Synergistic data-driven travel demand management based on phone records

Yan Leng  
Massachusetts Institute of Technology  
Media Lab  
Cambridge, MA 02139, USA  
yleng@mit.edu

Larry Rudolph  
Two Sigma LP & Massachusetts Institute of Technology  
New York, NY, USA  
larry.rudolph@twosigma.com

Jinhua Zhao  
Massachusetts Institute of Technology  
Department of Urban Studies and Planning  
Cambridge, MA 02139, USA  
jinhua@mit.edu

Haris N. Koutsopolous  
Massachusetts Institute of Technology  
Civil and Environmental Engineering  
Boston, MA 02115, USA  
haris@mit.edu

Alex ‘Sandy’ Pentland  
Massachusetts Institute of Technology  
Media Lab  
Cambridge, MA 02139, USA  
pentland@mit.edu

ABSTRACT

Traffic congestion has increasingly threatened urban development economically, societally, and environmentally. Leisure trips contribute to 79% of the total travel demand. However, leisure trips suffer from the ‘richer-get-richer’ effect, leading to congestion exacerbation. We address this issue with a novel data-driven travel demand management framework by recommending locations based on the phone records. In particular, we infer unobserved location preferences using Matrix Factorization from longitudinal mobility histories. We then formulate an constrained optimization problem to maximize preferences regarding recommended locations while accounting for constraints imposed by road capacity. Our case study shows that under full compliance rate, congestion falls by 52% at a cost of 31% less location satisfaction. Under 60% compliance rate, 41% travel delay is saved with a 17% reduction in satisfaction. This study highlights the effectiveness of the synergy among collective behaviors in improving system efficiency.

KEYWORDS

Travel demand management, Congestion alleviation, Call Detail Records, Constrained Optimization

1 INTRODUCTION

Traffic congestion presents increasingly severe world-wide threats to urban areas, causing huge time lost, monetary costs, environmental deterioration, energy wastes and fatal accidents. Take the United States as an example, the extra economic cost of congestion due to time and fuel was estimated to be $160 billion in 2014. Furthermore, negative congestion impacts spread from peak commuting hours to midday and overnight, generating approximately additional 41% of the total travel delay during these periods [14]. Meanwhile, U.S. Travel Association released the travel statistic of 2016, concluding that 79% of domestic trips taken are for leisure purposes, with a total number of 1.7 billion person-trips (U.S. residents) and $106.4 billion tax revenues 1 [8]. Naturally, an inquiry comes to mind is how to manage the leisure trips and flexible travel demand to partly address the congestion problem. Leisure travelers, by definition, are more flexible in terms of time and destinations than commuting trips. As an illustrative example shown in Figure 1, individuals’ preferences regarding (location · time) bundles are relatively indifferent from one another when above a certain threshold, indicating the elasticities of leisure trips [15]. This indicates that the paradigm to manage the travel demand of leisure purposes has the potential to relieve the congestion to a large degree [24]. Nonetheless, little research has geared towards managing flexible travel demand in particular.

Another concern, however, rises when people make decisions on destinations or activities are the ‘richer-get-richer’ or ‘preferential attachment’ effect. Namely, popular locations and activities will be recommended either via mouth-to-mouth spreading or various recommendation engines, further causing severe coordination failure. This calls for meticulous and optimized coordinations of travel destinations and travel time [13, 15, 27]. Centralized coordination platform for distributed trips has grown to become even more necessary with the increasing of leisure travel demand. Specifically, it increments internationally and domestically by around 4.2% to 6.6% since 2010 respectively [1]. We argue that balancing and coordinating traffic flow should be centrally optimized as a demand

---

1Top leisure trips of U.S. domestic travelers include visiting relatives, shopping, visiting friends, fine dining and rural sightseeing.
vices, such as the Call Detail Records (CDR) used in this study, offers various choice bundles and y-axis shows an individual’s preferences regarding each choice bundle (the higher the preference, the better the choice bundle for this individual). Individual preferences upon leisure activities are unconstrained within certain space of spatial and temporal choices. Trading primary choice with the secondary one does not sacrifice noticeable utility, which in turn, gain notable collective travel time savings.

management strategy by policy makers with an analysis of the interplay between user preferences and systematic travel delay. Furthermore, individual travel time can also be reduced to avoid the unpredicted travel delay caused by stochastic traffic flow [2].

The availability of large-scale geo-localized data from mobile devices, such as the Call Detail Records (CDR) used in this study, offers an unprecedented opportunity for location-based service providers and governments to understand human mobility patterns in order to provide better services and improve system performances [11, 23, 39]. CDR is an opportunistic dataset that is collected by mobile phone carriers for billing purposes - making it a widely-prevalent, readily available, longitudinal and even real-time data source. The comprehensive picture of population-wide travel behaviors, more inclusive than almost all other data sources, enables policy-makers and transportation agencies to intervene from a holistic perspective. Given the promising applications of CDR in transportation, it has rarely been applied to manage travel demand as a decision-making tool apart from understanding existing travel patterns [4, 15, 28]. The central power of policy-makers with a collection of information both involving travelers from the demand side and the detailed traffic infrastructure from the supply side, can be taken advantage of in order to build a travel demand management system not only satisfying personal preferences but also making efficient use of the system capacity.

To address the above mentioned issue, this paper proposes a data-driven travel demand management framework for policy-makers to mitigate the adverse effects of flexible travel demand while respecting the needs of all stakeholders using a personalized and distributed travel recommendation engine. We show that the spontaneous travel decisions lead to severe traffic congestion; while exiting location recommendation engines aggravate the situation.

In this work, we develop a novel recommendation engine for authorities to optimize system performance by trading-off between satisfied preferences and road congestions. We use Bayesian Non-negative Matrix Factorization (BNMF) to infer travelers’ implicit preferences, exploiting the underlying similarities of locations and travelers. We then formulate an inequality constrained optimization problem to maximize satisfied location preferences at the user level under pre-defined road congestion constraints. Lastly, we identify the target population for distributing recommendations by integrating the system with a next-location prediction approach using Recurrent Neural Network. With an implementation on the CDR data in Andorra under various compliance rates, we show the effectiveness of the method. Specifically, under a 100% compliance rate, a 52% reduction in travel delay (from 11.73 minutes per hour to 5.61 minutes per hour) only sacrifices 31% satisfaction regarding the recommendations. Similarly, under a 60% compliance rate, a 41% reduction in travel delay (to 6.98 minutes per hour) sacrifices 17% in satisfaction regarding the recommendations. The obvious advantage of the method lies in the slight sacrifice of individual utility in return of the benefits for a larger population and the society as a whole.

The paper is organized as follows. Section 2 summarizes current literature regarding travel demand management and location recommendations. The data sources this study rely on are described in section 3. Section 4 demonstrates the framework and the detailed steps, including traffic flow estimation, preference inference, system efficiency maximization and identification of the target population. Section 5 evaluates the method with an application to the country of Andorra under various compliance rates. Section 6 concludes the paper and discusses future directions.

2 RELATED WORKS

2.1 Travel demand management

Travel Demand Management (TDM) encompasses strategies that alter demand patterns to increase transportation system efficiency, in contrast with adding more capacities to the system [24]. With low cost and easy-to-implement characteristics, TDM has broad applications in energy savings, air quality improvements, peak period congestion alleviation, etc [19]. Different categories of TDM strategies include, economic policies, physical change measures, legal policies, and information or education measures [21, 24]. Our study implicitly connects with economic policies. The traditional strategies of economic policies include taxing vehicles, congestion pricing, lowering transit costs, etc [24, 37].

However, Most of the existing TDM strategies ignore flexible travel demand, which differentiates our work from prior research. More specifically, our study differentiates itself from existing TDM strategies in several aspects:

- We focus on travel demand with leisure purposes, the flexible nature of which can be manipulated at the destination and time-of-day levels.
- We propose to use a new large-scale and readily-available behavioral data, Call Detail Records, to support our TDM
strategy for policy makers. We argue that this can be developed into a new direction supplementing existing data sources mostly focus on travel surveys and census.

2.2 Location recommendations

The innate characteristic of learning users’ preference from sparse visitation patterns links our study with the location recommendation literature. In the early 2010s, several studies introduced traditional recommender engines to personalized location recommendation. Since then, a prosperous collection of papers have sprouted in this area. Ye (2011) [42] introduced user-based and item-based Collaborative Filtering (CF) to location recommendations using check-in data, based on the assumption that similar users have similar tastes and users are interested in similar Points of Interests (POIs). Berjani (2011) [10] employed the more effective and efficient matrix factorization in POI recommendations on check-in history. Quercia (2010) [35] is the only work that makes recommendations using mobile phone data, which, however, only applies an inefficient and hard-to-scale item-based CF [40, 42].

Yet, the application of existing methods, ignoring service capacity constraints, will result in traffic congestions and exacerbate travel experiences, no matter how sophisticated these methods are in inferring preferences. Therefore, our paper has the following improvements:

- We argue that capacity constraints, currently ignored by existing literature, are in-negligible factors of location recommendation. This is the central novelty of our study.
- We develop a framework to recommend locations for system efficiency based on Call Detail Record. It can be applied in other cities when call records (or other large-scale data sources, such as WiFi, GPS, Smart-card transactions, etc), traffic counts and road network GIS files are available.

2.3 Next-location prediction

Accurate next-location predictions, given previous footprints from these data sources, is a significant building block benefiting many areas, including mobile advertising, public transit planning, and urban infrastructure management [5, 22, 30, 34]. It is therefore an increasingly popular topic in pervasive computing based on GPS, bluetooth, check-in histories, etc. Obviously, different data sources vary in terms of resolutions at spatiotemporal scales, representativeness of the population and the availability of contextual information [16]. Most researchers build Markov models and predict longitudes and latitudes as continuous variables based on previous travel trajectories [6, 7, 20]. For example, Mathew [31] predicts next-location using a Hidden Markov Model with contextual information captured by hidden factors, such as activities and purposes. Domenico [16] and Alhasoun [5] utilize mobility correlations, either measured by social interactions or mutual information, to improve accuracies. Even though extensive researches have acceptable accuracy in predicting next locations on GPS or social media check-in histories, the performance is poor with Call Detail Record, which is sparse in space and time [5].

In this paper, we propose a new perspective by mapping mobility behaviors based on Call Detail Record as sentences. This enables us to introduce the application of Recurrent Neural Network, a successful tool in language modeling, into mobility prediction. The mapping and the implementations are described in section 4.5.

3 DATA

In this study, we rely upon three main data sources to understand travel demand patterns and transportation system performances, capturing the transportation characteristics of both the demand-side and the supply-side. Call Detail Record is the main behavioral datasets we utilize to profile travel demand. In addition to CDR, we make use of road network topology and capacity, as well as the regularly-collected traffic counts by the government.

3.1 Call Detail Record

The Call Detail Record (CDR) was originally collected for billing purposes. A record is stored when a mobile phone user connects to any antenna operated by the mobile carrier. Each CDR entry contains an encrypted user ID, start and end times of the phone call, the coordinates of the connected cell tower, the country the SIM card is registered in, the brand and the mode type of the phone. The metadata is shown in Figure 2.

3.2 Road network

To better understand transportation infrastructures which influences the relationship between travel demand and travel delay, we collect the road network GIS file from Andorra Transportation Department. In addition, we use some characteristics of road links, including the mapping from ID to road links, free-flow travel time, the number of lanes per direction, and the road capacity. These are all necessary in travel time estimation. Free-flow travel time was obtained from the Google Map API. Road capacities for road links with different attributes were obtained from a NCHRP report [32].
The goal of the methodology is to manage travel demand via sending recommendations in order not to spam users.

Locations - only when the two are inconsistent, recommendations are distributed, we predict the next locations based on the decision-makers. Finally, in order to decide to whom the recommendations are distributed, we predict the next locations based on users’ historical traces and compare them with the recommended locations - only when the two are inconsistent, recommendations are distributed in order not to spam users.

3.3 Traffic statistics

Ground-truth traffic statistics are a necessity to scale the CDR-based traffic flow to population-wise OD traffic flow. This is especially critical for CDR since the relatively low spatiotemporal resolution inherently determines that the CDR-based traffic flow is an imperfect proxy for understanding travel demand. We therefore use the publicly-available traffic counts at key intersections, which is collected by local transportation authorities with cameras to monitor internal mobility [18].

4 METHOD

The goal of the methodology is to manage travel demand via sending recommendations in terms of where and when to visit. In a nutshell, this is done in several steps as outlined in Figure 4. We first estimate travel demand and traffic flow along road links from mobile phone records. We then infer user location preferences based on location traces. With Bayesian Non-negative Matrix Factorization, we infer these implicit location preferences regarding all the locations from the sparse matrix, which correlates with the characteristics of both travelers and locations. Built upon the inferred preferences, an optimization problem is formulated with the objective of maximizing satisfactions regarding the recommendations subject to tolerable congestion levels, which can be controlled by the decision-makers. Finally, in order to decide to whom the recommendations are distributed, we predict the next locations based on users’ historical traces and compare them with the recommended locations - only when the two are inconsistent, recommendations are distributed in order not to spam users.

4.1 Definitions

The following terms are mentioned throughout the method. We formally define them in this section.

**User profile.** A user profile contains the longitude, latitude, timestamps, and the characteristics of the user. A user profile is generated for each user based on individual mobility traces. Recommendation generation is based on this profile and its characteristics.

**Realized trips.** Realized trips are calculated by the number of times individual traveled to location. Realized trips are used as a proxy for location preferences.

**Idealized trips/ Satisfaction regarding the recommendations.** Idealized trips are calculated by the inferred preference of traveler regarding the location with no observed precense. Similarly, it is a quantification to proxy travelers’ preferences regarding the locations with no observations.

**Tolerable excess throughput.** Tolerable excess throughput is determined according to the tolerable congestions pre-determined by the policy-makers.

4.2 Traffic flow estimation

Estimating traffic flow and identifying inflexible commuting flows are the first step in our study. This gives us an understanding and evaluation of the current traffic system conditions. However, traffic flow estimation from Call Detail Records is not a trivial problem, which require rigorous preprocessing and calibration. This is largely due to the relatively low spatial and temporal resolution of CDR, making the pre-processed traffic flows from CDR an imperfect measurements. Since the estimation of traffic flow is not the focus of this study, we exclude much details. Interested readers can refer to detailed implementation and analysis in [4, 26]. We summarize the processing into four steps: OD-matrix extraction, identifying commuting trips, scaling up to actual traffic flow, and traffic assignment.

We first extract the tower-to-tower Origin-Destination (OD) matrix, aggregated by the individual movements from one cell tower to another cell tower. In particular, we identify peak-hour commuting flows (local and non-local population can be identified by a specific field directly from the raw CDR). The time-variant traffic flow is critical in understanding traffic conditions as traffic flows vary heterogeneously across links and hours of the day. The OD pairs are assigned to road links that the traffic flows pass through according to the road network from the shape-file (as in equation 1). The last step is to scale the aggregated traffic flows from CDR to the actual vehicle-trips using traffic counts as the ground-truth collected by the government. This is calculated as in Equation 2. The working assumption is that the number of travelers using their cell phones on the highway is a constant fraction of the number of vehicles on the highway. Under this assumption, traffic flows derived from the Call Detail Records can be used to approximate actual traffic flow.

\[
R_{pt} = \sum_{kjt} O_{kt} D_{jt} d_{t}^{j}(p),
\]

\[
TC_{it} = R_{it} \times k_{it},
\]
where \( p \) is the index of road link, \( t \) is the index for hour of day, \( O_k D_j \) represents OD pair with origin \( j \) and destination \( k \), \( R_{pt} \) is the vehicle trips along road link \( p \) during time period \( t \), \( TC \) is the actual traffic counts and \( \kappa \) is the scaling factor. \( \delta^{kj}(p) \) is an indicator function, which equals 1 when traffic flow originating from \( k \) and targeting at \( j \) passes through link \( p \). Due to the heterogeneity of traffic flow across time and space, we calculate scaling factors separately for each road link during each time period.

### 4.3 Preference inference

As no explicit review or rating regarding locations is available in CDR, we propose to use visiting frequencies/realized trips as a proxy for location preferences. The problem this section tackles is to infer user’s preferences regarding all other locations.

We use Bayesian Non-negative Matrix Factorization, a latent factor model, to infer travelers’ preferences regarding locations with no observations. The decomposition of visiting frequency matrix characterizes both the locations and users by vector of factors inferred from location visiting patterns, mapping both travelers and locations to a joint latent factor space of dimensionality \( k \). The latent factor space determines why and how travelers prefer or dislike each location based on hidden characteristics. The hidden factors can be interpreted as personal interests or land use categories. High correspondence between location and user factors, based on the characterization of both the locations and the travelers, leads to a recommendation [29].

We use Bayesian Non-negative Matrix Factorization developed by Schmidt and Cemgil [12, 38] to decompose the visiting frequency matrix from the data into matrix \( U \) and matrix \( T \), characterizing the individual characteristics and location characteristics respectively as shown in Figure 5.

\[
\begin{align*}
\text{non-negative factors} \\
\sum_{\text{Data}} P &= U L^T + \text{noise}, \\
\text{Normal likelihood:} \\
p(U, L, \sigma^2) &= \Pi_{i,j} N(P_{i,j})(UL^T)_{ij}, \sigma^2, \\
\text{Marginal likelihood:} \\
p(P) &= \frac{p(U|P)}{p(P)}
\end{align*}
\]

We assume normal distribution of likelihood over matrix \( P \). Meanwhile, we assume exponential priors over \( U \) and \( L \), and inverse Gamma prior over noise variance \( \sigma^2 \). And \( \theta = (U, L\sigma^2) \)

### 4.4 Optimization for system efficiency

The central step of the proposed method is to optimize travelers’ location preferences with the constraint of acceptable congestions, granting authorities the freedom to trade-off between the two factors. An optimization problem is formulated to maximize preferences regarding location recommendations subject to road capacity constraints.

This paper formalizes the traffic problem by modeling destination and time choice as follows: each traveler \( i \) makes a choice of location \( j \) and the travel time period for \( t \). The choice is made based on personal utility - simplified as \( p_{ijt} \) - which is assumed to be the preferences regarding the location \( j \) inferred from the call records. Since travelers make myopic and selfish choices, the system settles into a suboptimal state. In a suboptimal state, the travel time delay of the whole system as well as the congestion are much higher due to the mis-coordination. The set of destination and time choice bundle that occurs when every traveler maximizes their utilities, satisfied preferences in our problem, is referred to as the user equilibrium flows [2, 15], which is similar to Wardrop’s principles in route choice [41].

The objective function of the formulated inequality constrained optimization model is to maximize the overall satisfied preferences regarding the locations to be recommended. The constraint function is determined by the acceptable congestion of the decision-makers. Mathematically, the problem can be written as:

\[
\begin{aligned}
\text{Maximize satisfied preferences:} \\
\text{subject to} \\
\text{Individual constraint (for simplification):} \\
\text{Capacity constraint:}
\end{aligned}
\]

where \( x_{ijt} \) is the binary decision variable indicating whether to recommend location \( j \) to tourist \( i \) during discretized time period \( t \in T \), \( \beta_{ij} \) is the inferred idealized preference of location \( j \) for traveler \( i \). \( D_{\text{fixed}} \) are \( C_{\text{road}} \) are vectors representing the fixed travel demand on and capacity of road links. \( R(\cdot) \) is the function mapping travel demand to road links as stated in equation 1. We form the Lagrange function of in-equality constrained optimization problem as:

\[
L(x_{ijt}, \lambda) = f(x_{ijt}) - \lambda g(x_{ijt}),
\]
\[ g(x_{ijt}) = C_{road} + \text{Extra throughput} - \sum_{j=1}^{J} \sum_{i=1}^{I} R(x_{ijt} - D_{\text{fixed}}) \]

\[ \text{Utility function, } U_{ijt} = \max_{j \in J} x_{ijt}. \]

One compelling characteristic of our optimization framework is that it can be easily extended to incorporate more factors from both the user characteristics and from the systematic performance perspective. We only consider users’ preferences in the utility function, \( U_{ijt} = \max_{j \in J} x_{ijt} \). More variables, such as travel time and travel monetary costs, can be added to better surrogate for the utility, for example \( U_{ijt} = \beta_0 + \beta_1 \max_{j \in J} x_{ijt} + \beta_2 \text{travel time} + \beta_3 \text{travel cost} \). Additional constraints can be added to incorporate more societal benefits, such as lower energy consumption, smaller negative environmental impact, and diversified social interactions, etc.

### 4.5 Identifying target population - Next-location prediction with Neural Network

In order to distribute recommendations more efficiently, the last step of the framework aims to identify the target population whose predicted next-locations are inconsistent with the one to be recommended. The problem can simply be formulated as predicting the next-locations of individuals based on historical location traces. We migrate the sophisticated techniques from language model into mobility. Specifically, we apply Recurrent Neural Network (RNN) to learn the dependences in mobility traces. RNN is an adaptation of the traditional feed forward neural network, which can process variable-length sequences of inputs[33, 36]. We argue that RNN is promising in next-location predictions for the following reasons. 1) Travelers visit locations in sequence and RNN reads in data sequentially. 2) The heterogeneity in mobility behaviors, the size of location traces and frequency of mobile phone usage, makes traditional machine learning techniques inapplicable. The ability to handle variable input lengths makes RNN promising in this situation [17]. 3) The embedding layer of RNN, mapping each input unit to vector space, can be accommodated to infer the real-valued representations of cell towers. This is useful in relating ‘similar’ cell towers close to each other in the vector space.

Specifically, cell tower traces for each individual are modeled as a sentence and each cell tower as a word. The mapping between the two is described in Table 1. We use a simple RNN architecture (as shown in Figure 6), comprising of an input layer, a hidden layer, a long-short-term-memory layer and output layer. The cell tower with the maximum probability is predicted to be the next location. The outputs are compared with the recommended locations to determine the actions to be taken by the policy-maker.

### 5 APPLICATION

We evaluate our method in Andorra, an European country near Spain and France. The transport in Andorra is mostly dependent on private automobiles with little public transport (bus or train) available. The anonymized CDR data used in this paper is provided by the monopoly mobile carrier in Andorra with 100% market share. Andorra is heavily relying on tourism, and we target 47743 tourists (including 20311 French tourists and 27432 Spanish tourists) visiting Andorra during a week on May of 2015 to provide personalized recommendations to improve system efficiency. The preferences of these tourists are learned from the mobility histories from the past 12 months. The country where the SIM card is registered is stored in each CDR record, which indicates the nationality of the mobile phone user. This enables us to identify tourists and local commuters. The spatial distribution of the cell towers are displayed in Figure 3, with different colors representing different cities where the tower is located.

We performed extensive simulations to gain insights into the interplay between satisfaction regarding the recommendations and the travel delay caused by congestion. To better understand how compliance rates, the probability that travelers will follow the recommendation, influence the effectiveness of our method, we vary it in the range of 20% and 100% across the population in order to evaluate the improvement on traffic conditions due to the recommendation system under different scenarios.

**Table 1: Mapping from mobility model to language models.**

<table>
<thead>
<tr>
<th>Language Model</th>
<th>Call detail records</th>
<th>Mobility model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word</td>
<td>Cell tower</td>
<td>Associate cell towers with hidden factors</td>
</tr>
<tr>
<td>Phrase</td>
<td>Short sequences of cell towers</td>
<td>Infer activities/ trip purposes</td>
</tr>
<tr>
<td>Sentence</td>
<td>Cell tower traces</td>
<td>Location predictions</td>
</tr>
</tbody>
</table>

**Figure 6: Infrastructure of RNN in next-location prediction**
Table 2: Results comparisons. $\Delta t$ is the average travel delay during the peak hour. Idealized trips measure travelers’ satisfaction regarding the recommended locations. Comp. is short for compliance rate. User optimal is the baseline model where recommendations are only based on user preferences.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$\Delta t$ (min/h)</th>
<th>Idealized trips</th>
</tr>
</thead>
<tbody>
<tr>
<td>User optimal</td>
<td>11.73</td>
<td>64925</td>
</tr>
<tr>
<td>100% comp.</td>
<td>5.61</td>
<td>44930</td>
</tr>
<tr>
<td>80% comp.</td>
<td>6.17</td>
<td>49997</td>
</tr>
<tr>
<td>60% comp.</td>
<td>6.98</td>
<td>53680</td>
</tr>
<tr>
<td>40% comp.</td>
<td>8.37</td>
<td>57442</td>
</tr>
<tr>
<td>20% comp.</td>
<td>10.40</td>
<td>61219</td>
</tr>
</tbody>
</table>

relationships between traffic flow and travel time based on the characteristics of the road infrastructure. One of the most widely used methods is the Bureau of Public Road function (BPR) [3], which models travel time as a function of the ratio between actual traffic volume and road capacity, volume-over-capacity (VOC) [9], as shown in Equation (11).

\[
t_{\text{sim}} = \frac{t_{ff}}{V/C} + \alpha(V/C)^{\beta}
\]  

Average travel delay:

\[
\Delta t = t_{\text{sim}} - t_{ff}
\]  

where $t_{ff}$ and $t_{\text{sim}}$ are the free flow and simulated travel time on the road segment. $\Delta t$ is the delay caused by congestion, $\alpha$ and $\beta$ are parameters used to characterize the non-linear relationship between $V/C$ and $t_{\text{sim}}$. The default parameter for the BPR equation are $\alpha = 0.15$ and $\beta = 4$.

We compare our method with a baseline model, referred to as user optimal method. This method make recommendations by maximizing overall user utilities - preferences regarding the recommended locations. Evaluation results are shown in Table 2, which summarize average travel delay per hour and idealized trips of: 1) user optimal method; 2) the proposed method under various compliance rates. Idealized trips, one of the metrics used to evaluate the method, are aggregated as $\sum_{i,j} P(y_j = j | \hat{y}_i = i)$. Under full compliance rate, the average travel delay is 5.61 minutes per hour with 44930 idealized trips, which indicates that with a 52% reduction in congestion time, and only 31% of idealized trips are sacrificed. When the compliance rate reduces to 80%, a 47% reduction in congestion time is achieved with only 23% of idealized trips being sacrificed. In general, the lower the compliance rate, the larger the idealized trips and the longer the travel delay.

The upper plot of Figure 7 shows the relationship between idealized trips and tolerable excess throughput, which demonstrates how individuals perceive their benefits under various level of congestion controls. Higher compliance rates satisfy larger individual benefits, which generate more traffic. The concave relationship reveals that preferences are satisfied more quickly in the beginning and slow down afterwards. The tolerable excess throughput enables decision-makers to manage congestion to an acceptable level.

The lower plot of Figure 7 reveals the interplay between idealized trips and average travel delay. Drawing a horizontal reference line across the plot indicates that for the same level of idealized trips, synergized behaviors generate less congestion. On the other hand, a vertical reference line will show that for the same travel delay, coordinated behaviors are more efficient. This calls for an effective schema to incentivize behavioral change to achieve synergistic effects.

**Identifying target population.** We also evaluate the performance of RNN on the prediction task to identify target population. As stated in the method section, we apply the approach specifically on tourists, which can be filtered by the country code from CDR. We do not exclude travelers with too few observations as long as their travel call includes more than one mobility point, suggesting the applicability of the method in sparse observation scenarios.

We introduce two baseline models, namely Naive model and Markov model. Naive model predicts the next location as the most frequent location in the mobility trace of the individual. The Markov model is built upon the contextual co-occurences between sequences of locations [20, 25, 30]. We tune various parameters to improve model performance and avoid over-fitting, which include activation function, dimension of embedding layer, drop out rates, sample size, batch size, and activation functions. The method is evaluated using accuracy $\gamma$, where $\gamma = \sum_{j=1}^{N} P(y_j = j, \hat{y}_j = j)$. The improvement $\tilde{\gamma}$ is evaluated by the normalized improvement comparing with the Naive model as $\tilde{\gamma} = \gamma - \gamma_{\text{naive}}$. The performances of the
model are shown in Table 3, which shows that we out-perform 34% in the accuracy comparing with the Naive model.

### 6 DISCUSSIONS

Traffic congestions have caused tremendous economic costs and generated severe urban issues across the world, deteriorating the environments, increase energy consumptions, wasting huge amounts of time, blocking emergency vehicles and calling for resources for additional infrastructures. Further, travel delays have expanded from commuting hours to midday and midnight by about 41% compare to free flow traffic [14]. Additional statistics shows that about 79% travels are non-work related, which are flexible in space and time by nature [8]. However, most existing transportation research targets peak-hour commuting traffic flows. In contrast, flexible travel demand is the population that this paper targets. In this work, we propose a low-cost solution utilizing existing information infrastructure to develop a centralized travel demand management tool - a space-time recommendation engine - to distribute traffic flows and divert travelers to various locations in order to gain more travel-time savings and improve travel experiences both for individuals and the society as a whole.

Our results suggest that uncoordinated individual travel decisions, without accounting for system efficiency, lead to traffic congestion; Furthermore, existing location recommendation or mouth-to-mouth influence exacerbates the situation due to the ‘richer-get-richer’ effect. In particular, we propose a travel demand management framework, recommending locations and time based on widely-existing and the most pervasive Call Detail Records based on a centralized optimization system accounting for both individual preferences and systematic capacity. The framework distinguishes itself from existing methods by different targeted demand - leisure travelers - and the trading-off between system capacity and individual preferences as a recommendation engine and travel demand management tool. The combination of recommendation and system efficiency is also a novel aspect of our study which are ignored by existing location recommendation engines. Our framework and pipeline are composed of four step: estimating traffic flow, infer location preferences, centralized optimizing for system efficiency and identify target population.

We simulate traffic flows with our method as a case study in Andorra, revealing a noticeable impact in reducing traffic congestion with moderate sacrifices on individual preferences. For example, under 100% compliance rate, there is a 52% reduction in travel delay (from 11.73 minutes per hour to 5.61 minutes per hour) with 31% dissatisfaction rate regarding the recommendations. Even with a much smaller compliance rate, under 60%, there is a 41% reduction in travel delay (to 6.98 minutes per hour) with only a 17% dissatisfaction rate. The simultaneous trade-off between congestion relief and overall satisfied location preferences from the simulation results indicates that moderate sacrifices for individual utilities lead to significant collective travel-time savings.

This research opens up several promising directions to extend the synergistic travel demand management framework with pervasive technologies. Given different objectives of the policy-makers, more systematic measures can be consolidated into the framework by enriching the optimization problem. Other systematic measures, such as air pollution, energy consumptions, can be factored into the constraints for a more systematic and customized evaluations and improvements of the system. In addition to the system-side extensions, detailed travel planning, including transportation modes, routes, and ranges, can be supplemented to augment the user utility functions.

In addition, the proof-of-concept simulations in this study, as a first attempt, demonstrate the effectiveness of the synergistic travel demand management tool in distributing travel delay and congestions. The natural future direction is to investigate the design of mechanisms to incentivize users to slightly sacrifice perceived immediate benefits for more societal gains. This includes designing information configurations for travelers, such as strategically present information accounting for travelers’ willingness to accept or using dynamic pricing mechanism to optimally shift travel demand. Moreover, it is interesting to propose a comprehensive framework of for practical applications, detailing the distribution channel, frequencies, target markets, from game-theoretic, behavioral economics and marketing science perspective.

### ACKNOWLEDGMENTS

The authors would like to thank Changing Places Group at MIT Media Lab, directed by Professor Kent Larson, for initiating the project and the collaboration. We acknowledge Luis Alonso, Arnaud Grignard, Nuria Macia for the helpful discussions. We would also like to thank Andorra Telecom for the data and technical support.

### REFERENCES

Synergistic data-driven travel demand management based on phone records


