Title

Incentive Systems for New Mobility Services

Authors

Ghafelebash, Ali
Razaviyayn, Meisam
Dessouky, Maged

Publication Date

2022-07-01

DOI

10.7922/G2FQ9TX1

Data Availability

The data associated with this publication are available at:
https://doi.org/10.7910/DVN/7VI4LX

Incentive Systems for New Mobility Services

July 2022

A Research Report from the National Center for Sustainable Transportation

Ali Ghafelebashi, University of Southern California

Meisam Razaviyayn, University of Southern California

Maged Dessouky, University of Southern California

National Center for Sustainable Transportation

METRANS Transportation Consortium USC | CSULB

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NCST-USC-RR-22-26	N/A	N/A

4. Title and Subtitle
Incentive Systems for New Mobility Services

5. Report Date
July 2022

6. Performing Organization Code
N/A

7. Author(s)
Ali Ghafelebashi, https://orcid.org/0000-0001-8339-7960
Meisam Razaviyayn, PhD, https://orcid.org/0000-0003-4342-6661
Maged Dessouky, PhD, https://orcid.org/0000-0002-9630-6201

8. Performing Organization Report No.
N/A

9. Performing Organization Name and Address
University of Southern California
METRANS Transportation Consortium
University Park Campus, VKC 367 MC:0626
Los Angeles, California 90089-0626

10. Work Unit No.
N/A

11. Contract or Grant No.
USDOT Grant 69A3551747114

12. Sponsoring Agency Name and Address
U.S. Department of Transportation
Office of the Assistant Secretary for Research and Technology
1200 New Jersey Avenue, SE, Washington, DC 20590

13. Type of Report and Period Covered
Final Research Report (January 2021 – December 2021)

14. Sponsoring Agency Code
USDOT OST-R

15. Supplementary Notes
DOI: https://doi.org/10.7922/G2FQ9TX1
Dataset DOI: https://doi.org/10.7910/DVN/7VI4LX

16. Abstract
With rapid population growth and urban development, traffic congestion has become an inescapable issue in large metropolitan regions. Research studies have proposed different strategies to control traffic, ranging from roadway expansion to transportation demand management programs. Among these strategies, congestion pricing and incentive offering schemes have been widely studied as reinforcements for traffic control in traditional traffic networks where each driver is a "player" in the network. In such a network, the "selfish" behavior of individual drivers prevents the entire network to reach a socially optimal operation point. In future mobility services, on the other hand, a large portion of drivers/vehicles may be controlled by a small number of companies/organizations. In such a system, offering incentives to organizations can potentially be much more effective in reducing traffic congestion rather than offering incentives directly to drivers. This research project studies the problem of offering incentives to organizations to change the behavior of their individual drivers (or individuals using their organization's services). The incentives are offered to each organization based on their aggregated travel time loss across all their drivers. This step requires solving a large-scale optimization problem to minimize the system-level travel time. We propose an efficient algorithm for solving this optimization problem. To evaluate the performance of the proposed algorithm, multiple experiments are conducted by Los Angeles traffic data. Our experiments show that the proposed algorithm can decrease the system-level travel time by up to 6.9%. Moreover, our experiments demonstrate that incentivizing organizations can be up to 8 times more efficient than incentivizing individual drivers in terms of incentivization monetary cost.

17. Key Words	18. Distribution Statement
New Mobility Services, Congestion Reduction, Incentive, Behavior Change, Travel Demand Management.	No restrictions.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	48	N/A

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

About the National Center for Sustainable Transportation

The National Center for Sustainable Transportation is a consortium of leading universities committed to advancing an environmentally sustainable transportation system through cutting-edge research, direct policy engagement, and education of our future leaders. Consortium members include: University of California, Davis; University of California, Riverside; University of Southern California; California State University, Long Beach; Georgia Institute of Technology; and University of Vermont. More information can be found at: ncst.ucdavis.edu.

Disclaimer

Acknowledgments

This study was funded, partially or entirely, by a grant from the National Center for Sustainable Transportation (NCST), supported by the U.S. Department of Transportation (USDOT) through the University Transportation Centers program. The authors would like to thank the NCST and the USDOT for their support of university-based research in transportation, and especially for the funding provided in support of this project.

NCST

Incentive Systems for New Mobility Services

A National Center for Sustainable Transportation Research Report

July 2022

Ali Ghafelebashi, Meisam Razaviyayn, and Maged Dessouky

Daniel J. Epstein Department of Industrial & Systems Engineering, University of Southern California

[page intentionally left blank]

TABLE OF CONTENTS

List of Tables

List of Figures

Incentive Systems for New Mobility Services

EXECUTIVE SUMMARY

With rapid population growth and urban development, traffic congestion has become an inescapable issue, especially in large cities worldwide. Many congestion reduction strategies have been proposed in the past, ranging from roadway extension, transportation systems management, and operations to demand management. In particular, as a demand management approach, congestion pricing and incentive offering schemes have been used as reinforcements for traffic control. Despite extensive research on congestion pricing mechanisms, almost all studies focus on traditional mobility systems, and little has been done for future mobility services. With recent technological advancements, the shape of mobility services is drastically changing. Traditionally, the driver is the car owner and is the ultimate decision-maker on his/her origin, destination, routing, and time of travel. In contrast, future mobility systems consist of different organizations and companies that completely (or partially) influence the behavior of individual human (or AI-based) drivers. Such organizations include car-sharing services (e.g., Zipcar, Turo), ride-hailing services (e.g., Uber, Lyft), crowdsourcing delivery systems (e.g., Amazon Flex, Instacart, DoorDash), navigation applications (e.g., Google Maps and Waze), and even companies producing autonomous cars with built-in navigation systems (e.g., Tesla), to name just a few.

In this work, we develop mechanisms for offering incentives to organizations and companies to change the behavior of individual drivers in their organization (or individuals using their organization's services). Such mechanisms can be more effective than traditional individual-level incentive offering mechanisms since each organization can control a large pool of individual drivers, thus moving the traffic flow toward the optimal "system-level" performance. In addition, such an "organization-level" incentive offering enjoys more flexibility than the individual "driver-level" incentive mechanisms. Our approach provides incentives to organizations to reduce the travel time of the system by indirectly influencing the behavior of individual drivers in these organizations. The provided model relies on historical data as well as demand estimates provided by organizations to predict traffic flow of the network; and provides incentives to organizations to reduce the system-level congestion.

We formulate the optimal incentive offering mechanism as a large-scale optimization problem and develop an efficient algorithm for solving it. We evaluate the performance of our method using data from the Los Angeles area. The Los Angeles region is ideally suited for being the validation area as one of the most congested cities in the US. Additionally, researchers at USC have developed the Archived Data Management System (ADMS) that collects, archives, and integrates a variety of transportation datasets from Los Angeles, Orange, San Bernardino, Riverside, and Ventura Counties. ADMS includes access to real-time traffic datasets from i) 9500 highway and arterial loop detectors providing data approximately every 1 minute, and ii) 2500 bus and train GPS location (AVL) data operating throughout Los Angeles County. Our experiments on this data show that the proposed model can reduce the total travel time of the system by up to 6.9% during peak traffic period in the morning. In addition, incentivizing

organizations can be up to 8 times more efficient than incentivizing individual drivers in terms of incentivization monetary cost, according to our model.

Introduction

According to the Transportation Statistics Annual Report of 2018, there was a 15.6% growth in vehicle-miles of travel and 19% increase in the total number of registered vehicles from 2000 to 2018. However, for the same period of time, public road and street mileage was just expanded by only 5.1% and total lane-miles merely augmented by 6% [1]. Due to an increase in the number of vehicles and traveled miles and the slow expansion of transportation infrastructure, cities are experiencing serious congestion. Today, traffic congestion is one of the most prevailing problems in large cities around the world causing lower quality of life and huge economic losses. INRIX, a company that provides traffic management services, estimates a $88 billion economy loss and 99 hours loss of time per driver in United State because of traffic congestion in 2019 [2]. In the Los Angeles area alone, each driver lost 103 hours in 2019 due to traffic congestion [3].

Beside economic losses, quality of air can be exacerbated by traffic congestion which can lead to worsening of the community's health. Transportation Research Board, one of the seven program units of the National Academies of Sciences, Engineering, and Medicine, reports emitted pollutants by vehicles as the main source of air pollution [4]. Duration and intensity of traffic congestion can aggravate emission levels [5] which leads to an increase in air pollution (specifically NO2) [6]. In addition, traffic congestion can increase the frequency of aggressive behavior of drivers and increase their stress level. In particular, Hennessy and Wiesenthal show an increase in the Likert scale, a psychometric scale which ranges from 0 = “low stress level” to 4 = “high stress level”, by more than two times from 0.8 to 1.73 during traffic congestion [7]. Higher levels of stress in drivers can result in a rise in accident occurrences [8].

Cambridge Systematics [9], a transportation consultancy firm that works on planning and policy related to movement of people and goods, categorizes the congestion reduction solutions into three groups:

1. Adding more capacity
2. Transportation System Management and Operation (TSM&O)
3. Demand Management

Strategy 1 consists of transportation infrastructure expansion such as adding more lanes to highways and building new roads. This solution may be effective for solving traffic congestion. However, different reasons such as local and national movements against this strategy and financial constraints have hindered taking this solution into action in recent years. It should be also noted that in congested metropolitan areas, it is unlikely to have congestion improvement by increasing the lane miles [10] because of the increase in the Vehicle Miles Traveled after lane miles increase [10]. TSM&O, the second strategy, aims at making the existing infrastructure more efficient and controls the short-term demand for the current network. TSM&O strategies, such as reversible commuter lanes, dynamic re-timing of traffic signals, providing information about travel conditions to travelers, and converting streets to one-way operation, are much more cost-efficient compared to constructing new transportation infrastructure in 1. Although the cost of TSM&O strategies is low, they are not enough to solve traffic congestion. The third strategy, Demand Management, includes Travel Demand Management (TDM), non-automotive travel modes, and land use management. The focus of Travel Demand Management is not on transportation infrastructure but better management of travel demand. Putting more people into fewer vehicles (e.g., ride-sharing), shifting the time of travel, and removing the need for travel altogether (e.g., teleworking) are a few examples of TDM. The need for substantial changes in drivers' lifestyle and inflexibility of workers' schedule are some of the main challenges of TDM strategies. To lower the travel rate of personal vehicles, more investment in the non-automotive form of transportation such as bus, rail transit systems, and bikeways can be employed as alternative strategies. In this report, we focus on a TDM approach.

In this research, we develop a mechanism to change the behavior of individual drivers in organizations. To be more specific, Demand Management which is the third category of Cambridge Systematics is our utilized framework. Pricing mechanisms in the literature can be considered as the closest approach to this work. There is an extensive study in both theory and practice for road pricing policies such as taxation or assigning a fee for entering a highway or road. [12] and [13] study the effect of monetary penalties on drivers' traveling behavior and how it results in a change in drivers' behavior (see the book [14], Ph.D. thesis , and the references therein). This category aims to motivate people to avoid congested roads and reduce the congestion in these roads by reducing their traffic flow. Multiple factors can be considered in these pricing mechanisms such as time [16], distance [17], or vehicle characteristics [18].

Although pricing schemes seem a good solution from a market point of view, issues such as equity concerns lead to challenges in the implementation of congestion pricing/taxation policies [19]. For instance, in some of the past implementations in Lyon, France; Mexico City, Mexico; and Genoa, Italy; the congestion pricing mechanism was not successful in the test phase because of low public acceptance [20]. Beside the equity barriers, policymakers have been hindered from implementing advanced congestion pricing schemes due to complexity and uncertainties in planning pricing policies [21]. Gu et al. [22] proposed a pricing model to meet equity and efficiency challenges. The proposed model charges drivers based on both distance and time. The model uses the amount of time spent in the traffic jam to charge the driver. Tradable credits (TCs) or tradable mobility permits (TPMs) are another token-based pricing mechanism [23]. In these schemes, drivers are allowed to trade certain tokens/credits via a market mechanism. There is a limit on the total number of vehicles in the trading system and the total available credits are usually considered as a fixed value – see [24] for a review article on this topic. Different mechanisms in this category have been put forward such as receiving free travel cards [25]. In addition, [26] and references therein have proposed multiple mathematical programming methods to model and define algorithms for such token-based schemes. [27] illustrates how tradable credits can be advantageous theoretically. Some economic sectors have made use of such token markets such as airport slot allocation [28]. However, design complexities [29] of such cap-and-trade programs have prevented implementation of these schemes in individual-level personal travels and daily commutes [30].

Recently, positive incentive schemes have attracted more attention from researchers. Rewarding policy is more popular in comparison with approaches that are based on charging a fee [31]. In addition, better effectiveness of rewarding desirable behavior compared to penalizing unpreferred ones has been shown in psychological theory of reactance [31]. Despite the effectiveness of rewarding policies in changing individual behavior [32], there are not many studies on the effectiveness of rewarding methods in the transportation area. The INSTANT project [34] has offered positive incentives to motivate travelers to avoid peak times for their commute. Rush hours or peak times are the time of the day with the heaviest congestion. The CAPRI project of Stanford [35] shows the effectiveness of rewarding policy in reducing congestion by offering positive incentives for rush hour avoidance and by changing travel mode to walking and biking. [36] has conducted a series of studies in the Netherlands on how effective offering positive incentives can be on discouraging drivers from traveling during rush hour. Different alternatives were incentivized in these projects such as avoiding rush hour (before and after the peak of the traffic in the morning), teleworking, or choosing alternative commute modes such as carpooling, cycling, and public transportation. [37] provides different levels of incentive to change drivers' departure time and their routing decision. Token as an alternative form of incentive was recently offered by [29] for different commuting options such as ride-sharing, different travel modes, and different routes. This study utilizes the travel history of users to learn their decisions and adjust choices based on their preferences. Users receive multiple choices for each trip and the value of an incentive of an alternative will be higher if it results in more congestion reduction in the network. Although these studies have shown short-term success in implementing rewarding policies, they did not necessarily lead to permanent changes in the behavior of individual participants. For instance, when the incentivization stopped in [38], participants returned to their previous behavior although the rewarding policy was successful in initially changing traveler's behavior during the experiment period.

Different transportation studies have utilized numerous options as an incentive. In Australia, [39] provided an early bird ticket program to participants to solve the problem of rail overcrowding during peak times. In Beijing, free WiFi and discounted ticket fares could effectively reduce the number of commuters during the morning rush hour [40]. To study how incentives can change the frequency of commuting by bus, [41] offered free bus tickets. In a similar study, [42] offered pre-paid bus tickets to college students in Germany to study its effectiveness in increasing the number of trips by bus. The Tripod project used token form incentives to reduce energy use [29]. The amount of received tokens were based on the amount of energy that a traveler saved. Users could redeem the earned tokens for goods and services at local businesses and agencies participating in Tripod. Users in [43] could collect credit in the smartphones that they received after joining the project. The credits could be redeemed for money or if it was more than a threshold, users could keep the provided smartphone. In the CAPRI project of Stanford [35], incentives were offered in the form of game points that users could collect by installing an app on their smartphone. Users could trade 100 points for $1 or use the collected points to play a game in which they may gain money and points or lose points. In a different study, [44] provides an algorithm to offer personalized incentives to drivers to reduce the traffic congestion by changing the routing decision of the drivers. These incentives could be personalized based on users' preferences.

In traditional congestion pricing and incentive offering mechanisms, incentives are offered directly to individual drivers to influence their decisions such as departure time and routing (Figure 1 (a)). In modern and future mobility services, many of these decisions are indirectly (or directly) made by organizations providing different transportation services. For example, navigation apps, which are regularly used by almost 70% of smartphone users [45], influence the routing decision of millions of drivers daily. Another example is crowdsourcing delivery platforms such as Amazon Flex, Instacart, and Doordash. According to a recent Morgan Stanley estimate, Amazon is already delivering almost half of its own packages using crowdsourcing delivery and the number of packages delivered by Amazon will surpass other delivery services such as USPS and FedEx by 2022 [68]. Another example of such organizations is ride-hailing organizations such as Uber and Lyft. In addition to these drastic changes in mobility services, the rise of autonomous cars will make future mobility services the ultimate decision-maker in routing, origin-destination selection, and the travel time in many applications. Thus, instead of incentivizing individual drivers, it is more advantageous to incentivize organizations to reduce congestion. Intuitively, since organizations have more flexibility and more power to change the traffic, it is expected that incentivizing organizations be more efficient than incentivizing individual drivers. Furthermore, a mechanism has more options in balancing the route selection across the large pool of drivers employed by the organization. Motivated by this idea, this research project develops an incentive offering mechanism to organizations to indirectly (or directly) influence the behavior of individual drivers (Figure 1 (b)) Our framework will be based on the following three-step procedure:

Step 1) The central planner receives organizations' demand estimates for the next time interval (e.g., next few hours)

Step 2) The central planner offers incentives to organizations to change their routes, travel time, and demand if necessary.

Step 3) Observe organizations' response and go back to Step1 for the next time interval.

The central planner (which is referred to as "Incentive Offering Platform" in Figure 1 (b) collects the organizations' demand estimates. We assume that organization will share the OD of their vehicles with the central planner. Based on the collected information, the central planner offers incentives to organizations to influence drivers' behavior. Then the central planner observes the organizations' responses to the incentives offered, and the entire process is continuously repeated in the network. In this process, the assigned route to each driver is fixed and does not change during the travel and driver is assumed to follow the route during the travel. When there is no incentivization, we assume that drivers select the route with the smallest travel time. The central planner is assumed to have the information of the routes with the smallest travel time given the OD.

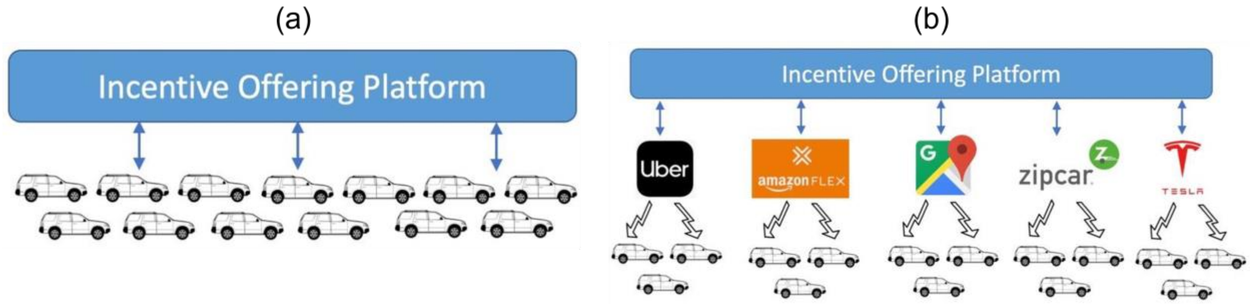


Figure 1. (a) Traditional platforms for offering incentives. (b) Presented platform for offering incentives.

The rest of this report is organized as follows. Section "Incentive Offering Mechanisms" introduces the basic notations and provides a description of our incentive offering mechanism for congestion reduction. We formulate an optimization problem to find the "optimal" incentive offering strategy. We then propose an algorithm for solving this optimization problem efficiently. Results of numerical experiments for the model are presented in Section "Numerical Experiments" using data from the Los Angeles area. Conclusions are provided in Section "Conclusion".

Incentive Offering Mechanisms

Given the origin-destination information of drivers of the organizations, the goal is to find the "optimal" strategy for offering organization--level incentives to them to reduce the traffic congestion of the system. To mathematically state the problem, we begin this section by elaborating on our notations. For further details of the notation, an example is provided in Appendix "An Example of the Model and Notations".

The traffic network is modeled by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. Vertices $\mathcal{V}$ of the graph are major ramps and intersections in the network and are connected by a set of edges $\mathcal{E}$. The direction of an edge in our directed graph is based on the direction of the road from which a driver can go from one node to another. The adjacency of two nodes is based on the possibility of driving directly from one node to another without visiting any other node. The total number of road segments/edges of the network is denoted by $|\mathcal{E}|$ (i.e., the cardinality of the set $\mathcal{E}$). A route in the network is a path in the graph and is denoted by one-hot encoding scheme. In other words, a given route is represented by a vector $r \in \{0,1\}^{|\mathcal{E}|}$ in which the k-th entry is one if the k-th edge is in route r and it is zero, otherwise. Let $T = \{0,1, \ldots, T\}$ denote the defined time horizon such that $t = 1$ is the starting time of the system. The vector $v_t \in R^{|\mathcal{E}|}$ denotes the traffic volume of road segments at time t in which the k-th entry is the total number of vehicles of road segment k at time t.

Let $\mathcal{N} = \mathcal{N}_1 \cup \ \ldots \cup \mathcal{N}_n$ denote the set of all drivers and $\mathcal{N}_i$ denote the set of drivers of organization i. If a driver works for multiple organizations, he or she will be counted as a different driver at each organization. Hence, $\mathcal{N}_1 \cap \ \ldots \cap \mathcal{N}_n = \emptyset$. For any driver $j \in \mathcal{N}_i$, let $\mathcal{R}_j \subseteq$

$\{0,1\}^{|\mathcal{E}|}$ denote the set of driver's possible route choices between their origin and destination. The binary variable $s_i^{r,j} \in \{0,1\}$ represents the assigned route to driver j of organization i. For this driver and given the route $\boldsymbol{r} \in \mathcal{R}$, the variable $s_i^{r,j} = 1$ if route $\boldsymbol{r}$ is assigned to the j^{th} driver of organization i; and $s_i^{r,j} = 0$, otherwise. Each driver can only be assigned to one route, i.e., $\sum_{r \in \mathcal{R}_j} s_i^{r,j} = 1$ Given any routing strategy assigned to drivers, we model the decision of the drivers deterministically due to the power of the organizations in enforcing their drivers' route.

If a driver works for an organization, we call it a *business driver*. In this project, we change the routing decision of business drivers via incentivizing their organizations. We assume that organizations accept our route assignments if the incentive offer can compensate for the change in their total travel time. Notice that when the organizations make their decision in accepting the offer, they do not have access to the offered route assignments to the other organizations. Hence, they can only check the travel time estimations based on the historical data. In the next section, we present our model and formulation in more detail. A complete list of notations used in this report can be found in the Appendix "List of Notations".

Model

In this section, we present our formulation to optimize a certain cost function of the system given the amount of available budget. In the presented formulation, we employ total travel as our cost function but any other cost function such as energy consumption or carbon emission can be used. We compute the system total travel time by summing the drivers' travel time of all road segments over all time periods in the horizon of interest:

$$F_{tt}(\hat{v}) = \sum_{\ell=1}^{|\mathcal{E}|} \sum_{t=1}^{|T|} \hat{v}_{\ell,t} \theta_{\ell,t}(\hat{v}_{\ell,t}) \quad (1)$$

where $\theta_{\ell,t}$ is the travel time of link l at time t (which itself is a function of the volume of the link at the given time). Here, $\hat{\boldsymbol{v}}$ is the vector of volume of links in which $\widehat{v_{\ell,t}}$ is the $(|\mathcal{E}| \times t + \ell)^{th}$ element of vector $\hat{\boldsymbol{v}}$ representing the volume of the ℓ^{th} link at time t. Given the volume vector, we estimate the travel time of the links at different times as described below.

Travel time value θ**:** Different functions have been proposed to capture the relation between volume and travel time. For instance, the Bureau of Public Roads (BPR) [50] defines a link congestion function in which the travel time of a road has nonlinear relation with its volume:

$$\theta(v) = f_{BPR}(v) = t_0 \left(1 + 0.15 \left(\frac{v}{w}\right)^4\right) \quad (2)$$

where $f_{\text{BPR}}(v)$ is the travel time of the drivers on the road segment given its traffic volume v; the parameter t_0 is the free flow travel time of the road segment; and w is the practical capacity of the road segment. In our experiments, to learn the parameters t_0 and w of the road segments in the Los Angeles area at different times of the day, we utilize the historical traffic data of the road segments. Given the $\theta(.)$ function in (2), in order to compute the total travel time of the system, one needs to compute the volume of the links. Next, we explain how the volume vector is computed in our model.

Volume vector $\hat{\boldsymbol{v}}$: The computation of volume vector $\hat{\boldsymbol{v}}$ requires (approximately) estimating the location of the drivers at different times based on their route knowledge. Clearly, by assigning a different route to a driver, the driver's impact on the values of the vector $\hat{\boldsymbol{v}}$ will be different because the driver's location will change by following a different route. Let us start with explaining our notation for route assignment: For each driver, we have a one-hot encoded vector describing which route has been assigned to the driver. Thus, for each driver we have a binary vector $\boldsymbol{s}_i^j \in \{0,1\}^{|\mathcal{R}|}$ in which only one element has a value of one and it corresponds to the assigned route to the j driver of organization i. As we need one vector for each driver, we can aggregate all our assignments in a matrix $\boldsymbol{S} \in \{0,1\}^{|\mathcal{R}| \times |\mathcal{N}|} = [\boldsymbol{S}_1 \boldsymbol{S}_2 \ldots \boldsymbol{S}_n]$ where $\boldsymbol{S}_i \in \{0,1\}^{|\mathcal{R}| \times |\mathcal{N}_i|}$, which is the assignment matrix of organization i with n being the number of organizations. As drivers cannot travel in the routes that are irrelevant to their OD pair, the corresponding elements in their assignment vector have a value of zero.

Given the driver's route entering the system at a specific time, we need to model the location of the individual in the upcoming times. In order to model the location of drivers in the system, we use the model developed by [51] in which the location of drivers are computed in a probabilistic fashion. This model can be presented by a matrix $\boldsymbol{R} \in [0,1]^{(|\mathcal{E}| \cdot |T|) \times |\mathcal{R}|}$ that estimates the probability of the presence of a driver in a given link at a specific time in the future (assuming that the driver is picking a specific route). Multiple ways to estimate matrix $\boldsymbol{R}$ is suggested in [51] including an approach based on the use of historical data. In our experiments in subsection "Simulation Model", the matrix $\boldsymbol{R}$ is computed based on the volume at User Equilibrium (UE) state of the system.

Given the matrix $\boldsymbol{R}$, it is easy to see that the vector

$$\hat{v} = \boldsymbol{RS1} \in \mathbb{R}^{|\mathcal{E}| \cdot |T|} \tag{3}$$

contains the expected number of vehicles in all the links at each time. Plugging the expression of $\hat{\boldsymbol{v}}$ in (1), we get the total travel time of the system as

$$F_{tt}(\hat{v}) = \sum_{\ell=1}^{|\mathcal{E}|} \sum_{t=1}^{|T|} (\boldsymbol{RS1})_{\ell,t} \theta((\boldsymbol{RS1})_{\ell,t}) = \sum_{\ell=1}^{|\mathcal{E}|} \sum_{t=1}^{|T|} (\boldsymbol{r}_{\ell,t} \boldsymbol{S1}) \theta(\boldsymbol{r}_{\ell,t} \boldsymbol{S1}) \tag{4}$$

where $\boldsymbol{r}_{\ell,t}$ is the row of matrix $\boldsymbol{R}$ which corresponds to link ℓ at time t.

To reduce the total travel time of the system, some drivers can be deviated to alternative routes to lower the traffic flow of congested links. To change the routing assignment of drivers, we need to offer incentives to their organizations such that it can compensate the organizations' financial loss caused by accepting our assignment. For simplicity, we use the increase in the total travel time to the organization as a measure of financial loss. Although we have estimated the travel time of the system from equation (4), we need to compute the "route travel times" to be able to compare the amount of change in travel time of each driver after offering incentives. Given the route travel times, we compute the incentives using a linear model which depends on the value of time (VOT) and the amount of increase in the travel time for each organization. In particular, we assume that, given the route assignment to organization i, the incentive value is

$$c_i = \alpha_i max\left\{0, \sum_{j\in\mathcal{N}_i} \boldsymbol{\delta}^\top \boldsymbol{s}_i^j - \gamma_i\right\} \tag{5}$$

where c_i is the incentive offered to organization i, $\alpha_i \in \mathbb{R}_+$ is the Value of Time (VOT) for a driver in organization i, $\boldsymbol{\delta} \in \mathbb{R}_+^{|\mathcal{R}|\cdot|T|}$ is the travel time of the route for each driver, and γ_i is the sum of the minimum travel time route of each driver of organization i in the absence of incentivization. When $\sum_{j\in\mathcal{N}_i} \boldsymbol{\delta}^\top \boldsymbol{s}_i^j - \gamma_i > 0$, it means that the organization's total travel time has increased compared to the baseline of having no incentive, and hence we will compensate the organization's loss. On the other hand, when $\sum_{j\in\mathcal{N}_i} \boldsymbol{\delta}^\top \boldsymbol{s}_i^j - \gamma_i < 0$, it means that the organization's travel time is improved after incentivization and hence no incentivization is required for this particular organization to participate. The details of our method for computing route travel time vector $\boldsymbol{\delta}$ is described next.

Route travel time vector $\boldsymbol{\delta}$: Estimation of the vector $\boldsymbol{\delta}$ requires the volume of links that can be derived based on the route assignment of the drivers. Let $\boldsymbol{S}$ denote the routing decision of the drivers. Given , we can estimate the volume vector $\boldsymbol{v}$ using (3). By utilizing BPR function (4) and the estimated volume vector $\boldsymbol{v}$, we are able to compute the speed of the links. Given the speed of each link, we can determine the vector $\boldsymbol{\delta}$ that contains the travel time of the different routes for different time slots and the vector $\eta \in \mathbb{R}_+^{K\cdot|T|}$ that contains the travel time of the fastest route for different OD pairs for different times. In order to do so, we rely on the method provided by [51]. Given the minimum travel time between OD pairs in η, we can compute the minimum travel time of organization i as $\gamma_i = (\boldsymbol{B}_i\eta)^\intercal \mathbf{1}$ where $\boldsymbol{B}_i \in \{0,1\}^{|\mathcal{N}_i|\times(K\cdot|T|)}$ is the matrix of shortest travel time assignment of drivers of organization i and $\mathbf{1}$ is the vector of ones. $\boldsymbol{B}_i\eta$ is the vector of shortest travel time between the OD pair for each driver and by summing the elements of this vector we get γ_i.

Proposed formulations: For minimizing the total travel time of the system via providing incentives to organizations, we propose to solve the following optimization problem:

$$\begin{aligned} \min_{\{S_i,c_i\}_{i=1}^n, \hat{v}} \quad & \sum_{\ell=1}^{|\mathcal{E}|} \sum_{t=1}^{|T|} \hat{v}_{\ell,t} \theta_{\ell,t}(\hat{v}_{\ell,t}) \\ s.t. \quad & \hat{v} = \sum_{i=1}^{n} \mathbf{R}\mathbf{S}_i \mathbf{1} \\ & \mathbf{D}\mathbf{S}_i\mathbf{1} = \mathbf{q}_i, \quad \forall i = 1,2,\dots,n \\ & \mathbf{S}_i^\top \mathbf{1} = \mathbf{1}, \quad \forall i = 1,2,\dots,n \\ & S_i \in \{0,1\}^{(|\mathcal{R}||\mathbf{T}|)\times(|\mathcal{N}_i|)}, \quad \forall i = 1,2,\dots,n \end{aligned} \tag{6}$$

$$\mathbf{S}_i^\top \boldsymbol{\delta} \leq \boldsymbol{b}_i \odot \mathbf{B}_i \eta, \quad \forall i = 1,2,\ldots,n$$
$$c_i \geq \alpha_i(\boldsymbol{\delta}^\top \mathbf{S}_i \mathbf{1} - \gamma_i), \quad \forall i = 1,2,\ldots,n$$
$$c_1 + \cdots + c_n \leq \Omega$$
$$c_i \geq 0, \quad \forall i = 1,2,\ldots,n$$

where $\widehat{v_{\ell,t}}$ is an element of vector $\hat{v}$ that corresponds to the volume of link ℓ at time t, $c_i \in \mathbb{R}_+$ is the cost of incentive assigned to organization i, $\boldsymbol{D} \in \{0,1\}^{(K\cdot|T|)\times(|\mathcal{R}|\cdot|T|)}$ is the matrix of route assignment of the OD pairs, $\boldsymbol{b}_i \in \mathbb{R}_+^{|\mathcal{N}_i|}$ denotes the factor by which the travel time of an assigned route can be larger than shortest travel time of the OD pair, $\boldsymbol{B}_i \in \{0,1\}^{|\mathcal{N}_i|\times(K\cdot|T|)}$ is the matrix of shortest travel time assignment of drivers of the organization i, and $\boldsymbol{q}_i \in \mathbb{R}^{K\cdot|T|}$ is the vector of the number of drivers of organization i for each OD pair at different times. Here, K is the number of OD pairs. If there are drivers in the system that do not work for any organization, we can consider them as a single organization that their decision matrix is initialized and has fixed values such that they are assigned to the fastest route (assuming they always select the shortest route). Same idea can be employed for organizations that have not joined the incentivization platform. In (6), the incentive value of each organization is being computed separately based on the increase in travel time of the organization. Hence, total cost is based on the increase in total travel time of the organizations and not the individuals. $\sum_{j\in\mathcal{N}_i} \boldsymbol{\delta}^\top \boldsymbol{s}_i^j - \gamma_i$ is the change of travel time of organization i after incentivization. In this term, changes of travel time of individual drivers can cancel each other out if some are increasing and some are decreasing. This cancelling out effect will lower the amount of incentive required for compensation of change in travel time of the organization. However, incentivizing individuals will require incentivizing all the individuals with increase in travel time. We explain the constraints in more detail below:

Constraint 1 ($\hat{\boldsymbol{v}} = \sum_{i=1}^n \boldsymbol{R}\,\boldsymbol{S_i}\mathbf{1}$): This constraint is the estimation of the volume for the different links at different times based on the routing assignments for the organizations.

Constraint 2 ($\boldsymbol{D}\boldsymbol{S}_i\mathbf{1} = \boldsymbol{q}_i$): This constraint makes sure that we assign the correct number of drivers for the routes between OD pairs. $\boldsymbol{S}_i\mathbf{1}$ represents the number of drivers that have been assigned to the different routes. We use matrix $\boldsymbol{D}$ to sum the number of drivers that are assigned to different routes between the same OD pair. $\boldsymbol{q}_i$ is the vector of the actual number of drivers of organization i that are travelling between the OD pairs and $\boldsymbol{D}\boldsymbol{S}_i\mathbf{1}$ must be equal to $\boldsymbol{q}_i$.

Constraint 3 ($\boldsymbol{S}_i^\top\mathbf{1} = \mathbf{1}$): This constraint simply states that we can only assign one route to each driver of organization i.

Constraint 4 ($\boldsymbol{S}_i \in \{0,1\}^{(|\mathcal{R}|\cdot|T|)\times|\mathcal{N}_i|}$): This constraint imposes binary structure on our decision parameters, where 0 means not assigning a route and 1 is assigning the route.

Constraint 5 ($\boldsymbol{S}_i^\top\boldsymbol{\delta} \leq \boldsymbol{b}_i \odot \boldsymbol{B}_i\eta$): This is our fairness constraint. Due to different reasons such as urgent deliveries by some of the organizations' drivers, they may not accept alternative routes with very large travel time compared to the fastest route. Moreover, the platform should consider fairness between different drivers in terms of the amount of deviation from the shortest travel time. The fairness constraint bounds the deviation of travel time of the assigned routes from the minimum travel time. $\boldsymbol{S}_i^{\intercal}\boldsymbol{\delta}$ represents the travel time of the assigned routes to drivers of organization i. $\boldsymbol{b}_i$ denotes the factor by which deviation is allowed for each driver.

Constraint 6 ($c_i \geq \alpha_i(\boldsymbol{\delta}^{\intercal}\boldsymbol{S}_i\mathbf{1} - \gamma_i)$ and $c_i \geq 0$): These two constraints guarantee (5).

Constraint 7 ($c_1 + \cdots + c_n \leq \Omega$): This is our budget constraint. The scalar c_i represents the cost of the incentive assigned to organization i. Ω is the total budget.

For further elaboration on model (6) and its constraints, an illustrative example is presented in Appendix "An Example of the Model and Notations".

Algorithm for Offering Incentives and A Distributed Implementation

The optimization problem (6) is of large size and includes binary variables ($\boldsymbol{S}_i, \forall i = 1, \dots, n$). Thus, solving it efficiently is a challenging task. In this subsection, we propose an efficient algorithm for solving it. First, we relax the binary constraint $\boldsymbol{S}_i \in \{0,1\}^{(|\mathcal{R}|\cdot|T|)\times|\mathcal{N}|}$ to convex constraint $\boldsymbol{S}_i \in [0,1]^{(|\mathcal{R}|\cdot|T|)\times|\mathcal{N}|}$and aim at finding a solution for the new relaxed problem

$$
\begin{aligned}
\min_{\{S_i,c_i\}_{i=1}^{n}, \hat{v}} \quad & \sum_{\ell=1}^{|\mathcal{E}|}\sum_{t=1}^{|T|} \hat{v}_{\ell,t}\theta_{\ell,t}(\hat{v}_{\ell,t}) \\
s.t. \quad & \hat{v} = \sum_{i=1}^{n} \mathbf{R}\mathbf{S}_i\mathbf{1} \\
& \mathbf{D}\mathbf{S}_i\mathbf{1} = \mathbf{q}_i, \quad \forall i = 1,2,\dots,n \\
& \mathbf{S}_i^{\top}\mathbf{1} = \mathbf{1}, \quad \forall i = 1,2,\dots,n \\
& \mathbf{S}_i \in [0,1]^{(|\mathcal{R}||\mathbf{T}|)\times(|\mathcal{N}_i|)}, \quad \forall i = 1,2,\dots,n \\
& \mathbf{S}_i^{\top}\boldsymbol{\delta} \leq \boldsymbol{b}_i \odot \mathbf{B}_i\eta, \quad \forall i = 1,2,\dots,n \\
& c_i \geq \alpha_i(\boldsymbol{\delta}^{\intercal}\mathbf{S}_i\mathbf{1} - \gamma_i), \quad \forall i = 1,2,\dots,n \\
& c_1 + \cdots + c_n \leq \Omega \\
& c_i \geq 0, \quad \forall i = 1,2,\dots,n
\end{aligned}
\tag{7}
$$

The cost/objective function of this problem is a summation of monomial functions with positive coefficients. Furthermore, $\theta_{\ell,t}$ is an affine mapping of the optimization variable $\boldsymbol{S}_i$. Since our domain is the nonnegative orthant and monomials are convex in this domain, the objective function is convex. As the constraints of this problem are convex, (7) becomes a convex optimization problem. Thus, standard solvers such as CVX [52] can be used to solve this problem. However, these solvers have large computational complexity because of utilizing methods such as interior point methods [53] with $O(n^3)$ iteration complexity where n is the number of variables. The reformulation is provided in the Appendix "Reformulated Optimization Model for ADMM Algorithm". As we discuss in Appendix "Review of ADMM", this reformulation is amenable to the ADMM method [54]-[57], which is a first-order method and scalable. The steps of the resulting algorithm is provided in Algorithm 1. The notations used in this algorithm are defined below.

$$\widetilde{D} = \begin{bmatrix} D & & \\ & \ddots & \\ & & D \end{bmatrix} \quad \Delta = \begin{bmatrix} \delta_p & & \\ & \ddots & \\ & & \delta_p \end{bmatrix} \quad \tilde{R} = [R \quad \dots \quad R] \quad \tilde{I} = [I \quad -I] \quad \tilde{c} = \begin{bmatrix} c \\ \mu \end{bmatrix}$$

$$\widetilde{\mathbf{1}} = \begin{bmatrix} \mathbf{1} \\ \mathbf{0} \end{bmatrix} \; \tilde{\alpha} = \begin{bmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{bmatrix} \quad \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \quad \tilde{u}^t = \begin{bmatrix} S_1^t \mathbf{1} \\ \vdots \\ S_n^t \mathbf{1} \end{bmatrix} \quad \gamma = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{bmatrix} \quad q = \begin{bmatrix} q_1 \\ \vdots \\ q_n \end{bmatrix}$$

$$\lambda_i = \begin{bmatrix} \lambda_{i,1} \\ \vdots \\ \lambda_{i,n} \end{bmatrix}, i = 1,3$$

Algorithm 1 ADMM algorithm for solving (7)

1: **Input**: Initial values: ω^0, $\mathbf{S}_i^0$, $\mathbf{H}_i^0$, $\mathbf{W}_i^0$, $\mathbf{Z}_i^0$, $\mathbf{u}^0$, β_i^0, $\tilde{\beta}^0$, $\tilde{\mathbf{c}}^0$, $\boldsymbol{\lambda}_{1,i}^0 \in \mathbb{R}^{|\mathcal{R}|\cdot|\mathbf{T}|}$, $\boldsymbol{\lambda}_2^0 \in \mathbb{R}^{|\mathcal{E}|\cdot|\mathbf{T}|}$, $\boldsymbol{\lambda}_{3,i}^0 \in \mathbb{R}^{K\cdot|\mathbf{T}|}$, $\boldsymbol{\lambda}_{4,i}^0 \in \mathbb{R}^{|\mathcal{R}|\cdot|\mathbf{T}|}$, $\boldsymbol{\Lambda}_{5,i}^0 \in \mathbb{R}^{(|\mathcal{R}|\cdot|\mathbf{T}|)\times|\mathcal{N}_i|}$, $\boldsymbol{\lambda}_{6,i}^0 \in \mathbb{R}^{|\mathcal{N}_i|}$, $\boldsymbol{\lambda}_7^0 \in \mathbb{R}^n$, $\boldsymbol{\Lambda}_{8,i}^0 \in \mathbb{R}^{(|\mathcal{R}|\cdot|\mathbf{T}|)\times|\mathcal{N}_i|}$, $\lambda_9^0 \in \mathbb{R}$, $\boldsymbol{\Lambda}_{10,i}^0 \in \mathbb{R}^{(|\mathcal{R}|\cdot|\mathbf{T}|)\times|\mathcal{N}_i|}$, Dual update step: ρ, Number of iterations: $\tilde{T}$.

2: **for** $t = 0, 1, \ldots, \tilde{T}$ **do**

3: **for** $\ell = 0, 1, \ldots, |\mathcal{E}|$ **do**

4: **for** $\hat{t} = 1, \ldots, |\mathbf{T}|$ **do**

5: $\omega_{\ell,\hat{t}}^{t+1} = \underset{\omega_{\ell,\hat{t}}}{\operatorname{argmin}}\, \omega_{\ell,\hat{t}}\theta_{\ell,t}(\omega_{\ell,\hat{t}}) + \boldsymbol{\lambda}_{2,(\ell,\hat{t})}^t(\omega_{\ell,\hat{t}} - \mathbf{r}_{\ell,\hat{t}}(\sum_{i=1}^n \mathbf{u}_i^t)) + \frac{\rho}{2}(\omega_{\ell,\hat{t}} - \mathbf{R}_{\ell,\hat{t}}(\sum_{i=1}^n \mathbf{u}_i^t))^2$

6: **end for**

7: **end for**

8: **for** $i = 1, \ldots, n$ **do**

9: $\mathbf{S}_i^{t+1} = (-\boldsymbol{\lambda}_{1,i}^t \mathbf{1}^\top - \boldsymbol{\Lambda}_{5,i}^t - \boldsymbol{\Lambda}_{8,i}^t - \boldsymbol{\Lambda}_{10,i}^t + \rho \mathbf{u}_i^t \mathbf{1}^{t\top} + \rho \mathbf{W}_i^t + \rho \mathbf{H}_i^t + \rho \mathbf{Z}_i^t)(\rho \mathbf{1}\mathbf{1}^\top + 3\rho \mathbf{I})^{-1}$

10: $\beta_i^{t+1} = \Pi\left(\frac{1}{\rho}(-\boldsymbol{\lambda}_{6,i}^t - \rho \mathbf{H}_i^{t\top}\boldsymbol{\delta} + \rho \mathbf{b}_i \odot (\mathbf{B}_i\boldsymbol{\eta}))\right)_{\mathbb{R}_+}$

11: **end for**

12: $\tilde{\mathbf{c}}^{t+1} = \Pi\left(\frac{1}{\rho}(\tilde{\mathbf{I}}^\top\tilde{\mathbf{I}} + \tilde{\mathbf{1}}\tilde{\mathbf{1}}^\top)^{-1}(\tilde{\mathbf{I}}^\top\boldsymbol{\lambda}_7^t - \lambda_9^t\tilde{\mathbf{1}} - \rho\tilde{\mathbf{I}}^\top(\boldsymbol{\alpha}\odot\gamma) + \rho\tilde{\mathbf{I}}^\top(\boldsymbol{\alpha}\odot(\boldsymbol{\Delta}^\top\mathbf{u}^t)) - \rho\tilde{\beta}\tilde{\mathbf{1}} + \rho\Omega\tilde{\mathbf{1}}\right)_{\mathbb{R}_+}$

13: $\mathbf{u}^{t+1} = \frac{1}{\rho}(\mathbf{I} + \tilde{\mathbf{R}}^\top\tilde{\mathbf{R}} + \tilde{\mathbf{D}}^\top\tilde{\mathbf{D}} + (\Delta\tilde{\alpha})(\Delta\tilde{\alpha})^\top)^{-1}(\boldsymbol{\lambda}_1^t + \tilde{\mathbf{R}}^\top\boldsymbol{\lambda}_2^t - \tilde{\mathbf{D}}^\top\boldsymbol{\lambda}_3^t - (\boldsymbol{\Delta}\tilde{\boldsymbol{\alpha}})\boldsymbol{\lambda}_7^t + \rho\tilde{\mathbf{u}}^{t+1} - \rho\tilde{\mathbf{R}}^\top\omega^{t+1} + \rho\tilde{\mathbf{D}}^\top\mathbf{q} + \rho(\boldsymbol{\Delta}\tilde{\boldsymbol{\alpha}})(\boldsymbol{\alpha}\odot\gamma) + \rho(\boldsymbol{\Delta}\tilde{\boldsymbol{\alpha}})(\tilde{\mathbf{I}}\tilde{\mathbf{c}}^{t+1}))$

14: **for** $i = 1, \ldots, n$ **do**

15: $\mathbf{W}_i^{t+1} = \frac{1}{\rho}(\mathbf{1}\mathbf{1}^\top + \mathbf{I})^{-1}(\rho\mathbf{1}\mathbf{1}^\top + \rho\mathbf{S}_i^{t+1} - \mathbf{1}\boldsymbol{\lambda}_{4,i}^{t\top} + \boldsymbol{\Lambda}_{5,i}^t)$

16: $\mathbf{H}_i^{t+1} = \frac{1}{\rho}(\boldsymbol{\delta}\boldsymbol{\delta}^\top + \mathbf{I})^{-1}(-\boldsymbol{\delta}\boldsymbol{\lambda}_{6,i}^{t\top} + \boldsymbol{\Lambda}_{8,i}^t - \rho\boldsymbol{\delta}\beta_i^{t+1\top} + \rho\boldsymbol{\delta}(\mathbf{b}_i\odot\mathbf{B}_i\boldsymbol{\eta})^\top + \rho\mathbf{S}_i^{t+1})$

17: $\mathbf{Z}_i^{t+1} = \mathbb{1}(\rho > \tilde{\lambda})\Pi\left(\left(\frac{1}{\rho-\tilde{\lambda}}\right)(\rho\mathbf{S}_i^{t+1} + \boldsymbol{\Lambda}_{10}^t - \frac{\tilde{\lambda}}{2})\right)_{[0,1]} + \mathbb{1}(\rho < \tilde{\lambda})\Pi\left(\left(\frac{1}{\rho-\tilde{\lambda}}\right)(\rho\mathbf{S}_i^{t+1} + \boldsymbol{\Lambda}_{10}^t - \frac{\tilde{\lambda}}{2})\right)_{\{0,1\}}$

18: **end for**

19: **for** $i = 1, \ldots, n$ **do**

20: $\boldsymbol{\lambda}_{1,i}^{t+1} = \boldsymbol{\lambda}_{1,i}^t + \rho(\mathbf{S}_i^{t+1}\mathbf{1} - \mathbf{u}_i^{t+1})$

21: $\boldsymbol{\lambda}_{3,i}^{t+1} = \boldsymbol{\lambda}_{3,i}^t + \rho(\mathbf{D}\mathbf{u}_i^{t+1} - \mathbf{q}_i)$

22: $\boldsymbol{\lambda}_{4,i}^{t+1} = \boldsymbol{\lambda}_{4,i}^t + \rho(\mathbf{W}_i^{t+1\top}\mathbf{1} - \mathbf{1})$

23: $\boldsymbol{\Lambda}_{5,i}^{t+1} = \boldsymbol{\Lambda}_{5,i}^t + \rho(\mathbf{S}_i^{t+1} - \mathbf{W}_i^{t+1})$

24: $\boldsymbol{\lambda}_{6,i}^{t+1} = \boldsymbol{\lambda}_{6,i}^t + \rho(\mathbf{H}_i^{t+1\top}\boldsymbol{\delta} + \beta_i^{t+1} - \mathbf{b}_i\odot\mathbf{B}_i\boldsymbol{\eta})$

25: $\boldsymbol{\Lambda}_{8,i}^{t+1} = \boldsymbol{\Lambda}_{8,i}^t + \rho(\mathbf{S}_i^{t+1} - \mathbf{W}_i^{t+1})$

26: $\boldsymbol{\Lambda}_{10,i}^{t+1} = \boldsymbol{\Lambda}_{10,i}^t + \rho(\mathbf{S}_i^{t+1} - \mathbf{Z}_i^{t+1})$

27: **end for**

28: $\boldsymbol{\lambda}_2^{t+1} = \boldsymbol{\lambda}_2^t + \rho(\omega^{t+1} - \mathbf{R}(\sum_{i=1}^n \mathbf{u}_i^{t+1}))$

29: $\boldsymbol{\lambda}_7^{t+1} = \boldsymbol{\lambda}_7^t + \rho(\boldsymbol{\alpha}\odot(\boldsymbol{\Delta}^\top\mathbf{u}^{t+1} - \boldsymbol{\delta}) - \tilde{\mathbf{I}}\tilde{\mathbf{c}}^{t+1})$

30: $\lambda_9^{t+1} = \lambda_9^t + \rho(\tilde{\mathbf{c}}^{t+1\top}\tilde{\mathbf{1}} + \tilde{\beta}^{t+1} - \Omega)$

31: **end for**

32: **Return:** $\mathbf{S}_i^{\tilde{T}}, \forall i = 1, \ldots, n$

In Algorithm 1, we use the projection operator $\Pi(\cdot)_{[0,1]}$ that projects elements of a matrix to the interval $[0,1]$. $\Pi(\cdot)_{\mathbb{R}_+}$ is also a projection operator but projects elements of a matrix to $\mathbb{R}_+$. Notice that in Algorithm 1, the computation load of the steps 9, 15, 16, and 17 is extensive because matrices $\boldsymbol{S_i}, \boldsymbol{W_i}, \boldsymbol{H_i}$ and $\boldsymbol{Z}_i$, $i = 1,2,\ldots,n$ have large sizes. However, each column in these matrices correspond to one driver and these steps are not coupled so we can perform the computation of each column in parallel by leveraging parallel computation.

In optimization problem (7) (and consequently (10)), all solutions $\boldsymbol{S}_i^*$ with a fixed value of $\boldsymbol{S}_i^*\boldsymbol{1} = \boldsymbol{u}^*$ lead to the same objective as long as $\boldsymbol{S}_i^{*\top}\boldsymbol{1} = \boldsymbol{1}$. Hence, there can be infinite number of solutions to our convex problem such that many are far from being binary. As we need to find a binary solution for our final variables, we add the following regularizer to the objective function of problem (10) to get a (approximately) binary solution:

$$\Re((\mathrm{Z}_j)_{i,(r,t)}) = -\frac{\tilde{\lambda}}{2}(\mathrm{Z}_j)_{i,(r,t)}((\mathrm{Z}_j)_{i,(r,t)} - 1) \quad (8)$$

where $\tilde{\lambda} \in \mathbb{R}_+$ is the regularization parameter and $\left(\boldsymbol{Z}_j\right)_{i,(r,t)} \in [0,1]$. This regularizer pushes the entries of matrix $Z_i, i = 1,2,\ldots,n$ toward the required binary domain of $\{0, 1\}$ by penalizing objective function when entries are far from this domain.

Although we are looking for a solution of problem (6), Algorithm 1 solves the relaxed version of this problem, i.e., problem (10). Therefore, we use the solution from Algorithm 1 to find a feasible point in (6). For this purpose, we solve the following mixed integer (linear) problem

$$\begin{aligned} \min_{\{S_i,c_i\}_{i=1}^n} \quad & \sum_{i=1}^{n} \|\boldsymbol{S}_i\mathbf{1} - u_i^*\|_1 \\ \text{s.t.} \quad & \mathbf{DS}_i\mathbf{1} = \mathbf{q}_i, \quad \forall i = 1,2,\ldots,n \\ & \mathbf{S}_i^\top\mathbf{1} = \mathbf{1}, \quad \forall i = 1,2,\ldots,n \\ & \mathbf{S}_i \in [0,1]^{(|\mathcal{R}||\mathbf{T}|)\times(|\mathcal{N}_i|)}, \quad \forall i = 1,2,\ldots,n \\ & \mathbf{S}_i^\top\boldsymbol{\delta} \leq \boldsymbol{b}_i \odot \mathbf{B}_i\eta, \quad \forall i = 1,2,\ldots,n \\ & c_i \geq \alpha_i(\boldsymbol{\delta}^\top\mathbf{S}_i\mathbf{1} - \gamma_i), \quad \forall i = 1,2,\ldots,n \\ & c_1 + \cdots + c_n \leq \Omega \\ & c_i \geq 0, \quad \forall i = 1,2,\ldots,n \end{aligned} \quad (9)$$

where $u_i^* = \boldsymbol{S}_i^{\tilde{T}}\mathbf{1}, \forall i = 1,2,\ldots,n$ is the optimal solution obtained by Algorithm 1. To solve problem (9), we can utilize off-the-shelf solvers such as Gurobi.

Numerical Experiments

We utilize data from the Los Angeles region to evaluate the performance of our proposed incentive model. Due to existing multiple routes between most origin-destination (OD) pairs, the Los Angeles area is a suitable region for evaluation. Moreover, the Archived Data Management System (ADMS) developed by researchers at the University of Southern California collects, archives, and integrates a variety of transportation datasets from Los Angeles, Orange, San Bernardino, Riverside, and Ventura Counties. ADMS contains real-time traffic data from 9500 highway and arterial loop detectors that are measured every 30 seconds and 1 minute respectively. While ADMS contains the traffic data of all highways, it does not contain the individuals' OD routing and OD information. Thus, we utilize our network flow information to make estimation of the origin-destination (OD) matrix. Row i and column j of the OD matrix correspond to the origin i and destination j respectively and its element is the number of drivers going from point i to point j. The OD matrix estimation problem is under-determined [59]-[61]. There are two categories of OD matrices: static and dynamic [61]. The majority of current dynamic OD estimation (DODE) methods are not computationally efficient for our data because of its high resolution. Moreover, many studies use a given data of the OD matrix [63]-

[66] to which we did not have access. Due to these challenges, we employ the algorithm provided by [51] for OD estimation.

Simulation Model

In our numerical experiments, we evaluate the performance of our incentivization method using data of Los Angeles area. More specifically, we first extract the sensor information including the location of the sensors from the Archived Data Management System (ADMS). Then, we extract the speed data and volume data of the selected sensors from ADMS. To create the graph of the network, we select sensors on ramps and highway intersections as nodes of the graph and use the data of sensors between them to create data of the connecting links. For distance of the nodes, we used Google Maps API and location data of the nodes. Figure 2 depicts the region by which we create our graph network. It has 12 nodes, 32 links, and 288.1 miles of road. Next, we use the created network graph and its speed data and volume data to estimate the OD pairs by the provided algorithm in [51]. The total number of estimated incoming drivers in each time interval is illustrated in Figure 3. For each OD pair, we find up to 3 different routing choices. First, we find the shortest path for each OD pair. Next, we remove the links in this route and find the second shortest path if it exists, and we do the same process for the third route. We model a region (Figure 2) which includes highways near Downtown Los Angeles and its neighborhood. In our network, the number of OD pairs is 144 and there are 270 paths between them in total.

We focus on incentivizing the organizations to change their behavior for the 7 AM to 8 AM interval (which is the rush hour based on the estimated number of incoming drivers in Figure 2. Although we have selected 7 AM to 8 AM as the incentivization time period, we also include 8 AM to 8:30 AM in our experiments because some of drivers entering between 7 AM and 8 AM may not finish their route before 8 AM. To track the effect of these drivers on the total travel time of the system, we include traffic flow from 8 AM to 8:30 AM in our analysis as well. The estimated total number of drivers incoming to the system between 6 AM to 9 AM by the OD estimation algorithm is depicted in Figure 3. The total number of drivers entering the system between 7 AM and 8:30 AM is 11985.

For our baseline, we use the volume of the network at the User Equilibrium (UE). Algorithm 2 in Appendix “User Equilibrium (UE) Algorithm” computes the volume of the network at UE. Algorithm 2 returns the matrix $\boldsymbol{R}_{\mathrm{UE}}$ and route travel time vector $\boldsymbol{\delta}_{\mathrm{UE}}$ at User Equilibrium given the volume (historical data) and OD estimation as inputs. To compute the cost of organizations' incentivization, we need to know the route travel times when drivers have made decisions based on the UE route travel time $\boldsymbol{\delta}_{\mathrm{UE}}$ Hence, we compute the new volume vector $\boldsymbol{v}_{\mathrm{new}} = \boldsymbol{R}_{\mathrm{UE}}\boldsymbol{S}_{\mathrm{UE}}\mathbf{1}$ where $\boldsymbol{S}_{\mathrm{UE}}$ is the assignment of drivers to the fastest route based on UE route travel time vector $\boldsymbol{\delta}_{\mathrm{UE}}$. Using the BPR function, volume vector $\boldsymbol{v}_{\mathrm{new}}$, and $\boldsymbol{\delta}_{\mathrm{UE}}$, we compute $\boldsymbol{\delta}$ that denotes the travel time of the routes if drivers make decision based on $\boldsymbol{\delta}_{\mathrm{UE}}$ and $\boldsymbol{\eta}$ denotes the minimum travel time between the different OD pairs.

Numerical Results

In this subsection, using our model and algorithm, we study the impact of organization incentivization for different budget values, number of organizations, VOTs, and percentage of drivers who are employed by the organizations in the incentivization program. The remaining drivers are assumed to be background drivers who follow the $\boldsymbol{\delta}_{\mathrm{UE}}$:

- Scenario I: Among the drivers entering the system between 7 AM and 8 AM, 10% of them (i.e., 812 drivers) belong to the organizations that we can incentivize.
- Scenario II: Among the drivers entering the system between 7 AM and 8 AM, 20% of them (i.e., 1624 drivers) belong to the organizations that we can incentivize.

Drivers in each organization are selected uniformly at random and all selected drivers of Scenario I are included in Scenario II to have a fair comparison between the two scenarios. Our base Value of Time (VOT) is derived from the estimation of [67] which is $2.63 per minute or $157.8 per hour. Default number of organizations in our experiments is 10.

The percentage of travel time decrease with incentivization as compared to a system with when no incentivization scheme is used are presented in Figure 4 for both Scenarios I and II. The no incentivization system solution basically assumes all drivers as background drivers. We observe that by increasing the available budget, the amount of decrease in travel time increases (as expected). This decrease is more in larger budgets in Scenario II because the model has access to more drivers to select and has more flexibility to recommend alternative routes. For the purpose of sensitivity analysis, we also provide travel time decrease for both Scenario I and II with a different VOT of $\frac{157.8}{2} = 78.$ per hour in Figure 4. The comparison of results for different VOTs in Figure 4 shows that for a very large budget, the decrease in travel time is almost similar. This is because none of the models utilize the entire budget at a $10,000 budget. However, with a smaller VOT and the budget of $2000 there is a large gap between Scenarios I and II because Scenario II has access to more drivers to deviate. In our plots, the budget of $0 shows the case when the incentivization platform is absent. In Figure 5, we present the total incentivization cost for different budgets in both Scenario I and Scenario II when there are 10 organizations in the system. This cost increases when the available budget is more, as expected. This pattern shows that the platform can utilize the resources when it has access to more money. Figure 6 shows the cost per deviated driver for the two scenarios. Although the gap between the total cost of Scenarios I and II is small, the cost per driver is significantly smaller in Scenario II. As can be seen in this figure, the cost per driver is smaller in Scenario II due to more flexibility the model has in choosing the drivers efficiently. As Table 1 shows, the number of selected drivers in Scenario II is larger because there are more drivers for selection.

Table 1. Distribution of the number of drivers that were assigned to an alternative route.

Scenario	Budget			
	$200	$800	$2000	$10000
I	33	57	85	130
II	51	94	130	222

The number of organizations in the system can alter the total travel time decrease and the cost. Figure 7 illustrates the percentage decrease of travel time and total cost when there are different number of organizations in the system. As an extreme case, we also include the case that each organization contains one driver (i.e., we incentivize individuals rather than organizations). In Figure 7, we observe larger cost for reducing the same amount of travel time decrease when there are more organizations in the system. The intuitive reason behind this observation is as follows. For each organization, after incentivization, some drivers lose time and some gain travel time. At the organization level, the time changes of drivers can cancel each other out, and hence we may not need to significantly compensate the organization. When the number of organizations increases, the cancelling effect becomes weaker and the incentivization costs more. This also explains why incentivizing organizations is much more cost efficient than incentivizing individual drivers.

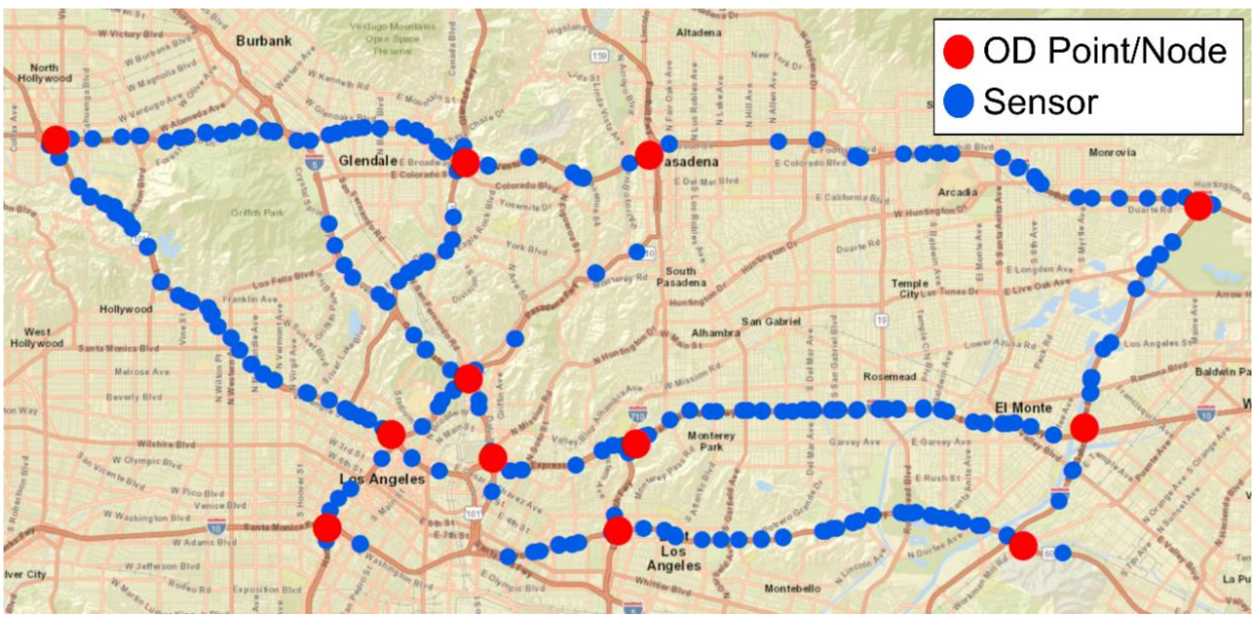


Figure 2. Studied region.

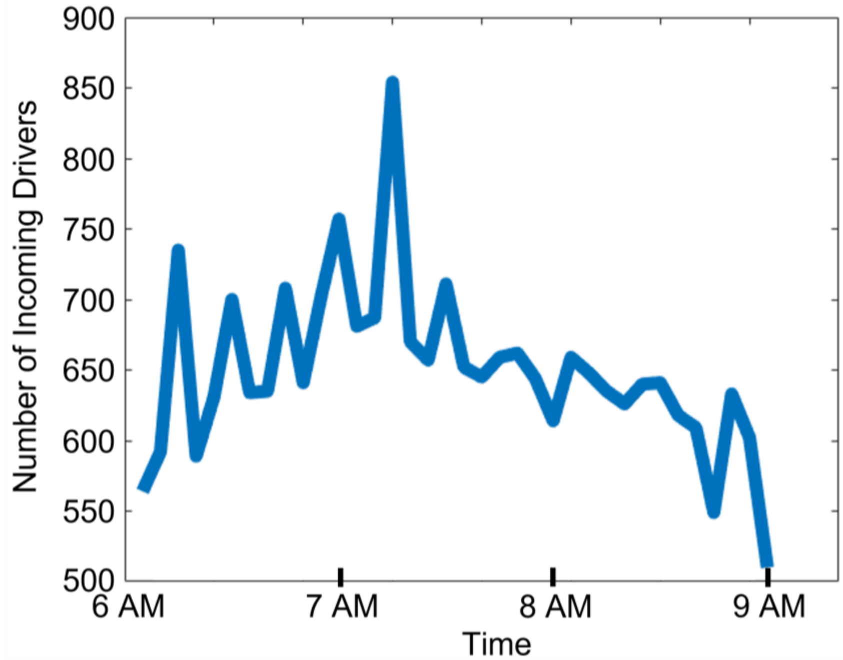


Figure 3. Total estimated number of drivers entering the system (in 5 minute intervals).

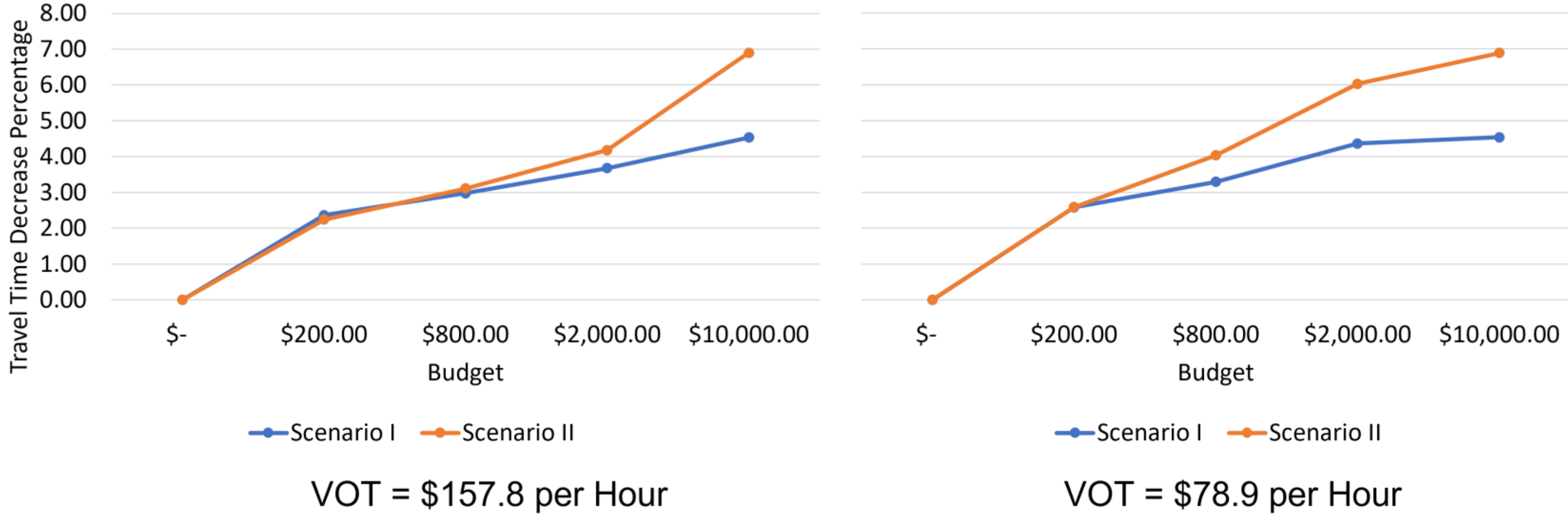


Figure 4. Percentage of travel time decrease with different budgets at different VOTs.

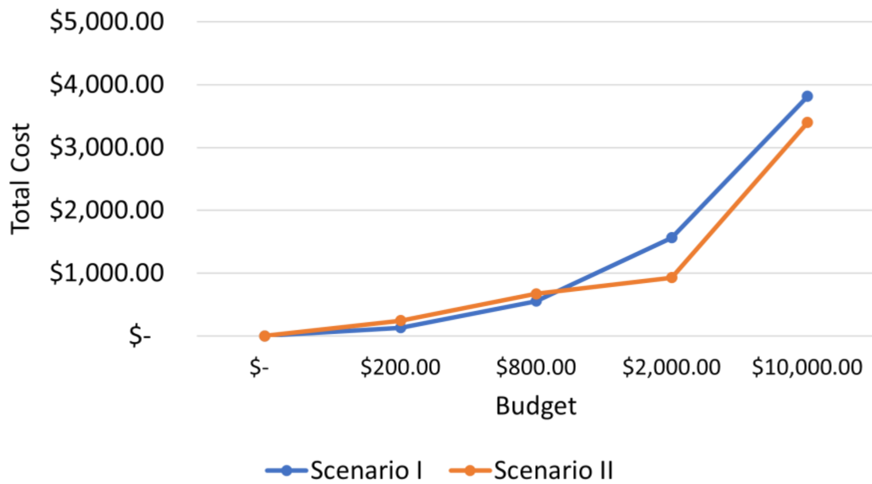


Figure 5. Total cost of incentivization of 10 organizations with different budgets in Scenario I and II and VOT=$157.8/Hour.

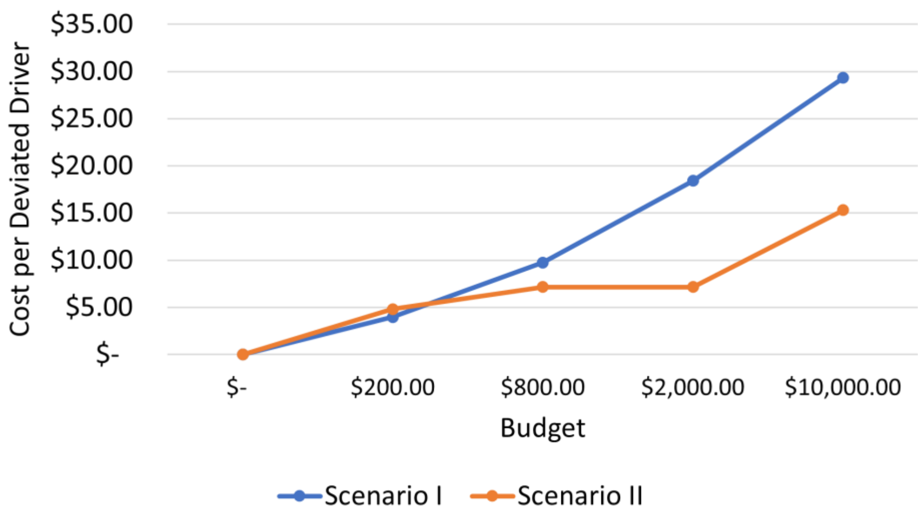


Figure 6. Cost of incentivization per Deviated Drivers of 10 organizations with different budgets in Scenario I and II and VOT=$157.8/Hour.

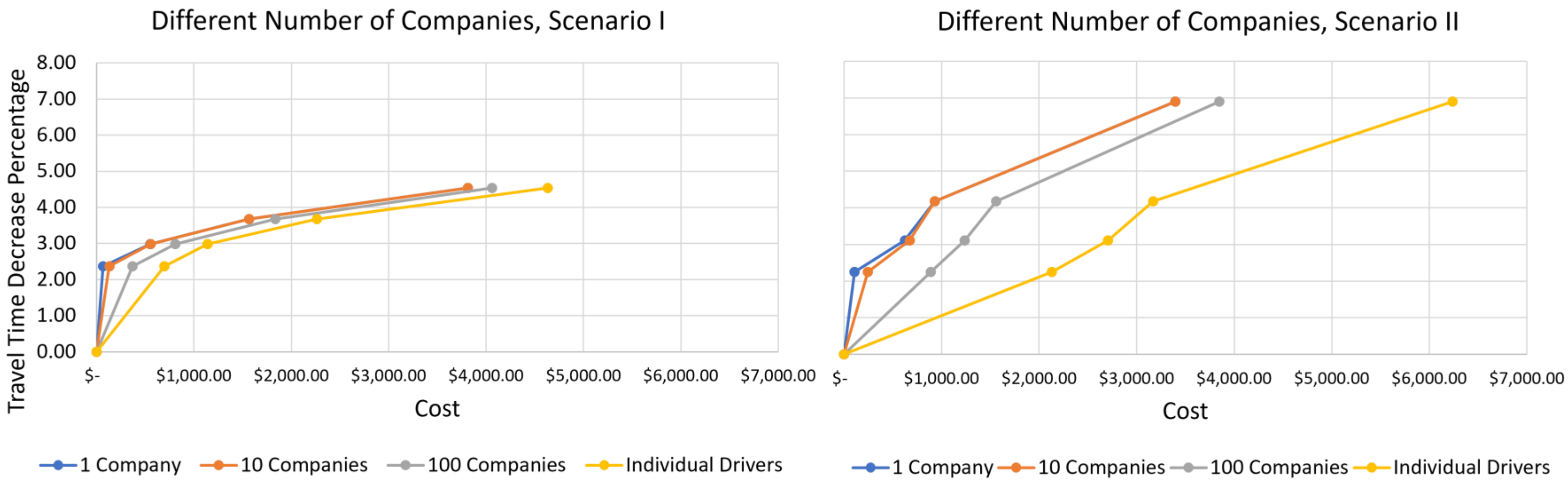


Figure 7. Cost of incentivization and travel time decrease percentage for different number of organizations in Scenario I and II and VOT=$157.8/Hour.

Conclusion

In this project, we study the problem of incentivizing organizations to reduce traffic congestion. To this end, we developed a mathematical model and provided an algorithm for offering organization-level incentives. In our framework, a central planner collects the origin-destination and routing information of the organizations. Then, the central planner utilizes this information to offer incentive packages to organizations to incentivize a system-level optimal routing strategy. In particular, we focused on minimizing the total travel time of the network. However, other utilities can be used in our framework. Finally, we employed data from Archived Data Management System (ADMS) to evaluate the performance of our model and algorithm. A 6.90% reduction in the total travel time of the network was reached by our framework in the experiments. More importantly, we observed that incentivizing companies/organization is more cost efficient than incentivizing individual drivers. As future work, it is important to study the effect of incentivization to change the departure time or the demand. This is particularly relevant in future mobility services because many of them, such as delivery services, are flexible in terms of trip time to a certain degree. In addition, we can consider stochastic nature of making decision in routing by individual drivers. Moreover, we can extend the incentivization framework to the case that not all the organizations accept their received offer. All the codes for this project can be found in [70]-[71].

References

[1] Michael J. Sprung, Matthew Chambers, and Sonya Smith-Pickel. chapter Transportation Statistics Annual Report 2018. Dec 2018. Tech Report.

[2] “INRIX 2019 global traffic scorecard,” https://inrix.com/scorecard/, accessed: 2021-02-1

[3] “INRIX 2019 Los Angeles traffic scorecard,” https://inrix.com/press-releases/2019-traffic-scorecard-us/, accessed: 2021-12-28.

[4] Congestion mitigation and air quality improvement program: Assessing 10 years of experience. http://www.trb.org/main/blurbs/160904.aspx.

[5] Health Effects Institute. Panel on the Health Effects of Traffic-Related Air Pollution. Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects. Number 17. Health Effects Institute, 2010.

[6] Kai Zhang and Stuart Batterman. Air pollution and health risks due to vehicle traffic. Science of the total Environment, 450:307–316, 2013.

[7] Dwight A Hennessy and David L Wiesenthal. Traffic congestion, driver stress, and driver aggression. Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 25(6):409–423, 1999.

[8] Melvin L Selzer and Amiram Vinokur. Life events, subjective stress, and traffic accidents. American Journal of Psychiatry, 131(8):903–906, 1974.

[9] Cambridge Systematics. Traffic congestion and reliability: Trends and advanced strategies for congestion mitigation. Technical report, United States. Federal Highway Administration, 2005.

[10] Robert B Noland and Lewison L Lem. A review of the evidence forinduced travel and changes in transportation and environmental policy in the US and the UK. Transportation Research Part D: Transport and Environment, 7(1): 1–26, 2002.

[11] Susan Handy and Marlon G Boarnet. Impact of highway capacity and induced travel on passenger vehicle use and greenhouse gas emissions. California Environmental Protection Agency, Air Resources Board, Retrieved August 28: 2015, 2014.

[12] Arthur C Pigou. The economics of welfare Macmillan and Co. London, United Kingdom, 1920.

[13] Frank H Knight. Some fallacies in the interpretation of social cost. The Quarterly Journal of Economics, 38(4):582–606, 1924.

[14] Erik Verhoef, Michiel C.J. Bliemer, Linda Steg, and Bert van Wee, editors. Pricing in Road Transport. Number 4192 in Books. Edward Elgar Publishing, July 2008.

[15] Dirk Hendrik Van Amelsfort. Behavioural responses and network effects of time-varying road pricing. 2009.

[16] Nan Zheng, Guillaume Rérat, and Nikolas Geroliminis. Time-dependent area-based pricing for multi- modal systems with heterogeneous users in an agent-based environment. Transportation Research Part C: Emerging Technologies, 62:133–148, 2016.

[17] Carlos F Daganzo and Lewis J Lehe. Distance-dependent congestion pricing for downtown zones. Transportation Research Part B: Methodological, 75:89–99, 2015.

[18] Shu Zhang, Ann M Campbell, and Jan F Ehmke. Impact of congestion pricing schemes on costs and emissions of commercial fleets in urban areas. Networks, 73(4):466–489, 2019.

[19] Jasper Knockaert, Yin-Yen Tseng, Erik T Verhoef, and Jan Rouwendal. The Spitsmijden experiment: A reward to battle congestion. Transport Policy, 24:260–272, 2012.

[20] Charles Raux and Stéphanie Souche. The acceptability of urban road pricing: A theoretical analysis applied to experience in lyon. Journal of Transport Economics and Policy (JTEP), 38(2):191–215, 2004.

[21] Ziyuan Gu, Zhiyuan Liu, Qixiu Cheng, and Meead Saberi. Congestion pricing practices and public acceptance: A review of evidence. Case Studies on Transport Policy, 6(1):94–101, 2018.

[22] Ziyuan Gu, Sajjad Shafiei, Zhiyuan Liu, and Meead Saberi. Optimal distance-and time-dependent area-based pricing with the network fundamental diagram. Transportation Research Part C: Emerging Technologies, 95:1–28, 2018.

[23] Erik Verhoef, Peter Nijkamp, and Piet Rietveld. Tradeable permits: their potential in the regulation of road transport externalities. Environment and Planning B: Planning and Design, 24(4):527–548, 1997.

[24] Wenbo Fan and Xinguo Jiang. Tradable mobility permits in roadway capacity allocation: Review and appraisal. Transport Policy, 30:132–142, 2013.

[25] John Thøgersen and Berit Møller. Breaking car use habits: The effectiveness of a free one-month travelcard. Transportation, 35(3):329–345, 2008.

[26] Guangmin Wang, Ziyou Gao, Meng Xu, and Huijun Sun. Models and a relaxation algorithm for continuous network design problem with a tradable credit scheme and equity constraints. Computers & Operations Research, 41:252–261, 2014.

[27] Theodore Tsekeris and Stefan Voß. Design and evaluation of road pricing: state-of-the-art and methodological advances. NETNOMICS: Economic Research and Electronic Networking, 10(1):5–52, 2009.

[28] Hideki Fukui. An empirical analysis of airport slot trading in the United States. Transportation Research Part B: Methodological, 44(3):330–357, 2010.

[29] Carlos Lima Azevedo, Ravi Seshadri, Song Gao, Bilge Atasoy, Arun Prakash Akkinepally, Eleni Christofa, Fang Zhao, Jessika Trancik, and Moshe Ben-Akiva. Tripod: sustainable travel incentives with prediction, optimization, and personalization. In Transportation Research Board 97th Annual Meeting, 2018.

[30] Nico Dogterom, Dick Ettema, and Martin Dijst. Tradable credits for managing car travel: a review of empirical research and relevant behavioural approaches. Transport Reviews, 37(3):322–343, 2017.

[31] Jack Williams Brehm. A Theory of Psychological Reactance. New York: Academic Press, 1966.

[32] David M Kreps. Intrinsic motivation and extrinsic incentives. The American Economic Review, 87(2):359–364, 1997.

[33] Kent C Berridge. Reward learning: Reinforcement, incentives, and expectations. In Psychology of Learning and Motivation, volume 40, pages 223–278. Elsevier, 2000.

[34] Deepak Merugu, Balaji S Prabhakar, and N Rama. An incentive mechanism for decongesting the roads: A pilot program in Bangalore. In Proc. of ACM NetEcon Workshop. Citeseer, 2009.

[35] Jia Shuo Yue, Chinmoy V Mandayam, Deepak Merugu, Hossein Karkeh Abadi, and Balaji Prabhakar. Reducing road congestion through incentives: a case study. 2015.

[36] Michiel Bliemer, Matthijs Dicke-Ogenia, and Dick Ettema. Rewarding for avoiding the peak period: A synthesis of three studies in the Netherlands. In European Transport Conference 2009. Citeseer, 2009.

[37] Xianbiao Hu, Yi-Chang Chiu, and Lei Zhu. Behavior insights for an incentive-based active demand management platform. International Journal of Transportation Science and Technology, 4(2):119–133, 2015.

[38] Vivek Kumar, Chandra R Bhat, Ram M Pendyala, Daehyun You, Eran Ben-Elia, and Dick Ettema. Impacts of incentive-based intervention on peak period traffic: experience from the Netherlands. Transportation Research Record, 2543(1):166–175, 2016.

[39] Graham Currie. Free fare incentives to shift rail demand peaks–medium-term impacts. Technical report, 2011.

[40] Zheng Zhang, Hidemichi Fujii, and Shunsuke Managi. How does commuting behavior change due to incentives? an empirical study of the Beijing subway system. Transportation Research Part F: Traffic Psychology and Behaviour, 24:17–26, 2014.

[41] Satoshi Fuji and Ryuichi Kitamura. What does a one-month free bus ticket do to habitual drivers. Transportation, 30:81–95, 2003.

[42] Sebastian Bamberg, Icek Ajzen, and Peter Schmidt. Choice of travel mode in the theory of planned behavior: The roles of past behavior, habit, and reasoned action. Basic and Applied Social Psychology, 25(3):175–187, 2003.

[43] J Knockaert, M Bliemer, D Ettema, D Joksimovic, A Mulder, J Rouwendal, and D Van Amelsfort. Experimental design and modelling Spitsmijden. Utrecht, Consortium Spitsmijden, 2007.

[44] Ali Ghafelebashi, Meisam Razaviyayn, and Maged Dessouky, “Congestion reduction via personalized incentives,” 2021.

[45] Amy He, “People continue to rely on maps and navigational apps,” https://www.emarketer.com/content/people-continue-to-rely-on-maps-and-navigational-apps-emarketer-forecasts-show, accessed: 2021-12

[46] Riley Panko, “The popularity of google maps: Trends in navigation apps in 2018,” https://themanifest.com/app-development/trends-navigation-apps, accessed: 2021-12-12

[47] Noortje Groot, Bart De Schutter, and Hans Hellendoorn, “Toward system-optimal routing in traffic networks: A reverse Stackelberggame approach,”IEEETransactionsonIntelligentTransportationSystems, vol. 16, no. 1, pp. 29–40, 2014

[48] Meisam Razaviyayn, Zhi-Quan Luo, Paul Tseng and Jong-Shi Pang, “A Stackelberg game approach to distributed spectrum management,”Mathematicalprogramming, vol. 129, no. 2, pp. 197–224, 2011.

[49] Chenfeng Xiong, Mehrdad Shahabi, Jun Zhao, Yafeng Yin, Xuesong Zhou, and Lei Zhang. An integrated and personalized traveler information and incentive scheme for energy efficient mobility systems. Transportation Research Part C: Emerging Technologies, volume 113, pages 57–73, 2020.

[50] United States. Bureau of Public Roads. Traffic assignment manual for application with a large, high speed computer, volume 37. US Department of Commerce, Bureau of Public Roads, Office of Planning, Urban Planning Division, 1964.

[51] Wei Ma and Zhen Sean Qian. Estimating multi-year 24/7 origin-destination demand using high-granular multi-source traffic data. Transportation Research Part C: Emerging Technologies, 96:96–121, 2018.

[52] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014.

[53] Michael Grant and Stephen Boyd. Graph implementations for nonsmooth convex programs. In V. Blondel, S. Boyd, and H. Kimura, editors, Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences, pages 95–110. Springer-Verlag Limited, 2008. http://stanford.edu/~boyd/graph_dcp.html.

[54] Stephen Boyd, Neal Parikh, and Eric Chu. Distributed optimization and statistical learning via the alternating direction method of multipliers. Now Publishers Inc, 2011.

[55] Daniel Gabay and Bertrand Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Computers & Mathematics with Applications, 2(1):17–40, 1976.

[56] Roland Glowinski and A Marroco. Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de dirichlet non linéaires. ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique, 9(R2):41–76, 1975.

[57] Mingyi Hong, Zhi-Quan Luo, and Meisam Razaviyayn. Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems. SIAM Journal on Optimization, 26(1):337– 364, 2016.

[58] Bingsheng He and Xiaomingzuylen Yuan, "On the o(1/n) convergence rate of the douglas–rachford alternating direction method,"SIAMJournalonNumericalAnalysis, vol. 50, no. 2, pp. 700–709, 2012

[59] Henk J Van Zuylen and Luis G Willumsen. The most likely trip matrix estimated from traffic counts. Transportation Research Part B: Methodological, 14(3):281–293, 1980.

[60] Ennio Cascetta. Estimation of trip matrices from traffic counts and survey data: a generalized least squares estimator. Transportation Research Part B: Methodological, 18(4-5):289–299, 1984.

[61] Michael GH Bell. The real time estimation of origin-destination flows in the presence of platoon dispersion. Transportation Research Part B: Methodological, 25(2-3):115–125, 1991.

[62] Sharminda Bera and KV Rao. Estimation of origin-destination matrix from traffic counts: the state of the art. European Transport Trasporti Europei, 2011.

[63] Panchamy Krishnakumari, Hans van Lint, Tamara Djukic, and Oded Cats. A data driven method for od matrix estimation. Transportation Research Part C: Emerging Technologies, 2019.

[64] Stefano Carrese, Ernesto Cipriani, Livia Mannini, and Marialisa Nigro. Dynamic demand estimation and prediction for traffic urban networks adopting new data sources. Transportation Research Part C: Emerging Technologies, 81:83–98, 2017.

[65] Marialisa Nigro, Ernesto Cipriani, and Andrea del Giudice. Exploiting floating car data for time- dependent origin–destination matrices estimation. Journal of Intelligent Transportation Systems, 22(2):159–174, 2018.

[66] JinYoung Kim, Fumitaka Kurauchi, Nobuhiro Uno, Takeshi Hagihara, and Takehiko Daito. Using electronic toll collection data to understand traffic demand. Journal of Intelligent Transportation Systems, 18(2):190–203, 2014.

[67] Juan Camilo Castillo, "Who benefits from surge pricing?" Available at SSRN3245533, 2020.

[68] Michael Sheetz, "The popularity of google maps: Trends in navigation apps in 2018," https://www.cnbc.com/2019/12/12/analyst-amazon-delivering-nearly-half-its-packages-instead-of-ups-fedex.html, accessed: 2021-12-28.

[69] Esri, "ArcGIS." [Online]. Available: https://www.esri.com/

[70] "Github codes," https://github.com/ghafeleb/Incentive_Systems_for_New_Mobility_Services, accessed: 2021-12-30

[71] "Incentive systems for new mobility services codes," https://doi.org/10.7910/DVN/7VI4LX, accessed: 2021-12-3

Data Summary

Products of Research

Research products of this work will be submitted in peer-reviewed journal articles, book chapters and/or conference proceedings targeted towards the transportation science research community. All the resulting code for offering incentives are shared with Caltrans and NCST through HARVARD Dataverse platform (https://doi.org/10.7910/DVN/7VI4LX). In addition, the codes will be shared on (https://github.com/ghafeleb/Incentive_Systems_for_New_Mobility_Services). Our goal is to make the use of our algorithm and its implementation convenient to the transportation science research community. The data format and the content of files shared on HARVARD Dataverse is described next.

Data Format and Content

1. data\\capacity\\Mar2May_2018_new_5-22_link_capacity_region_toy.csv: capacity of links in CSV format
2. data\\capacity\\Mar2May_2018_new_5-22_link_s_0_region_toy.csv: free flow speed of links in miles per hour in CSV format
3. Mar2May_2018_new_5-22_link_tt_0_minutes_region_toy.csv: free flow travel time of links in minutes in CSV format
4. DPFE_files\\Q_vector\\2021_11_09_08_55_25\\python3\\2018-05-01.pickle: OD estimation in pickle format
5. DPFE_files\\R_matrix\\2021_11_09_08_55_25\\2018-05-01.pickle: matrix R in pickle format for Python 2
6. DPFE_files\\R_matrix\\2021_11_09_08_55_25\\2018-05-01_pck.pickle: matrix R in pickle format for Python 3
7. DPFE_files\\tt\\2021_11_09_08_55_25\\2018-05-01_pck.pickle: travel time of paths in pickle format for Python 3
8. Data\\region_toy\\link_length_meter_region_toy_original.csv: length of links in meter in CSV format
9. Data\\region_toy\\link_length_mile_region_toy_original.csv: length of links in miles in CSV format
10. data\\speed_volume\\Mar2May_2018_region_toy_AVG5_5-22_with_linkID_pad\\my_link_avg_count_data_AVG5min_5-22_region_toy_pad.pickle: volume data of links in pickle format
11. data\\speed_volume\\Mar2May_2018_region_toy_AVG5_5-22_with_linkID_pad\\my_link_avg_spd_data_AVG5min_5-22_region_toy_pad.pickle: speed data of links in CSV format
12. data\\YAML\\region_toy_create_graph.yaml: properties of data used to run create_graph.py code
13. data\\YAML\\region_toy_DataCreator1.yaml: properties of data used to run DataCreator1.py code

14. data\\YAML\\region_toy_DataCreator2.yaml: properties of data used to run DataCreator2.py code
15. data\\YAML\\region_toy_HistoricalData.yaml: properties of data used to run HistoricalData.py code
16. data\\YAML\\region_toy_realCost.yaml: properties of data used to run realCost.py code
17. data\\YAML\\region_toy_runDet.yaml: properties of data used to run runDet.py code

Data Access and Sharing

The codes are free to share and open to the public. In particular, we have uploaded the codes on HARVARD Dataverse platform (https://doi.org/10.7910/DVN/7VI4LX). Moreover, we will share the code on the GitHub repository (https://github.com/ghafeleb/Incentive_Systems_for_New_Mobility_Services).

As input to our codes, we used Archived Data Management System (ADMS) that collects, archives, and integrates a variety of transportation datasets from Los Angeles, Orange, San Bernardino, Riverside, and Ventura Counties. ADMS includes access to real-time traffic datasets from i) 9500 highway and arterial loop detectors providing data approximately every 1 minute, and ii) 2500 bus and train GPS location (AVL) data operating throughout Los Angeles County. We can share a sample of our run upon a reasonable request.

Results of this work will be published in peer-reviewed scientific journals, books published in English, conference proceedings, or as peer-reviewed data reports.

Reuse and Redistribution

USC's policy is to encourage, wherever appropriate, research data to be shared with the general public through internet access. This public access will be regulated by the university in order to protect privacy and confidentiality concerns, as well to respect any proprietary or intellectual property rights. Administrators will consult with the university's legal office to address any concerns on a case-by case basis, if necessary. Terms of use will include requirements of attribution along with disclaimers of liability in connection with any use or distribution of the research data, which may be conditioned under some circumstances.

Appendix

List of Notations

$\mathcal{G}$: Directed graph of the traffic network

$\mathcal{V}$: Set of nodes of graph $\mathcal{G}$ which correspond to major intersections and ramps

$\mathcal{E}$: Set of edges of graph $\mathcal{G}$ which correspond to the set of road segments

$|\mathcal{E}|$: Total number of road segments/edges in the network $\mathcal{G}$ (i.e., the cardinality of the set $\mathcal{E}$)

r: Route vector

T: Time horizon

$|T|$: Number of time units (i.e., the cardinality of T)

v_t: Volume vector of road segments at time t

$\mathcal{N}$: Set of all drivers

$\mathcal{N}_i$: Set of drivers of organization i

$|\mathcal{N}|$: Total number of drivers (i.e., the cardinality of the set $\mathcal{N}$)

$\mathcal{N}_i$: Total number of drivers of organization i (i.e., the cardinality of the set $\mathcal{N}_i$)

$\mathcal{R}_j$: Set of possible route options for driver j

$\mathcal{R}$: Total set of possible route options for all OD pairs

$|\mathcal{R}|$: Number of possible route options (i.e., the cardinality of the set $\mathcal{R}$)

$s_i^{r,j}$: Decision parameter indicates whether route r is assigned to driver j from organization i

T_r: The exact travel time for route r

$F_{tt}(.)$: Total travel time function

ℓ: An index for a link/road in the network which is an edge in graph $\mathcal{G}$

$\theta_{l,t}$: Travel time of link l at time t

$\hat{v}$: The vector of estimated volume of links at different times in the horizon

$\widehat{v_{\ell,t}}$: The $(|\mathcal{E}| \times t + \ell)^{th}$ element of vector $\hat{v}$ representing the volume of ℓ^{th} link at time t

t_0: The free flow travel time of the link

v: The traffic volume of the link

w: The practical capacity of the link

s_i^j: The binary route assignment vector of driver j from organization i

$f_{\mathrm{BPR}}(.)$: BPR function

S: Decision matrix of all drivers

S_i: Decision matrix of drivers of organization i

R: The matrix of probability of a driver being at each link given their route

D: The matrix of route assignment of the OD pairs

q: The vector of number of drivers for each OD pair

q_i: The vector of number of drivers of organization i for each OD pair

$\boldsymbol{\delta}$: The vector of travel time of routes at different times

η: The vector of shortest travel time between different OD pairs at different times

b_i: This vector contains the factors by which the travel time of assigned routes can be larger than shortest travel time of the drivers of organization i

B_i: The matrix of shortest travel time assignment of drivers of organization i

α_i: The Value of Time for organization i

α: The vector of Value of Time for different organizations

c_i: The of cost of incentive offered to organization i

γ_i: Total travel time of organization i in the abscense of incentivization platform

Ω: Budget

$r_{\ell,t}$: The row of matrix R that corresponds to link ℓ at time t

K: The number of OD pairs

e: An edge of graph $\mathcal{G}$ which corresponds to a road segments in the traffic network

Reformulated Optimization Model for ADMM Algorithm

To solve optimization problem (7) efficiently, we present a distributed algorithm based on this reformulation

$$
\begin{aligned}
\min_{\substack{\mathbf{S},\omega,\mathbf{H},\mathbf{W}\\ \mathbf{Z},\mathbf{u},\{\beta_i\}_{i=1}^{n}\\ u_i,\tilde{\beta},c}} \quad & \sum_{\ell=1}^{|\mathcal{E}|}\sum_{t=1}^{|T|} \widehat{v_{\ell,t}}\,\theta_{\ell,t}(\widehat{v_{\ell,t}}\,) \\
& -\frac{\tilde{\lambda}}{2}\sum_{r=1}^{\mathcal{R}}\sum_{t=1}^{|\mathrm{T}|}\sum_{i=1}^{n}(Z_i)_{r,t}((Z_i)_{r,t}-1) \\
s.t. \quad & \mathbf{S}_i\mathbf{1}=\mathbf{u}_i, \quad \forall i=1,2,\dots,n \\
& \boldsymbol{\omega}=\widetilde{\boldsymbol{R}}\boldsymbol{u} \\
& \widetilde{\boldsymbol{D}}\boldsymbol{u}=\boldsymbol{q}, \quad \forall i=1,2,\dots,n \\
& \mathbf{W}_i^{\top}\mathbf{1}=\mathbf{1}, \quad \forall i=1,2,\dots,n \\
& \mathbf{S}_i=\mathbf{W}_i, \quad \forall i=1,2,\dots,n \\
& \mathbf{H}_i^{\top}\boldsymbol{\delta}_{\boldsymbol{p}}+\beta_i=\boldsymbol{b}_i\odot\mathbf{B}_i\eta, \quad \forall i=1,2,\dots,n \\
& \mathbf{S}_i=\mathbf{H}_i, \quad \forall i=1,2,\dots,n \\
& \beta_i\geq 0, \quad \forall i=1,2,\dots,n \\
& \mathbf{Z}_i\in\,[0,1]^{(|\mathcal{R}|\cdot|\mathbf{T}|)\times(|\mathcal{N}_i|)}, \quad \forall i=1,2,\dots,n \\
& \tilde{I}\tilde{c}=\alpha\odot(\Delta u-\gamma) \\
& \tilde{c}\geq 0, \quad \tilde{c}^{\top}\tilde{\mathbf{1}}+\tilde{\beta}=\Omega, \quad \tilde{\beta}\geq 0 \\
& \mathbf{S}_i=\mathbf{Z}_i, \quad \forall i=1,2,\dots,n
\end{aligned}
\tag{10}
$$

Details of Alternating Direction Method of Multipliers (ADMM)

Before explaining the steps of our proposed algorithm, let us first explain Alternating Direction Method of Multipliers (ADMM), which is a main building block of our framework.

Review of ADMM

$$
\min_{w,z} h(w)+g(z) \qquad s.t. \;\; Aw+Bz\;=\;c,
$$

where $w\in R^{d_1}, z\in R^{d_2}$, $c\in R^k$, $A\in R^{k\times d_1}$, and $B\in R^{k\times d_2}$. By forming the augmented Lagrangian function

$$
\mathcal{L}(w,z,\lambda)\triangleq h(w)+g(z)+\langle\lambda,Aw+Bz-c\rangle+\frac{\rho}{2}\|Aw+Bz-c\|_2^2
$$

each iteration of ADMM applies alternating minimization to the primal variables and gradient ascent to the dual variables. More precisely, at iteration r, ADMM uses the update rules:

Primal Update: $$w^{r+1} = \arg\min_{w} \mathcal{L}(w, z^r, \lambda^r),$$
$$z^{r+1} = \arg\min_{z} \mathcal{L}(w^{r+1}, z, \lambda^r) \tag{11}$$
Dual Update: $$\lambda^{r+1} = \lambda^r + \rho(Aw^{r+1} + Bz^{r+1} - c)$$

This algorithm is well studied in the optimization literature (see [54] for a monograph on the use of this algorithm in convex distributed optimization and [57] for its use in non-convex continuous optimization).

ADMM for Solving (7)

Let

$$
\begin{aligned}
&\mathcal{L}(\{\mathrm{S}_i\}_{i=1}^n, \{H_i\}_{i=1}^n, \{\mathrm{W}_i\}_{i=1}^n, \{\mathrm{Z}_i\}_{i=1}^n, \omega, \tilde{\beta}, \{\beta_i\}_{i=1}^n, \tilde{c}, \{\mathrm{u}_i\}_{i=1}^n) \\
&\triangleq F_u(\omega) - \frac{\tilde{\lambda}}{2}\sum_{i=1}^{n}\sum_{j=1}^{|\mathcal{N}_i|}\sum_{\ell=1}^{|\mathcal{E}|}\sum_{t=1}^{|T|}(\mathrm{Z}_j)_{i,(r,t)}((\mathrm{Z}_j)_{i,(r,t)}) \\
&+ \sum_{i=1}^{n} \mathbb{I}_{\mathbb{R}_+^{|\mathcal{N}_i|}}(\beta_i) + \mathbb{I}_{\mathbb{R}_+^{2n}}(\tilde{c}) + \mathbb{I}_{\mathbb{R}_+}(\tilde{\beta}) + \mathbb{I}_{[0,1]}((\mathrm{Z}_j)_{i,(r,t)}) \\
&+ \sum_{i=1}^{n}\langle \lambda_{1,i}, \mathrm{S}_i\mathbf{1} - u_i\rangle + \langle \lambda_2, \omega - \tilde{R}u\rangle + \sum_{i=1}^{n}\langle \lambda_{3,i}, \mathrm{D}u_i - q_i\rangle \\
&+ \sum_{i=1}^{n}\langle \lambda_{4,i}, \mathrm{W}_i\mathbf{1} - \mathbf{1}\rangle + \sum_{i=1}^{n}\langle \Lambda_{5,i}, \mathrm{S}_i - W_i\rangle \\
&+ \sum_{i=1}^{n}\langle \lambda_{6,i}, \mathrm{H}_i^\top\delta_p - \beta_i - b_i\cdot(B_i\eta)\rangle + \langle \lambda_7, (\Delta\tilde{\alpha})^\top u - \tilde{\alpha}\delta - \tilde{I}\tilde{c}\rangle \\
&+ \sum_{i=1}^{n}\langle \Lambda_{8,i}, S_i - H_i\rangle + \langle \lambda_9, \tilde{c}^\top\tilde{\mathbf{1}} + \tilde{\beta} - \Omega\rangle + \Lambda_{10,i}, \mathrm{S}_i - Z_i\rangle \\
&+ \frac{\rho}{2}\sum_{i=1}^{n}\|S_i\mathbf{1} - u_i\|^2 + \frac{\rho}{2}\sum_{i=1}^{n}\|S_i\mathbf{1} - u_i\|^2 + \frac{\rho}{2}\|\omega - \tilde{R}u\|^2 \\
&+ \frac{\rho}{2}\sum_{i=1}^{n}\|Du_i - q_i\|^2 + \frac{\rho}{2}\sum_{i=1}^{n}\|W_i\mathbf{1} - \mathbf{1}\|^2 + \frac{\rho}{2}\sum_{i=1}^{n}\|S_i - W_i\|^2 \\
&+ \frac{\rho}{2}\sum_{i=1}^{n}\|\mathrm{H}_i^\top\delta_p + \beta_i - b_i\cdot(B_i\eta)\|^2 + \frac{\rho}{2}\|(\Delta\tilde{\alpha})^\top u - \tilde{\alpha}\delta - \tilde{I}\tilde{c}\|^2
\end{aligned} \tag{12}
$$

$$+\frac{\rho}{2}\sum_{i=1}^{n}\|S_i - H_i\|^2 + \frac{\rho}{2}\|\tilde{c}^{\top}\mathbf{1} + \tilde{\beta} - \Omega\|^2 + \frac{\rho}{2}\sum_{i=1}^{n}\|S_i - Z_i\|^2$$

be the augmented Lagrangian function of (7) with the set of Lagrange multipliers $\{\{\lambda_1\}_{i=1}^{n}, \lambda_2, \dots, \{\Lambda_{10}\}_{i=1}^{n}$ and $\rho > 0$ be the primal penalty parameter. Then, ADMM solves (7) by the following iterative scheme

$$\omega_{(\ell,t)}^{t+1} = \underset{\omega_{(\ell,t)}}{\operatorname{argmin}} \;\; \omega_{(\ell,t)}\theta_{\ell,t}\big(\omega_{(\ell,t)}\big) + (\lambda_2^t)_{\ell,\hat{t}}\left(\omega_{\ell,\hat{t}} - r_{\ell,\hat{t}}\left(\sum_{i=1}^{n} u_i\right)\right) + \frac{\rho}{2}\left(\omega_{\ell,\hat{t}} - r_{\ell,\hat{t}}\left(\sum_{i=1}^{n} u_i^t\right)\right)$$

$$S_i^{t+1} = \underset{S_i}{\operatorname{argmin}} \;\; \langle \lambda_{1,i}^t, S_i\mathbf{1} - u_i^t\rangle + \langle \Lambda_{5,i}^t, S_i - W_i^t\rangle + \langle \Lambda_{8,i}^t, S_i - H_i^t\rangle + \langle \Lambda_{10,i}^t, S_i - Z_i^t\rangle$$
$$+ \frac{\rho}{2}\|S_i\mathbf{1} - u_i^t\|^2 + \frac{\rho}{2}\sum_{i=1}^{n}\|S_i\mathbf{1} - u_i^t\|^2 + \frac{\rho}{2}\|S_i - W_i^t\|^2 + \frac{\rho}{2}\|S_i - H_i^t\|^2$$
$$+ \frac{\rho}{2}\|S_i - Z_i^t\|^2, \quad \forall i = 1,2,\dots,n$$

$$\beta_i^{t+1} = \underset{\beta_i}{\operatorname{argmin}} \;\; \mathbb{I}_{\mathbb{R}_+^{|\mathcal{N}_i|}}(\beta_i) + \langle \lambda_{6,i}^t, H_i^{t\top}\delta_p + \beta_i^t - b_i \cdot (B_i\eta)\rangle + \frac{\rho}{2}\left\|H_i^{t\top}\delta_p + \beta_i^t - b_i \cdot B_i\eta\right\|^2, \quad \forall i$$
$$= 1,2,\dots,n$$

$$\tilde{c}^{t+1} = \underset{\tilde{c}}{\operatorname{argmin}} \;\; \mathbb{I}_{\mathbb{R}_+^{2n}}(\tilde{c}) + \langle \lambda_7^t, (\Delta\tilde{\alpha})^{\top}u^t - \tilde{\alpha}\delta - \tilde{\mathrm{I}}\tilde{c}\rangle + \langle \lambda_9^t, \tilde{c}^{\top}\tilde{\mathbf{1}} + \tilde{\beta} - \Omega\rangle$$
$$+ \frac{\rho}{2}\left\|(\Delta\tilde{\alpha})^{\top}u^t - \tilde{\alpha}\delta - \tilde{\mathrm{I}}\tilde{c}\right\|^2 + \frac{\rho}{2}\left\|\tilde{c}^{\top}\mathbf{1} + \tilde{\beta}^t - \Omega\right\|^2$$

$$u^{t+1} = \underset{u}{\operatorname{argmin}} \;\; \langle \lambda_1^t, \tilde{u}^{t+1} - u\rangle + \langle \lambda_2^t, \omega - \tilde{R}u\rangle + \langle \lambda_2^t, \tilde{D}u - q\rangle + \langle \lambda_7^t, (\Delta\tilde{\alpha})^{\top}u - \tilde{\alpha}\delta - \tilde{\mathrm{I}}\tilde{c}^{t+1}\rangle$$
$$+ \frac{\rho}{2}\|\tilde{u}^{t+1} - u\|^2 + \frac{\rho}{2}\left\|\omega - \tilde{R}u\right\|^2 + \frac{\rho}{2}\left\|\tilde{D}u - q\right\|^2 + \frac{\rho}{2}\left\|(\Delta\tilde{\alpha})^{\top}u - \tilde{\alpha}\delta - \tilde{\mathrm{I}}\tilde{c}^{t+1}\right\|^2$$

$$W_i^{t+1} = \underset{W_i}{\operatorname{argmin}} \;\; \langle \lambda_{4,i}^t, W_i\mathbf{1} - \mathbf{1}\rangle + \langle \Lambda_{5,i}^t, S_i^{t+1} - W_i\rangle + \frac{\rho}{2}\|W_i\mathbf{1} - \mathbf{1}\|^2 + \frac{\rho}{2}\left\|S_i^{t+1} - W_i\right\|^2, \quad \forall i$$
$$= 1,2,\dots,n$$

$$H_i^{t+1} = \underset{H_i}{\operatorname{argmin}} \;\; \langle \lambda_{6,i}^t, H_i^{\top}\delta_p + \beta_i^{t+1} - b_i \cdot (B_i\eta)\rangle + \langle \Lambda_{8,i}^t, S_i^{t+1} - H_i\rangle$$
$$+ \frac{\rho}{2}\left\|H_i^{\top}\delta_p + \beta_i^t - b_i \cdot (B_i\eta)\right\|^2 + \frac{\rho}{2}\left\|S_i^{t+1} - H_i\right\|^2, \quad \forall i = 1,2,\dots,n$$

$$Z_i^{t+1} = \underset{Z_i}{\operatorname{argmin}} \, \mathbb{1}\big(\rho > \tilde{\lambda}\big)\mathbb{I}_{[0,1]^{(|\mathcal{R}|\cdot|\mathrm{T}|)\times|\mathcal{N}_i|}}(Z_i) + \mathbb{1}\big(\rho > \tilde{\lambda}\big)\mathbb{I}_{\{0,1\}^{(|\mathcal{R}|\cdot|\mathrm{T}|)\times|\mathcal{N}_i|}}(Z_i)$$
$$- \frac{\tilde{\lambda}}{2}\sum_{j=1}^{|\mathcal{N}_i|}\sum_{\ell=1}^{|\varepsilon|}\sum_{t=1}^{|T|}(Z_j)_{i,(r,t)}((Z_j)_{i,(r,t)} - 1) + \langle \Lambda_{10,i}^t, S_i^{t+1} - Z_i\rangle$$
$$+ \frac{\rho}{2}\sum_{i=1}^{n}\left\|S_i^{t+1} - Z_i\right\|^2, \quad \forall i = 1,2,\dots,n$$

$$\tilde{\beta}^{t+1} = \underset{\tilde{\beta}}{\operatorname{argmin}} \;\; \mathbb{I}_{\mathbb{R}}\big(\tilde{\beta}\big) + \langle \lambda_9^t, \tilde{c}^{t+1\top}\tilde{\mathbf{1}} + \tilde{\beta} - \Omega\rangle + \frac{\rho}{2}\left\|\tilde{c}^{t+1\top}\mathbf{1} + \tilde{\beta} - \Omega\right\|^2$$

$$\lambda_{1,i}^{t+1} = \lambda_{1,i}^t + \rho(S_i^{t+1}\mathbf{1} - u_i^{t+1}), \; \forall i = 1,2,\dots,n$$

$$\lambda_2^{t+1} = \lambda_{1,i}^t + \rho\left(\omega^{t+1} - R\left(\sum_{i=1}^{n} u_i^{t+1},\right)\right)$$
$$\lambda_{3,i}^{t+1} = \lambda_{3,i}^t + \rho(Du_i^{t+1}\mathbf{1} - q_i),\ \forall i = 1,2,\dots,n$$
$$\lambda_{4,i}^{t+1} = \lambda_{4,i}^t + \rho\left(W_i^{t+1\top}\mathbf{1} - \mathbf{1}\right),\ \forall i = 1,2,\dots,n$$
$$\Lambda_{5,i}^{t+1} = \Lambda_{5,i}^t + \rho(S_i^{t+1} - W_i^{t+1}),\ \forall i = 1,2,\dots,n$$
$$\lambda_{6,i}^{t+1} = \lambda_{6,i}^t + \rho\left(H_i^{t+1\top}\delta_p + \beta_i^{t+1} - b_i \odot B_i\eta\right),\ \forall i = 1,2,\dots,n$$
$$\lambda_7^{t+1} = \lambda_7^t + \rho(\alpha \odot (\Delta^\top u^{t+1} - \delta) - \tilde{\mathrm{I}}\tilde{c}^{t+1})$$
$$\Lambda_{8,i}^{t+1} = \Lambda_{8,i}^t + \rho(S_i^{t+1} - H_i^{t+1}),\ \forall i = 1,2,\dots,n$$
$$\lambda_9^{t+1} = \lambda_9^t + \rho\left(\tilde{c}^{t+1\top}\mathbf{1} + \tilde{\beta}^{t+1} - \Omega\right)$$
$$\Lambda_{10,i}^{t+1} = \Lambda_{10,i}^t + \rho(S_i^{t+1} - Z_i^{t+1}),\ \forall i = 1,2,\dots,n$$

The primal update rules can be simplified as

$$\omega_{\ell,\hat{t}}^{t+1} = \underset{\omega_{\ell,\hat{t}}}{\operatorname{argmin}}\ \ \omega_{\ell,\hat{t}}\theta_{\ell,t}(\omega_{\ell,\hat{t}}) + \lambda_{2(\ell,\hat{t})}^t\left(\omega_{\ell,\hat{t}} - r_{\ell,\hat{t}}\left(\sum_{i=1}^{n} u_i\right)\right) + \frac{\rho}{2}\left(\omega_{\ell,\hat{t}} - R_{\ell,\hat{t}}\left(\sum_{i=1}^{n} \mathrm{u}_i^t\right)\right)^2, \forall i = 1,2,\dots,|\varepsilon|, \forall \hat{t} = 1,2,\dots,\tilde{T}$$
$$S_i^{t+1} = \left(-\lambda_{1,i}^t\mathbf{1}^\top - \Lambda_{5,i}^t - \Lambda_{8,i}^t - \Lambda_{10,i}^t + \rho u_i\mathbf{1}^{t_\top} + \rho W_i^t + \rho H_i^t + \rho Z_i^t\right)(\rho\mathbf{1}\mathbf{1}^\top + 3\rho I)^{-1}, \forall i = 1,2,\dots,n$$
$$\beta_i^{t+1} = \Pi\left(\frac{1}{\rho}\left(-\lambda_{6,i}^t - \rho H_i^{t\top}\delta_p + \rho b_i \odot (B_i\eta)\right)\right)_{\mathbb{R}_+}, \forall i = 1,2,\dots,n$$
$$\tilde{c}^{t+1} = \Pi\left(\frac{1}{\rho}\left(\tilde{I}^\top\tilde{I} + \tilde{\mathbf{1}}\tilde{\mathbf{1}}^\top\right)^{-1}\left(\tilde{I}^\top\lambda_7^t - \lambda_9^t\tilde{\mathbf{1}} - \rho\tilde{I}^\top(\alpha\odot\gamma) + \rho\tilde{I}^\top\left(\alpha\odot(\Delta^\top u^t)\right) - \rho\tilde{\beta}\tilde{\mathbf{1}} + \rho\Omega\tilde{\mathbf{1}}\right)\right)_{\mathbb{R}_+}$$
$$u^{t+1} = \frac{1}{\rho}\left(I + \tilde{R}^\top\tilde{R} + \widetilde{D}^\top\widetilde{D} + (\Delta\tilde{\alpha})(\Delta\tilde{\alpha})^\top\right)^{-1}\left(\lambda_1^t + \tilde{R}^\top\lambda_2^t - \widetilde{D}^\top\lambda_3^t - (\Delta\tilde{\alpha})\lambda_7^t + \rho\tilde{u}^t - \rho\tilde{R}^\top\omega^{t+1} + \rho\widetilde{D}^\top q + \rho(\Delta\tilde{\alpha})(\alpha\odot\gamma) + \rho(\Delta\tilde{\alpha})(\tilde{I}\tilde{c}^{t+1})\right)$$
$$Z_i^{t+1} = \mathbb{I}(\rho > \tilde{\lambda})\Pi\left(\left(\frac{1}{\rho - \tilde{\lambda}}\right)\left(\rho S_i^{t+1} + \Lambda_{10}^t - \frac{\tilde{\lambda}}{2}\right)\right)_{[0,1]} + \mathbb{I}(\rho < \tilde{\lambda})\Pi\left(\left(\frac{1}{\rho - \tilde{\lambda}}\right)\left(\rho S_i^{t+1} + \Lambda_{10}^t - \frac{\tilde{\lambda}}{2}\right)\right)_{\{0,1\}}, \forall i = 1,2,\dots,n$$
$$W_i^{t+1} = \frac{1}{\rho}(\mathbf{1}\mathbf{1}^\top + I)^{-1}\left(-\mathbf{1}\lambda_{4,i}^{t_\top} + \Lambda_{5,i}^t + \rho\mathbf{1}\mathbf{1}^\top + \rho S_i^{t+1}\right), \forall i = 1,2,\dots,n$$
$$H_i^{t+1} = \frac{1}{\rho}\left(\delta_p\delta_p^\top + I\right)^\top\left(-\delta_p\lambda_{6,i}^{t_\top} + \Lambda_{8,i}^t - \rho\delta_p\beta_i^\top + \rho\delta_p(b_i \cdot B_i\eta)^\top + \rho S_i^{t+1}\right), \forall i = 1,2,\dots,n$$

User Equilibrium (UE) Algorithm

In our numerical experiments, we use the volume at User Equilibrium state of the system as our baseline. To compute the volume at User Equilibrium, we present Algorithm 2. Before we present the details of Algorithm 2, let us explain some notations used in this algorithm. Vector $v \in \mathbb{R}_+^{|\mathbb{E}|\cdot|T|}$ denotes the volume of links at different time slots. $\tilde{S} \in \{0,1\}^{(|\mathcal{R}|\cdot|T|)\times|\mathcal{N}|}$ is the matrix of route assignment of all the drivers in the system at different times such that each driver is assigned to the fastest route. Speed vector $\mu \in \mathbb{R}_+^{|\mathbb{E}|\cdot|T|}$ includes the speed of the links at different times. The travel time of paths at different times is presented in vector $\widetilde{\delta_p} \in \mathbb{R}_+^{|\mathcal{R}|\cdot|T|}$. In this algorithm, we rely on the method presented by [51] to compute matrix R based on the volume (v) and compute $\widetilde{\delta_p}$ based on speed (μ).

Algorithm 2 UE Algorithm

Input: Step size: α_{UE}, Number of iterations: $\tilde{T}$.
Compute $\mathbf{R}$ using volume vector $\mathbf{v}_0$ (historical data)
Compute route travel time vector $\tilde{\boldsymbol{\delta}}_{p,0}$ based on speed vector $\boldsymbol{\mu}_t$
for $t = 1, 2, \ldots, \tilde{T}$ **do**
    Create the assignment matrix $\tilde{\mathbf{S}}_t$ based on route travel time vector $\tilde{\boldsymbol{\delta}}_{p,t-1}$
    $\tilde{\mathbf{v}}_t = \mathbf{R}_{t-1}\tilde{\mathbf{S}}_t\mathbf{1}$
    $\mathbf{v}_t = (1-\alpha_{\mathrm{UE}})\mathbf{v}_{t-1} + \alpha_{\mathrm{UE}}\tilde{\mathbf{v}}_t$
    Compute $\mathbf{R}_t$ based on volume vector $\mathbf{v}_t$
    Compute speed vector $\boldsymbol{\mu}_t$ utilizing $\mathbf{v}_t$ and BPR function
    Compute $\tilde{\boldsymbol{\delta}}_{p,t}$ employing $\boldsymbol{\mu}_t$
end for
Return: $\boldsymbol{\delta}_{\mathrm{UE}} = \tilde{\boldsymbol{\delta}}_{p,\tilde{T}}$ and $\mathbf{R}_{\mathrm{UE}} = \mathbf{R}_{\tilde{T}}$

An Example of the Model and Notations

In this section, we provide a small example of a network for further description of our model and notations. Consider the network

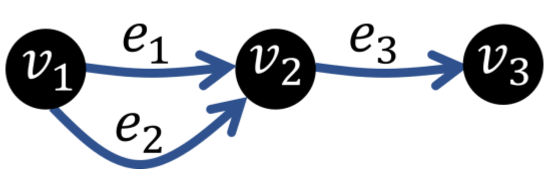


Figure 8. Network example $\mathcal{G}_1$.

where $\mathcal{V} = \{v_1, v_2, v_3\}$ is the set of nodes and $\mathcal{E} = \{e_1, e_2, e_3\}$ is the set edges (roads). Details of the links and attributes are represented in Table 2. The (origin, destination) pair is $(v_, v_3)$. There are two routes going from origin to destination as illustrated in Table 3. The time horizon set is $\boldsymbol{T} = \{1, 2, 3\}$ and each time is 0.2 hour. To estimate the location of drivers at each time, we need matrix $\boldsymbol{R} \in [0,1]^{9\times 6}$ as follows

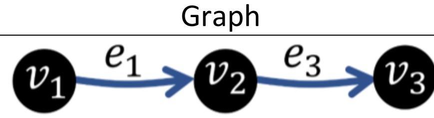
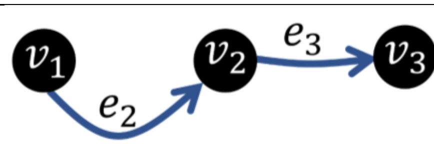
$$
R = \begin{array}{r} \\ \\ \\ t_2 = 1, e_1 \\ t_2 = 1, e_2 \\ t_2 = 1, e_3 \\ t_2 = 2, e_1 \\ t_2 = 2, e_2 \\ t_2 = 2, e_3 \\ t_2 = 3, e_1 \\ t_2 = 3, e_2 \\ t_2 = 3, e_3 \end{array}
\begin{array}{cccccc}
t_1 & t_1 & t_1 & t_1 & t_1 & t_1 \\
=1 & =1 & =2 & =2 & =3 & =3 \\
r_1 & r_2 & r_1 & r_2 & r_1 & r_2 \\
1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0.5 & 0 & 0.5 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0.5 & 0 & 0.5 & 0
\end{array}
$$

where t_1 is the entrance time of the driver and t_2 is the driver's arrival time at the road.

Table 2. Set of edges.

	Length (Mile)	Speed (mph)	Travel Time (Hour)
e_1	5	50	0.1
e_2	10	50	0.2
e_3	5	50	0.1

Table 3. Set of routes.

	r	Graph
Route 1 $e_1 \rightarrow e_3$	$r_1 = \begin{bmatrix}1\\0\\1\end{bmatrix}$	
Route 2 $e_2 \rightarrow e_3$	$r_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}$	

Assume there are two organizations in the system. The Value of Time for organization 1 is \$2.63 per minute and Value of Time for organization 2 is \$1.315 per minute so $\alpha_1 = 2.63$ and $\alpha_2 = 1.315$. There are three drivers in the system and $\mathcal{N} = \{d_1, d_2\, d_3\}$ such that $\mathcal{N}_1 = \{d_1, d_2\}$ and $\mathcal{N}_2 = \{d_3\}$.

Given that drivers d_1 and d_2 have to finish their travel as soon as possible and the travel time of driver d_3 can delayed up to two times of the shortes travel time, matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ and vectors $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ will be defined as follows:

OD Assignment

$$B_1 = \begin{matrix} d_1 \\ d_2 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Travel Time Multiplier

$$b_1 = \begin{matrix} d_1 \\ d_2 \end{matrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

OD Assignment

$$B_2 = d_3 \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}$$

Travel Time Multiplier

$$b_2 = d_3 \begin{pmatrix} 2 \end{pmatrix}$$