli. Efficient routing algorithms for multiple vehicles with no explicit communications. *IEEE Transactions on Automatic Control*, 54(10):2302–2317, 2009.

[3] M. Asif, J. Dauwels, C. Goh, A. Oran, E. Fathi, M. Xu, M. Dhanya, N. Mitrovic, and P. Jaillet. Spatio and temporal patterns in large-scale traffic speed prediction. *IEEE Transactions on Intelligent Transportation Systems*, (15):797–804, 2014.

[4] S. Blandin, D. Work, P. Goatin, B. Piccoli, and A. Bayen. A general phase transition model for vehicular traffic. *SIAM Journal on Applied Mathematics*, 71(1):107–127, 2011.

[5] D. R. Choffnes and F. E. Bustamante. An integrated mobility and traffic model for vehicular wireless networks. In *Proceedings of the 2Nd ACM International Workshop on Vehicular Ad Hoc Networks*, pages 69–78, 2005.

[6] J. Cortes, S. Martinez, T. Karatas, and F. Bullo. Coverage control for mobile sensing networks. *IEEE Transactions on Robotics and Automation*, 20(2):243–255, 2004.

[7] R. Ganti, M. Srivatsa, and T. Abdelzaher. On limits of travel time predictions: Insights from a new york city case study. In *IEEE 34th ICDCS*, pages 166–175, June 2014.

[8] M. Grant and S. Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, 2014.

[9] Hara Associates Inc. and Corey, Canapary& Galanis. Taxi user surveys. `http://www.sfmta.com/sites/default/files/Draft%20SF%20UserSurvey%2055%20WEB%20version04042013.pdf`, 2013.

[10] J. Herrera, D. Work, R. Herring, X. Ban, Q. Jacobson, and A. Bayen. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C*, 18(4):568–583, 2010.

[11] A. Howard, M. Matarić, and G. Sukhatme. Mobile sensor network deployment using potential fields: A distributed, scalable solution to the area coverage problem. In *Distributed Autonomous Robotic Systems 5*. Springer Japan, 2002.

[12] Y. Huang and J. W. Powell. Detecting regions of disequilibrium in taxi services under uncertainty. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, number 10, pages 139–148, 2012.

[13] K.-D. Kim. Collision free autonomous ground traffic: A model predictive control approach. In *Proceedings of the ACM/IEEE 4th ICCPS*, number 10, pages 51–60, 2013.

[14] B. Li, D. Zhang, L. Sun, C. Chen, S. Li, G. Qi, and Q. Yang. Hunting or waiting? discovering passenger-finding strategies from a large-scale real-world taxi dataset. In *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pages 63–38, 2011.

[15] Z. Liao. Real-time taxi dispatching using global positioning systems. *Commun. ACM*, 46(5):81–83, May 2003.

[16] G. Naus, J. Ploeg, M. V. de Molengraft, W. Heemels, and M. Steinbuch. Design and implementation of parameterized adaptive cruise control: An explicit model predictive control approach. *Control Engineering Practice*, 18(8):882 – 892, 2010.

[17] R. Negenborn, B. D. Schutter, and J. Hellendoorn. Multi-agent model predictive control for transportation networks: Serial versus parallel schemes. *Engineering Applications of Artificial Intelligence*, 21(3):353 – 366, 2008.

[18] New York City Taxi and Limousine Commission. Taxi of tomorrow survey results. `http://www.nyc.gov/html/tlc/downloads/pdf/tot_survey_results_02_10_11.pdf`, 2011.

[19] C. Petres, Y. Pailhas, P. Patron, Y. Petillot, J. Evans, and D. Lane. Path planning for autonomous underwater vehicles. *IEEE Transactions on Robotics*, 23(2):331–341, 2007.

[20] M. Piorkowski, N. Sarafijanovic-Djukic, and M. Grossglauser. A parsimonious model of mobile partitioned networks with clustering. In *First International Communication Systems and Networks and Workshops (COMSNETS)*, pages 1–10, 2009.

[21] J. W. Powell, Y. Huang, F. Bastani, and M. Ji. Towards reducing taxicab cruising time using spatio-temporal profitability maps. In *Proceedings of the 12th International Conference on Advances in Spatial and Temporal Databases*, number 19, pages 242–260, 2011.

[22] M. Qu, H. Zhu, J. Liu, G. Liu, and H. Xiong. A cost-effective recommender system for taxi drivers. In *20th ACM SIGKDD International Conference on KDD*, pages 45–54, 2014.

[23] K.-T. Seow, N. H. Dang, and D.-H. Lee. A collaborative multiagent taxi-dispatch system. *IEEE Transactions on Automation Science and Engineering*, 7(3):607–616, 2010.

[24] Y. Wan and S. Roy. A scalable methodology for evaluating and designing coordinated air-traffic flow management strategies under uncertainty. *IEEE Transactions on Intelligent Transportation Systems*, 9(4):644–656, 2008.

[25] J. Yang, P. Jaillet, and H. Mahmassani. Real-time multivehicle truckload pickup and delivery problems. *Transportation Science*, 38(2):135–148, 2004.

[26] D. Zhang, T. He, S. Lin, S. Munir, and J. A. Stankovic. Dmodel: Online taxicab demand model from big sensor data in. In *IEEE International Congress on Big Data*, pages 152–159, 2014.

[27] R. Zhang and M. Pavone. Control of robotic mobility-on-demand systems: a queueing-theoretical perspective. In *Proceedings of Robotics: Science and Systems*, July 2014.