Transportation Research Part C 120 (2020) 102814

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Modeling user interaction with app-based reward system: A graphical model approach integrated with max-margin learning

Jingshuo Feng [a], Shuai Huang [a,*], Cynthia Chen [b]

[a] Department of Industrial and Systems Engineering, University of Washington, Seattle, WA 98195, USA
[b] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, USA

ARTICLE INFO

Keywords:
Travel behavior
Graphical model
Max-margin learning
App-user interaction data
Personalized behavior model

ABSTRACT

In recent years, there has been a rapid growth of smart apps that could interact with users and implement personalized rewards to coordinate and change user behavior. Understanding user behavior is an enabling factor for the success of these promising apps. However, existing statistical models for modeling user behavior encounter limitations. Choice models based on Random Utility Maximization (RUM) commonly assume that the data collection is independent with the human behavior. However, when users interact with the apps, the real potential and also the real challenge for modeling user behavior is that the apps not merely are data collection tools, but also change users' behaviors. In this work, we model the user behavior as a graphical model, examine our hypothesis that existing choice models are not suitable, and develop an interesting computational strategy using max-margin formulation to overcome the learning challenge of the our proposed graphical model that is named the Latent Decision Threshold (LDT) model.

1. Introduction

Much research attention has been focused on Transportation Demand Management (TDM) to mitigate traffic congestion and a number of related costly urban issues, such as pollution, noise and road user safety (Mayeres et al., 1996; Katzev, 2003). TDM strategies are designed and widely deployed in most metropolitan regions to modify travel behavior patterns by providing incentives or costs to certain travel behaviors (Arian et al., 2018; Zhu et al., 2019; Meyer, 1999). Examples include increasing parking costs in peak periods, offering low-cost public transportation, and monetary rewards, etc (Ben-Elia and Ettema, 2011b,a; Nijkamp and Shefer, 1998; Rouwendal and Verhoef, 2006; Rothengatter, 1982; Schuitema and Steg, 2008; Bamberg et al., 2003). These TDM strategies are usually developed based on population level, and researches show that those generic strategies have limited success (Giuliano, 1992; Stopher, 2004; Möser and Bamberg, 2008). Individuals respond differently to these TDM strategies (Habibian and Kermanshah, 2013), given their diverse demographic and socio-economic characteristics (Elias and Shiftan, 2012). Such diversity makes generic incentive ineffective, and thus, personalized incentive strategies are potentially more promising.

In recent years, the rapid proliferation of smart personal technologies makes it possible to interact with commuters and offer incentives individually (Arian et al., 2018; Zhu et al., 2019; Chen et al., 2016). For example, Zhu et al. (2019) proposed a new personalized TDM system which could offer personalized promotions to each commuter to change his/her behaviors. Ideally, the promotions with tailored rewards would be designed based on statistically accurate modeling of the user behavior, by learning from the interaction history between the commuter and the system. It has been reported in Zhu et al. (2019) that the personalized TDM

* Corresponding author.
E-mail address: shuaih@uw.edu (S. Huang).

https://doi.org/10.1016/j.trc.2020.102814
Received 20 February 2020; Received in revised form 18 September 2020; Accepted 19 September 2020
Available online 28 September 2020

system is quite effective in changing users' behaviors, i.e., the acceptance rate of the promoted suggestions reaches 68%, which leads to a significant travel time saving and congestion mitigation on the transportation system.

To fully unleash the potential of these app-based reward systems, understanding user behavior through the user–app interactions is an enabling factor in further intervention strategy development and optimization of user experience. The application of user models can improve the personalized recommendations (Chaptini, 2005; Jiang et al., 2014) and it is essential for assigning proper personalized incentives such as monetary rewards in order to encourage behavioral changes. Discrete choice models based on the theory of Random Utility Maximization (RUM) (Chorus et al., 2013; McFadden et al., 1973; Ben-Akiva et al., 1985; Hess et al., 2018) such as logit model are widely used in learning and understanding behavioral preferences for their good interpretability. However, they commonly assume that the data collection is independent with the human behavior, in other words, the technology for data collection is usually considered as irrelevant and will not interfere with the object to be observed and modeled.

The connection between the data collection and user behavior in the new system needs a more advanced statistical model to characterize. While not exactly concerning the same problem as ours, there have been some studies in the literature on the endogeneity between choices and the decisions, or choices and the travel behaviors. Guevara and Ben-Akiva (2010) deals with the endogeneity caused by model misspecification (i.e., omission of the attributes) using control function and latent variable. It does not relate the data collection procedure with the change of user behavior caused by a reward mechanism typical in nowadays smart TDM systems. Guevara and Hess (2019) corrects the endogeneity which comes from the stated preferences and revealed preferences. Ben-Akiva et al. (2012) describes a framework of incorporating contexts like social network as the decision being made may also be affected by family, friends, and other choices being given. Both works do not concern personalized reward systems. Hancock et al. (2018) uses a dynamic model from decision field theory to characterize the changing preferences with sequential choices and decisions, which concerns a different problem from ours.

Thus, none of the above approaches are fully applicable to the unique challenge posed by the app-based reward system where rewards are influenced by the alternative's attributes and personal preferences. To fill in this gap, in this work we propose a graphical model to depict the user–app interaction and to model the user behavior. Graphical model is a generic term referring to a family of multivariate statistical models that specifically model the interactions among variables and derive their data-generating process (Koller and Friedman, 2009). Using the graph, the model can represent the conditional dependencies among the variables, which enables graphical model to reveal the relationship between variables and makes it easy to interpret. Some existing hybrid choice models could also be potentially cast into the framework of graphical model, e.g., similar efforts have been undertaken in the literature for a range of linear models that show in the framework of graphical models many existing models find a unifying framework (Roweis and Ghahramani, 1999). A distinct concept that separates our proposed method with the existing works is the concept of "decision threshold", which is the tipping point where a user may change his/her decision between alternatives under a combined influence of the alternative's attributes and personal preferences. To model this new mechanism we develop the graphical model named the Latent Decision Threshold (LDT) model. It can characterize a user's decision behavior considering the attributes of the alternatives, the user's preferences, and the user's decision threshold between the alternatives. As the tipping point encodes a nonlinear relationship among variables, our proposed LDT model is a new type of graphical model that tailors the user–app interaction mechanism and provides a more delicate understanding of how a user makes a decision in a particular situation influenced by rewards.

Many graphical models and algorithms are very computational costly (Lee and Shi, 2001; Wei and Tanner, 1990), particularly when the interactions among variables are not all linear, as in our proposed LDT model. Therefore, another important contribution of this work is that we further reveal an interesting connection between the parameter estimation problem of the LDT model and max-margin learning. Using this connection, we resort to the computationally efficient solutions of max-margin formulation as a better approach than algorithms such as Expectation–Maximization algorithm (EM) which is usually used when a graphical model has latent variables.

The remainder of this paper is organized as follows. In Section 2, we introduce an example of app-based reward systems from a personalized TDM system, and develop a characterization of the user–app interaction process using existing models for user behavior such as logit model and mixed logit model (MLM). In Section 3, we develop the formulation of the proposed LDT model and construct an efficient parameter estimation algorithm based on the max-margin learning principle. Comprehensive simulation studies are conducted in Section 4, to compare the proposed LDT model with existing models. A real case study is presented in Section 5. Section 6 includes a brief conclusion and discussion.

2. Background

2.1. Personalized TDM system

An example of app-based reward systems can be seen in Zhu et al. (2019). They introduce a personalized TDM system which could offer tailored promotions with personalized incentives to each commuter to change his/her travel behaviors. When a commuter is about to depart, he/she can request a trip in the app. Then a promoted alternative travel plan will be generated based on the app's knowledge of this commuter, i.e., based on the decision choice model that can be learned from previous interactions between the user and the app system. The alternative may differ from the user's original travel plan in some attributes, such as departure time and total travel time. To encourage the user to accept the promotion, a certain amount of reward points is assigned and the commuter will be awarded if he/she accepts the promotion and changes the travel plan accordingly.

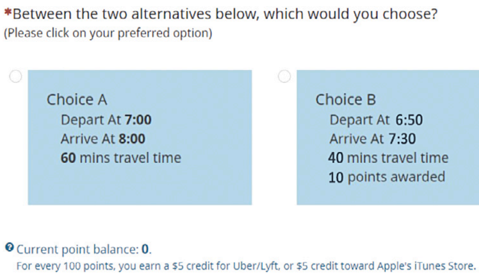

Fig. 1. An example of the alternatives and the choice question: the original plan and one promoted alternative plan with a certain amount of reward points.

Fig. 1 shows an example of user scenario of this personalized TDM system. Before the trip, the user will be asked to choose between Choice A (which is the original travel plan) and Choice B (a promotion). In this case, the attributes that characterize the choices include Schedule Delay Early (SDE), Schedule Delay Late (SDL), and Travel Time Saving (TTS). The use of SDE and SDL as two variables is needed because it has been found that people have different preferences on departing earlier or later if they are asked to change their travel plan (Ben-Elia and Ettema, 2011a,b). As Choice A is the original travel plan, which is also often called the default plan, the system will determine the amount of reward points needed for the potential acceptance of the promoted alternative Choice B (Zhu et al., 2019). In other words, a personalized incentive should be derived based on the alternatives' attributes and the user's preferences to encourage the user to switch to the new alternative.

2.2. The logit models

In this study we concentrate on a class of Random Utility Maximization (RUM) known as logit models. The RUM theory assumes that each alternative is characterized by M attributes, denote as $\boldsymbol{x} = [x_1, \ldots, x_M]^\top$. A user is assumed to have preferences on these attributes, and the user will make the final decision (denote as y) based on the attributes of the alternatives and his/her own preferences on these attributes (denoted as $\boldsymbol{\beta}$). Put simply, the problem in modeling decision behavior is to find $\boldsymbol{\beta}$.

The RUM theory assumes that a user will select the alternative with the highest utility, and utility is a concept quantifying the attractiveness of an alternative (Ben-Akiva et al., 1985; Hess et al., 2018; Hensher, 1994) which is defined to be related to the attributes, for example:

$$U_i = V_i + \epsilon_i = \beta_0 + \sum_{m=1}^{M} \beta_m x_{im} + \epsilon_i = \beta_0 + \boldsymbol{\beta}^\top \boldsymbol{x}_i + \epsilon_i,$$

where $V_i = \beta_0 + \boldsymbol{\beta}^\top \boldsymbol{x}_i$ is the measurable systematic utility and ϵ_i is the random utility (Ben-Akiva et al., 1985). The parameter $\boldsymbol{\beta}$ shows the preferences of the user, i.e., a positive β_m indicates that this attribute x_m of the alternative is attractive to the user, granting higher utility to the alternative and making it more likely to be chosen. With the multinomial logit (MNL) formulation (McFadden et al., 1973), the logit model based on RUM can be expressed as follows, where the probability that the user chooses alternative i is given by:

$$Pr(y = i) = \frac{\exp(V_i)}{\sum_{j=1}^{J} \exp(V_j)} = \frac{\exp(\beta_0 + \boldsymbol{\beta}^\top \boldsymbol{x}_i)}{\sum_{j=1}^{J} \exp(\beta_0 + \boldsymbol{\beta}^\top \boldsymbol{x}_j)}. \tag{1}$$

For applications where the user chooses among two alternatives, $\boldsymbol{x}$ could be the differences between the attributes of the two alternatives (Hess et al., 2018; McFadden et al., 1975). To apply logit model on the new personalized TDM system as shown in Fig. 1, we use $y = -1$ to denote that the user chooses Choice A, and use $y = 1$ to denote that the user chooses Choice B. The four attributes are $x_{SDE} = 10$, $x_{SDL} = 0$, $x_{TTS} = 20$, and the monetary reward points $r = 10$. According to the logit model, the probability of accepting the promotion ($y = 1$) will be:

$$Pr(y = 1) = \frac{\exp(V_B)}{1 + \exp(V_B)}, \text{ where } V_B = \beta_0 + \beta_{SDE} x_{SDE} + \beta_{SDL} x_{SDL} + \beta_{TTS} x_{TTS} + \beta_r r. \tag{2}$$

As the RUM theory is used for population-level modeling, a more related model in our context is the mixed logit model (MLM) (also called the random parameters logit model) that allows the parameters of the individuals vary. It can model the heterogeneity of the population (Greene and Hensher, 2003; Hensher and Greene, 2003; Bastin et al., 2010) by modeling the individual-level parameters $\boldsymbol{\beta}^{(n)}$'s as random samples drew from a multivariate normal distribution. For example, the probability that individual n will accept the alternative can be expressed in the form (Train, 2009; Hensher and Greene, 2003; Milton et al., 2008):

$$Pr(y = 1) = \int_{\boldsymbol{\beta}^{(n)}} \frac{\exp(V_B)}{1 + \exp(V_B)} f(\boldsymbol{\beta}^{(n)}) d\boldsymbol{\beta}^{(n)}, \tag{3}$$

where $f(\boldsymbol{\beta}^{(n)})$ is the density function for all the individual-level parameters $\boldsymbol{\beta}^{(n)}$, and it can be specified as a multivariate normal distribution with unknown mean and covariance. We can treat all the preference parameters as random parameters and the systematic utilities V_B in Eq. (3) are based on such personal preferences, i.e., $V_B = \beta_0^{(n)} + \beta_{SDE}^{(n)} x_{SDE} + \beta_{SDL}^{(n)} x_{SDL} + \beta_{TTS}^{(n)} x_{TTS} + \beta_r^{(n)} r$. MLM will be a baseline in our study to be compared with our proposed LDT model, and results are shown in Sections 4 and 5.

Table 1
Estimated coefficients using population-level Logit model on the personalized TDM system data.

	$\hat{\beta}_0$	$\hat{\beta}_{SDE}$	$\hat{\beta}_{SDL}$	$\hat{\beta}_{TTS}$	$\hat{\beta}_r$
Estimated coefficients	1.9161	−0.0210	−0.0290	0.0504	−0.0202
p-value	0.000***	0.000***	0.000***	0.327	0.000***

Table 2
Results of the Mixed Logit Model (MLM) on the personalized TDM system data.

	$\hat{\beta}_0^{(n)}$	$\hat{\beta}_{SDE}^{(n)}$	$\hat{\beta}_{SDL}^{(n)}$	$\hat{\beta}_{TTS}^{(n)}$	$\hat{\beta}_r^{(n)}$
% of Counter-intuitive sign	–	15.81%	1.76%	21.25%	38.84%

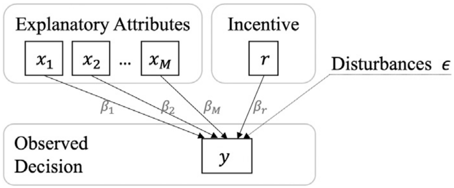

Fig. 2. An illustration of the logit model based on RUM.

2.3. Limitations of the logit models

We observe some counter-intuitive results when applying the logit model and MLM to analyze the data collected in Zhu et al. (2019). Table 1 shows the estimated coefficients of the logit model for the population. The fact that $\hat{\beta}_r$ in this logit model is negative literally implies the higher the rewards, the lower the probability the user could accept the promotion. If this were true, it will lead to a consequence that we cannot change user's travel behavior by providing more rewards. On top of this conceptual difficulty to understand the model, we also observe an unusual statistical phenomenon: the estimated coefficients are close to zero. Later in-depth analysis in Section 5 will reveal that this may be due to the large heterogeneity of the users. Hence we also apply mixed logit model (MLM) as shown in Eq. (3) to learn personalized models. The results of MLM are shown in Table 2.

MLM did not alleviate the difficulty in interpretation; rather, more counter-intuitive signs of $\hat{\boldsymbol{\beta}}^{(n)}$ are observed. For instance, a considerable portion of $\hat{\beta}_r^{(n)}$ is negative. This paradox leads us to hypothesize that we could not take it at its face value, i.e., $\hat{\beta}_r^{(n)}$ not only estimates $\beta_r^{(n)}$ but also something else. This indicates that the data-generating mechanism is different from the one assumed by the logit model and MLM. The same observation could be made on $\hat{\beta}_{TTS}^{(n)}$, as time saving should also be an encouraging factor for users and is supposed to have positive sign (Ben-Elia and Ettema, 2011a,b).

2.4. Analysis of the paradox

The counter-intuitive results shown in Tables 1 and 2 imply a mismatch of the modeling framework as illustrated in Eq. (1) with the underlying mechanism of the users' decision-making process when using app-based reward system. To illustrate this, Fig. 2 shows a conceptual understanding of the data-generating mechanism assumed by the logit models. It takes all the attributes including the reward as a set of variables with additive effects to predict the final decision. This "flattened" treatment of all attributes in the decision-making behavior as a set of variables with interchangeable positions in the same layer, while each variable's effect is additive, is probably an oversimplification of the problem that caused the counter-intuitive results.

To provide a remedy for this problem, first, it is worthy of pointing out that in existing app-based TDM systems it is a common strategy that the rewards assigned to the promotions are usually related to the attributes, e.g., in Zhu et al. (2019), when an alternative travel plan is generated, the reward points for this alternative are derived based on the attributes and the estimated user preferences. In other words, r is often correlated with other attributes $\boldsymbol{x}$. This correlation or multicollinearity has been known to be a hazard to models that have linear forms where the logit models are no exception. Lemma 1 articulates the danger of applying those models with one-layer flattened structure to model the user behavior in interactive apps and offers an explanation of the unstable and counter-intuitive estimations.

Lemma 1. *In logit models, when the reward r is linearly related to the attributes $\boldsymbol{x}$, there will be an infinite set of estimated coefficients $\hat{\boldsymbol{\beta}}$ that can lead to same utilities, and therefore, all maximize the log-likelihood of the model. In some of these solutions, $\hat{\beta}_r$ may be negative.*

The proof is provided in Appendix. Lemma 1 indicates that the parameter estimations of the logit model will be unstable and misleading when the attributes are correlated with rewards r. It is then of interest to see if this is true in the personalized TDM system data we have used in this study. Table 3 shows the population-level correlations between rewards r and the other three attributes, i.e., x_{SDE}, x_{SDL} and x_{TTS}, by pooling all users' data together. The result shows that r is significantly correlated with other three attributes. We also check the individual-level correlations (i.e., using each user's own data), and the average correlations

Table 3
Correlations between rewards r and other attributes $\boldsymbol{x}$ in personalized TDM system data.

	x_{SDE} and r	x_{SDL} and r	x_{TTS} and r
Correlation	0.0296	−0.0733	−0.1890
p-value	0.006***	0.000***	0.000***
Average of the personal correlations	0.1036	−0.0141	−0.0526
Average of the absolute values of personal correlations	0.3299	0.2993	0.2058

Table 4
Estimated coefficients using linear regression model for r on $\boldsymbol{x}$ on the personalized TDM system data.

	$\hat{\gamma}_0$	$\hat{\gamma}_{SDE}$	$\hat{\gamma}_{SDL}$	$\hat{\gamma}_{TTS}$
Estimated coefficients	39.558	0.0106	−0.0793	−0.7485
p-value	0.000***	0.481	0.000***	0.000***

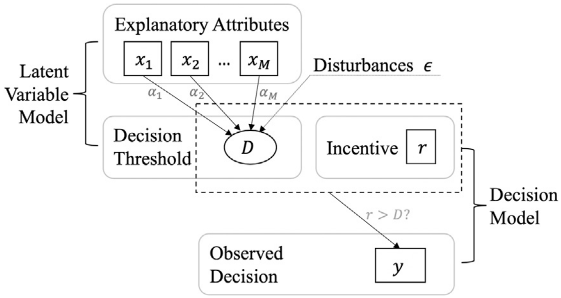

Fig. 3. An illustration of the Latent Decision Threshold (LDT) model. Squares indicate observed variables and the ellipse indicates the latent variable.

over all users are reported in Table 3. We can observe that there are users whose individual-level correlations between rewards and other attributes are high. Note that here we report the absolute values because the magnitude reflects the strength of correlation.

To further examine if the condition in Lemma 1 exists in this data, we build a population-level linear regression model for r using $\boldsymbol{x}$ as predictors, i.e., $r = \gamma_0 + \boldsymbol{\gamma}^\top \boldsymbol{x} + \epsilon$. Table 4 shows the result, clearly suggesting that there is significant correlation between the reward and other attributes. While this analysis is done on the population level, i.e., we pooled all users' data and build one regression model, we also built personalized regression models, i.e., we built a regression model for each individual, using only his/her own data, and a similar observation could be obtained.

Although a common way to deal with multicollinearity among the variables is to apply variable selection techniques before modeling, i.e., to remove the highly correlated variables, it is not suitable in our case because all the variables are important and contextually meaningful although statistically correlated, and the reward r, which is the highly correlated variable, is a key component for the success of personalized reward system and should be included in the model. Principal Component Analysis (PCA) (Wold et al., 1987) is another common practice in dealing with multicollinearity as it projects related variables into a new coordinate system such that the new features are not linearly correlated. However, it will lead to a loss of interpretability as the variables are transformed into new features. Putting all together, we conclude that there is a significant gap between the existing literature with the personalized TDM systems we aim to study. To fill in this gap, we propose a graphical model to characterize the data-generating mechanism in these user–app interactions in the following sections.

3. The Latent Decision Threshold (LDT) model

3.1. Model formulation

In this section, we present the proposed Latent Decision Threshold (LDT) model. Motivated by the limitations of the RUM-based models discussed in Section 2, the LDT model aims to provide a fair characterization of the data-generating mechanism underlying users' decision-making, which is shown in Fig. 3. Unlike the logit models shown in Fig. 2 that flatten the multi-layered mechanism, the LDT model uses a middle layer of latent variable to model the decision threshold and its interaction with the reward r. Thus, LDT model consists of two parts: the latent variable model, and the decision model.

LDT model postulates that each alternative the user faces has a certain cost or difficulty, which is related to the explanatory attributes. For example, in transportation demand management, the user might risk being late for work if asked to depart 10 min later than he/she usually does. It might cost an individual more time or more fuel on the road if he/she has to change the route. And it may be hard for some users to change their schedules (Ben-Elia and Ettema, 2011a,b). Therefore, TDM strategies either design regulations like toll (Nijkamp and Shefer, 1998; Rouwendal and Verhoef, 2006) or increase parking costs (Rothengatter, 1982; Schuitema and Steg, 2008) to reduce the traffic demand, or offer incentives like monetary rewards (Ben-Elia and Ettema, 2011a,b) or free public transport (Bamberg et al., 2003) to compensate such cost and encourage people to switch to a desirable travel plan.

The second part of the LDT model, the decision model, characterizes the comparison mechanism. For a user to accept a choice, the given incentive should be strong enough to offset the cost or difficulty. The tipping point cost or difficulty which is comparable with the reward, is called the **Decision Threshold** (denote as D), and any amount larger than D would change the user's decision. In other words, if the incentive exceeds the threshold, the user will accept the promotion, otherwise, he/she will reject it. Hence, this decision threshold can be considered as the minimum reward the system needs to provide to the user for a given alternative. Eq. (4) shows the decision-making part of the LDT model:

$$y = \begin{cases} 1, & r > D; \\ -1, & r \le D. \end{cases} \tag{4}$$

This design will lead to the following consolidated formulation for the decision model:

$$y(r - D) \ge 0, \tag{5}$$

which can facilitate the development of our optimization solution in Section 3.2.

The first part of the LDT model, the latent variable model, characterizes user preferences on the attributes and captures the relation between the attributes and the latent variable, the decision threshold D. In this work, we use the linear model as shown in Eq. (6).

$$D(\boldsymbol{x}) = \alpha_0 + \alpha_1 x_1 + \cdots + \alpha_M x_M + \epsilon = \alpha_0 + \boldsymbol{\alpha}^\top \boldsymbol{x} + \epsilon, \tag{6}$$

where parameter $\boldsymbol{\alpha}$ reflects user preferences on the attributes. Here we use $\boldsymbol{\alpha}$ to represent the user preferences instead of $\boldsymbol{\beta}$ which has been used in logit models, to distinguish the two because of their different meanings. Their signs imply different interpretations. As discussed in Section 2.2, a positive β_m indicates that attribute x_m is attractive to the user, granting higher utility to the alternative and make it more likely to be chosen. However, a positive α_m means that the decision threshold will be increased, and make it harder for user to accept the promoted alternative. The magnitude of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ also differs, because α_m indicates the "cost" of change for attribute x_m in terms of the reward, but β_m is not comparable with reward unless normalized by β_r, i.e., β_m/β_r.

3.2. Model estimation

As shown in Fig. 3, there are two types of unknown parameters: the preferences on the attributes, $\boldsymbol{\alpha}$, and the latent variable, the decision threshold D. It is important to recognize that if we know $\boldsymbol{\alpha}$, we could readily derive D based on Eq. (6). Thus, D is an intermediate parameter.

As LDT is a graphical model, we could use the Expectation–Maximization (EM) algorithm to estimate the latent variable and the other parameters in iterative steps, i.e., the Expectation step estimates the sufficient statistics of the unobserved variable, given the observed data and current estimates of the coefficients, while the Maximization step takes the estimated complete data and estimates the coefficients (Dempster et al., 1977). However, EM algorithm is not ideal here. First, to use the EM algorithm, a joint likelihood function is needed which requires additional probabilistic assumptions for all the variables including the latent variable. Construction of likelihood function also asks for conditional probability distribution of the observed variables conditional on the latent variable. Also, when the distribution for the latent variable is continuous as here in our case, the computation of integral is usually needed. However, not like in some other graphical models, here we have no closed-form for the E-step due to the unique mechanism outlined in Eq. (4). It will lead to extra computational difficulty since an approximation method is often needed (Lee and Shi, 2001; Wei and Tanner, 1990).

Thus, we propose a computational method that is based on the max-margin learning, as LDT model bears such an interesting structure that could be utilized. We then no need to estimate the intermediate latent variable D. To see that, we can consolidate Eqs. (5) and (6) as one inequality for each binary choice scenario:

$$y(r - \alpha_0 - \boldsymbol{\alpha}^\top \boldsymbol{x}) \ge 0. \tag{7}$$

Here $\boldsymbol{x}$, r and y are the observed attributes, reward, and the final decision, respectively for each binary choice scenario. α_0 and $\boldsymbol{\alpha}$ are individual-specified parameters to be estimated. This reformulation removes the need to involve D, and Eq. (7) encodes all the constraints that we need the estimated LDT model to satisfy.

To enforce regularization to overcome the risk of overfitting, we adopt the principle of max-margin and develop the following formulation (for each individual with all his/her data points):

$$\begin{aligned} \min_{\alpha_0, \boldsymbol{\alpha}, \xi} \quad & \frac{1}{2}\boldsymbol{\alpha}^\top \boldsymbol{\alpha} + C \sum_t \xi_t \\ \text{s.t.} \quad & y_t(r_t - \alpha_0 - \boldsymbol{\alpha}^\top \boldsymbol{x}_t) \ge 1 - \xi_t, \text{ for all } t, \\ & \xi_t \ge 0, \text{ for all } t, \end{aligned} \tag{8}$$

by introducing small violations ξ_t's. This follows the idea of non-separable Support Vector Machine (SVM) to tolerate violations on a certain level (Cortes and Vapnik, 1995; Chen et al., 2004). The objective of this learning formulation is similar as in SVM, to balance the margin of classifying hyperplane and the small violations, controlled by the tuning parameter of C. This algorithm naturally fits into the LDT model and the comparison nature in decision-making when there are incentives as shown in Fig. 3. In addition to that, it can also help us incorporate robustness into the model learning to be resilient against noise, i.e., with the max-margin objective (Chechik et al., 2008) and the introduction of the violation parameters (ξ) to prevent overfitting of the model, especially when data is limited (Floares et al., 2017). The max-margin learning also has very good interpretability, which is very important for understanding user behavior. There are multiple solvers available for this convex optimization problem. In our work, we use CVXR for estimating the model (Grant and Boyd, 2008, 2014).

The constraint structure in Eq. (8) is almost the same as that in the SVM formulations (Cortes and Vapnik, 1995). The difference between the two here lies on the fact that, if we rewrite the LDT formulation in the standard form of SVM, we will have a varying offset in the constraints, i.e., here, because r_t varies from case to case. In summary, the max-margin formulation in Eq. (8) effectively addresses the computational issues in our graphic model that have a unique mechanism of the latent variable. It does not demand additional assumptions in variable distribution or conditional probabilities. Furthermore, the maximized soft margin is capable of obtaining stable estimations with the existence of errors in data. The combination of the advantages from both the graphical model and the computationally efficient max-margin algorithm makes LDT model suitable and effective for modeling user–app interactions.

4. Simulation studies

In this section, we evaluate the performances of the proposed LDT model, and examine our hypothesis about the data-generating mechanism that is beyond the user behavior data to explain the counter-intuitive results as mentioned in Section 2.3. Our approach is to design an experiment that generates data by the data-generating mechanism as shown in Fig. 3 and compare the performances of LDT with the logit models.

4.1. Data generation

The simulation design follows the data-generating mechanism described in Section 3.1 where each alternative consists of a set of explanatory attributes $\boldsymbol{x}$ and a certain amount of reward r, and the users will make the decision based on his/her own preferences $\boldsymbol{\alpha}$ towards the attributes. To account for the various complexities that can be encountered in real world, our simulation experiments include several different aspects.

First, in real-world decision behaviors, some attributes are usually disliked (or always welcomed) by the users, e.g., in transportation, changing of scheduled time and extra toll cost can be regard as barrier attributes (commuters usually do not like it), while time or fuel saving can be encouraging attributes which most people would prefer (Ben-Elia and Ettema, 2011a,b). This characteristic can be reflected in the parameters $\boldsymbol{\alpha}$ as always being positive or negative. For example, we design three attributes in the following study where two of them are barriers (x_1, x_2) and another is an encouraging attribute x_3, and the corresponding parameters will follow:

$$\boldsymbol{\alpha} = [\alpha_1 \geq 0,\ \alpha_2 \geq 0,\ \alpha_3 \leq 0]^\top. \tag{9}$$

This constraint also provides a perspective for evaluating the interpretability of the model, which is further discussed in Section 4.2.

Another layer of complexity is the heterogeneity or diverse cohorts in user preferences, i.e., diverse personal preferences for N different users ($\{\alpha_0^{(n)}, \boldsymbol{\alpha}^{(n)}\},\ n = 1, \ldots, N$). Multiple methods can achieve this setting. In our work, we utilize the Gaussian Mixture Distribution to generate $\boldsymbol{\alpha}^{(n)}$. To be specific, $\boldsymbol{\alpha} \sim N(\boldsymbol{\mu}_{\boldsymbol{\alpha}}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}})$ where $\boldsymbol{\mu}_{\boldsymbol{\alpha}}$ has multiple choices, for example:

$$\boldsymbol{\mu}_{\boldsymbol{\alpha}}^1 = [1, 2, -2]^\top,\ \boldsymbol{\mu}_{\boldsymbol{\alpha}}^2 = [2, 1, -2]^\top,\ \boldsymbol{\mu}_{\boldsymbol{\alpha}}^3 = [2, 2, -1]^\top.$$

Other choices will also work as well. After we have obtained all the $\{\alpha_0^{(n)}, \boldsymbol{\alpha}^{(n)}\}$ from these distributions, we truncate the $\boldsymbol{\alpha}$ that does not fit the condition outlined in Eq. (9) to be 0. With the simulated personal preferences for a user, we further generate the alternatives for the user to choose. An alternative is defined by the attributes (x_1, x_2, x_3), which we generate from uniform distributions, i.e., $x_m \sim Unif(0, 30), m = 1, 2, 3$.

The reward mechanism, i.e., the strategy to assign r for the promotions may differ in different apps, or with different levels of knowledge about the users. To determine the rewards for the alternatives, we design three different reward approaches as follows.

- **Random Reward approach** It randomly gives out rewards regardless of the attributes of the choices. This may represent the situations where very few knowledge about user preferences are known so that the rewards are determined randomly. We generate it using normal distribution where the mean is equal to the average of all generated decision thresholds. The variance of the normal distribution is the same as that of all generated decision thresholds. In this way the overall acceptance rate for the promotions would be close to 0.5, which will ensure we have a balanced dataset, i.e., about half of the promotions would be accepted by users. Based on the aforementioned approach in generating the data, the random rewards are generated via $r \sim N(35, 45^2)$.
- **Contribution Reward approach** Although users preferences vary and are unknown, the "contribution" of any alternative to the transportation system can be determined. For example, we can generate more reward points to those who leave home 20 min earlier than those who leave home 2 min earlier. Specifically, here we can define $r = 10 + 1.67x_1 + 1.67x_2 - 1.67x_3 + \epsilon_c$ with $\epsilon_c \sim N(0, 5^2)$ to ensure we have a balanced dataset (for the same reason explained in the random reward approach).
- **Predictive Reward approach** This reward mechanism builds on the assumption that the app system has obtained adequate knowledge about every user's preferences and can assign rewards that equal the decision threshold. Here, we design $r_P = \alpha_0^{(n)} + \boldsymbol{\alpha}^{(n)\top}\boldsymbol{x} + \epsilon_P$ for any user n. The random term $\epsilon_P \sim N(0, 5^2)$ is used here for the same reason mentioned in the contribution reward approach.

4.2. Performance metrics

We will compare the proposed LDT model with MLM, since both build models at individual level. A population-level logit model is also estimated to provide another baseline for the potential interest of readers. As all the parameters involved in our models are numerical, we use the root of mean squared error ($RMSE$) to evaluate estimation performances of the models. For example, to evaluate how well LDT estimates the latent decision threshold (D), the $RMSE$ is defined as:

$$RMSE = \sqrt{\frac{1}{N \times T} \sum_{n,t} (D_t^{(n)} - \hat{D}_t^{(n)})^2}, \tag{10}$$

where n refers to the nth user and t refers to the tth decision made by this user. Smaller $RMSE$ indicates that the model can retrieve the latent variable (decision threshold, D) better, in other words, the model performs better. For logit models, we estimate the reward where a user will accept the promotion (decision threshold, D), i.e., by solving r for equation $Pr(y = 1) = Pr(y = -1) = 0.5$ with estimated coefficients. This seems similar to the concept "tipping point" used in the LDT model, but it is worthy of pointing out that the hypothesized data-generating mechanisms of the two models are quite different as shown in Figs. 2 and 3. Besides $RMSE$, we also use the mean absolute error (MAE) and mean absolute percentage error ($MAPE$) to evaluate how well the models can estimate the latent variables and user preferences. For example, for a parameter α, MAE and $MAPE$ are defined as follows:

$$MAE = \frac{1}{N} \sum_{n} |\alpha^{(n)} - \hat{\alpha}^{(n)}|, \quad MAPE = \frac{1}{N} \sum_{n} \frac{|\alpha^{(n)} - \hat{\alpha}^{(n)}|}{|\alpha^{(n)}|}. \tag{11}$$

In order to have a fair comparison between LDT and the logit models, we normalize $\hat{\boldsymbol{\beta}}$'s by $\hat{\beta}_r$ so that it also represents the "cost" of change in terms of reward.

On the other hand, as shown in Table 2, for this personalized TDM system where users interact with the system, the logit models meet difficulties in interpreting the estimated coefficients, i.e., a large proportion of them are counter-intuitive. We also evaluate this aspect of interpretability in our study using the sign of the estimated coefficients, i.e., as outlined in Eq. (9). If $\hat{\boldsymbol{\alpha}}$ shows different signs as in Eq. (9), it is a sign error. For each element in $\hat{\boldsymbol{\alpha}}$, we calculate the fraction of sign errors (sign error rate, SER). Note that logit models also have a parameter for the reward, β_r, which should be positive (Ben-Elia and Ettema, 2011a,b). Thus, we also calculate the SER for the reward coefficient for the logit models.

Finally, since the decision-making behavior is cast as a binary classification form, for each learned model, we also report the classification accuracy, recall, and precision on the testing set. Classification accuracy is an overall evaluation, which is defined as the rate of correct classifications. Recall and precision represent two different aspects of the accuracy. Recall, also known as sensitivity, indicates how well the model can detect the positive. It is defined as the fraction of true positives (actually positive and also classified as positive) among all those cases which are actually positive. Precision represents the accuracy within positive classifications. It is defined as the fraction of true positives among those cases which are classified as positive.

All the experiments are conducted on R (version 4.0.2) on the platform of x86_64-apple-darwin17.0 (64-bit) under macOS 10.15.6 (2.2 GHz 6-Core Intel Core i7, 6 GB 2400 MHz DDR4). The run times for the simulation experiments are reported in the results as well, i.e., in Tables 5–7. Our code is publicly available on GitHub repository (https://github.com/feng-jings/LDTmodel).

4.3. Results

For each reward approach, we run the simulation experiment including the data generation and model fitting for 100 replicates. The overall performance over the 100 replicates is reported using the average of the metrics listed in Section 4.2. Simulation results on $N = 1000$ users are shown in Tables 5 to 7 that correspond to the three reward approaches, respectively.

An overall observation is that the proposed LDT model outperforms MLM in parameter estimation (reflected by $RMSE$, MAE and $MAPE$) and interpretability (reflected by SER). Both models have similar classification performances (reflected by classification accuracy, recall and precision), and are superior than the population-level logit model. Specifically, from the perspective of coefficient estimation, the estimation errors of LDT model is smaller than that of MLM. The SER also shows MLM often lead to more counter-intuitive results. Among the three reward approaches, the mixed logit model performs worse under the predictive reward approach case than the other two in terms of the coefficient estimation and interpretability.

It is worthy of pointing out that, when normalizing the parameters and calculating the tipping points (decision threshold) for MLM, it could suffer from numerical instability because of very low $\hat{\beta}_r$, so that in deriving the statistics reported in Tables 5–7 we have to drop some experiments that have enormously large metrics. In other words, the $RMSE$, MAE and $MAPE$ values of the logit models shown in the Tables 5–7 are actually underestimated (some will be over 10^4).

The observations aforementioned are consistent with the analysis presented in Section 2.3. According to Lemma 1, when variables are correlated, i.e., the reward is related to other attributes, the estimated coefficients by the logit and mixed logit models may not accurately report the true preference parameters as shown by the high estimation errors measured by $RMSE$, MAE, $MAPE$, and suffer from an interpretability problem presented by SER. However, although its estimated parameters may be inaccurate individually, the estimated model itself may still lead to the same utility so that the classification performances are still comparable. This observation is further reinforced by the experiment using the predictive reward approach (i.e., that computes the reward based on each user's own preferences, so it holds the highest correlation of the three reward approaches) which shows the model performances of MLM under this approach is the worst according to Tables 5–7.

Table 5
Model performances of Logit, MLM, and LDT models, over 100 replicates of simulation using **random reward approach**.

	Logit	MLM	LDT
Time (secs)	**0.8754**	191.5911	602.3220
Testing classification accuracy	0.7908	0.8862	**0.8909**
Recall	0.8044	0.8820	**0.8918**
Precision	0.8068	0.8811	**0.8942**
RMSE: Decision threshold	41.7415	38.4610	**16.1519**
RMSE: Coefficient #1	1.6482	1.3611	**0.9928**
RMSE: Coefficient #2	1.6436	1.3572	**0.9930**
RMSE: Coefficient #3	1.6448	1.3573	**1.0669**
MAE: Decision threshold	34.8529	31.8577	**12.9827**
MAE: Coefficient #1	1.4157	1.1649	**0.8230**
MAE: Coefficient #2	1.4154	1.1647	**0.8264**
MAE: Coefficient #3	1.4197	1.1590	**0.8720**
MAPE: Decision threshold	0.9307	0.8277	**0.4509**
MAPE: Coefficient #1	0.9086	0.7341	**0.5533**
MAPE: Coefficient #2	0.9086	0.7351	**0.5561**
MAPE: Coefficient #3	0.9078	0.7325	**0.5699**
SER: Coefficient #1	–	0.1188	**0.1090**
SER: Coefficient #2	–	0.1193	**0.1070**
SER: Coefficient #3	–	0.1210	**0.1100**
SER: Coefficient reward	–	0.0260	–

Table 6
Model performances of Logit, MLM, and LDT models, over 100 replicates of simulation using **contribution reward approach**.

	Logit	MLM	LDT
Time (secs)	**0.8014**	167.1454	621.2343
Testing classification accuracy	0.5699	0.8797	**0.8869**
Recall	0.6472	0.8858	**0.8924**
Precision	0.5831	0.8833	**0.8945**
RMSE: Decision threshold	41.7543	38.1875	**15.5102**
RMSE: Coefficient #1	1.6493	1.3339	**0.8838**
RMSE: Coefficient #2	1.6444	1.3248	**0.8895**
RMSE: Coefficient #3	1.6480	1.3238	**0.9383**
MAE: Decision threshold	34.8650	31.6239	**11.9106**
MAE: Coefficient #1	1.4168	1.1471	**0.6950**
MAE: Coefficient #2	1.4162	1.1464	**0.6993**
MAE: Coefficient #3	1.4161	1.1345	**0.7034**
MAPE: Decision threshold	0.9313	0.8236	**0.4230**
MAPE: Coefficient #1	0.9097	0.7215	**0.5100**
MAPE: Coefficient #2	0.9093	0.7256	**0.5110**
MAPE: Coefficient #3	0.9110	0.7250	**0.4989**
SER: Coefficient #1	–	0.0817	**0.0615**
SER: Coefficient #2	–	0.0753	**0.0613**
SER: Coefficient #3	–	0.0806	**0.0614**
SER: Coefficient reward	–	0.0401	–

A closer look at the coefficients estimated by MLM and LDT models provide an additional explanation of the model performances. Fig. 4 illustrates the details of coefficient estimations of MLM and LDT model using random reward approach (in one simulation replicate). Note that, for reason that has been explained in Section 3.1 (i.e., the paragraph below Eq. (6)), in order to be comparable with the coefficients of the LDT model, the coefficients in MLM need to be flipped and normalized (by $\hat{\beta}_r^{(n)}$). The red lines refer to zero. Compared with the real values on the top, it can be observed that besides the higher fraction of sign violation, the estimations from MLM are more dispersed.

Fig. 5 further shows how well the two models perform on the estimation of the latent decision threshold (D) using predictive reward approach. The first two panels correspond to MLM, where the middle panel zooms in the shadowed area of the left panel to give a better view of how the majority of the results of the MLM estimation deviates from the true model. The right panel corresponds to the LDT model. It can be seen that the proposed LDT method could yield a nice estimation of the latent decision thresholds (D) of the alternatives. It implies that the LDT model is a better choice in explaining user decision behavior and has promising potential to be used as the analytic foundation where personalized promotion algorithms could be developed.

Table 7
Model performances of Logit, MLM, and LDT models, over 100 replicates of simulation using **predictive reward approach**.

	Logit	MLM	LDT
Time (secs)	**0.8084**	176.4998	610.6803
Testing classification accuracy	0.6296	0.6581	**0.7566**
Recall	0.7846	0.7639	**0.7927**
Precision	0.6558	0.6905	**0.7947**
RMSE: Decision threshold	43.6269	42.9456	**4.3299**
RMSE: Coefficient #1	1.7137	1.6719	**0.2380**
RMSE: Coefficient #2	1.7086	1.6667	**0.2387**
RMSE: Coefficient #3	1.6988	1.6521	**0.2948**
MAE: Decision threshold	36.6210	35.9834	**3.3852**
MAE: Coefficient #1	1.4790	1.4397	**0.1918**
MAE: Coefficient #2	1.4787	1.4396	**0.1929**
MAE: Coefficient #3	1.4655	1.4213	**0.2290**
MAPE: Decision threshold	0.9968	0.9736	**0.1441**
MAPE: Coefficient #1	0.9795	0.9424	**0.1367**
MAPE: Coefficient #2	0.9790	0.9424	**0.1382**
MAPE: Coefficient #3	0.9661	0.9373	**0.1638**
SER: Coefficient #1	–	0.0635	**0.0556**
SER: Coefficient #2	–	0.0576	**0.0537**
SER: Coefficient #3	–	0.0621	**0.0451**
SER: Coefficient reward	–	0.3327	–

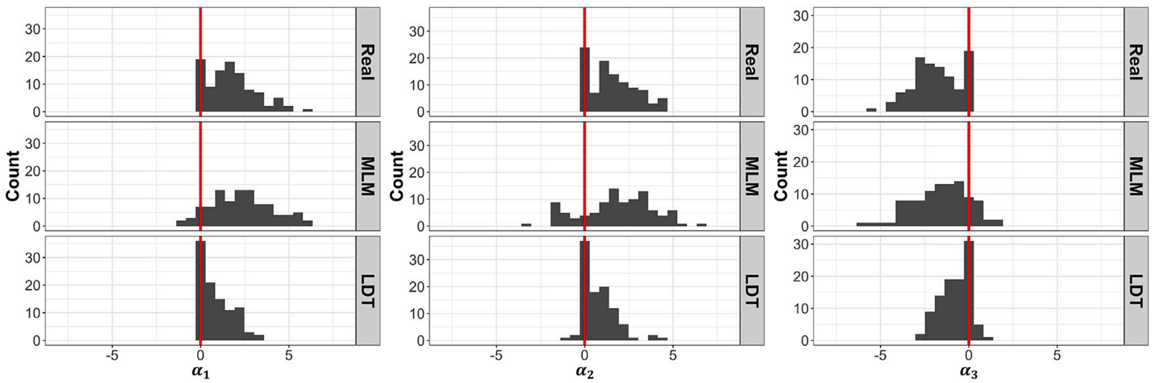

Fig. 4. Coefficient estimation performances of MLM and LDT model, on the simulated dataset using random reward approach (extreme values from unstable estimates of the MLM model have been eliminated for better presentation).

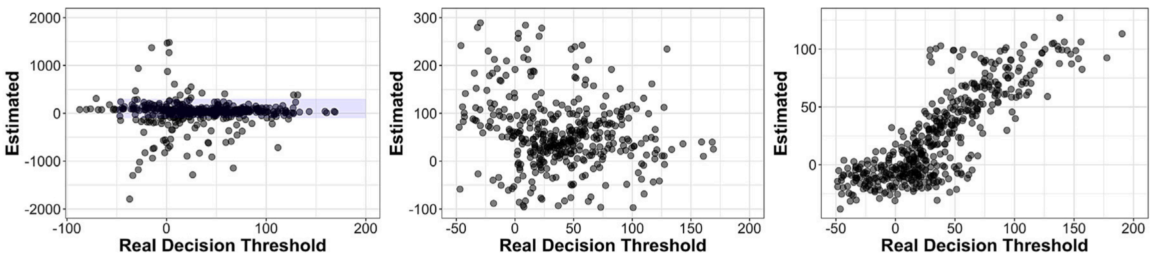

Fig. 5. Latent variable (the decision threshold) estimation performances of MLM and LDT model, on the simulated dataset using predictive reward approach. Left panel: latent variable estimation given by MLM; Middle panel: zooming in the shadowed area of the left panel; Right panel: latent variable estimation given by the LDT model.

5. Real case study results

In this section, we apply our new LDT model on a real-world dataset collected from the personalized TDM system (Zhu et al., 2019) depicted in Section 2.1. After proper cleaning of the data, it consists of 828 respondents and each of them was asked to make 13 rounds of decisions between two alternatives, i.e., whether accept a promotion or stick to the original travel plan. There were

Table 8
Model performances of Logit, MLM, and LDT models, on the personalized TDM system data.

	Logit	MLM	LDT
Testing classification accuracy	0.6677	0.7834	**0.7920**
Recall	0.8152	0.8389	**0.8640**
Precision	0.7329	0.8300	**0.8392**
SER: Coefficient SDE	–	0.1581	**0.1422**
SER: Coefficient SDL	–	**0.0176**	0.1314
SER: Coefficient TTS	–	0.2125	**0.1315**
SER: Coefficient reward	–	0.3884	**–**
Standard Deviation: Intercept	–	98.9605	**24.3670**
Standard Deviation: Coefficient SDE	–	7.6052	**1.0097**
Standard Deviation: Coefficient SDL	–	7.9815	**1.7143**
Standard Deviation: Coefficient TTS	–	17.6733	**2.0290**

Table 9
The answers of User ID 625 (negative $\hat{\alpha}_{SDE}$).

Question No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Delay early (min)	30	0	10	0	0	0	0	60	0	30	60	30	10
Delay late (min)	0	30	0	30	10	30	10	0	30	0	0	0	0
Time save (min)	2	2	2	5	2	2	5	2	5	2	5	5	2
Reward (points)	20	20	20	20	20	40	20	20	40	10	20	20	20
Decision	1	−1	1	1	1	1	1	1	1	1	1	1	1

387 female respondents and 441 male respondents. Commuters with different daily commuting times are investigated (range from 10 to 60 min) and their average daily commuting time was 31.58 min (Zhu et al., 2019). Promotions for a user are designed based on the user's self-reported background. The attributes that characterize the alternatives include departure time change, i.e., either several minutes earlier (x_{SDE}) or several minutes later (x_{SDL}), and travel time saving (x_{TTS}). The system will determine the amount of reward points r needed for the potential acceptance of the promoted alternative to encourage the user to accept the choice. For each user, we use the first 10 rounds as the training data and the last 3 as the testing data. Among the 828 individuals, there are 174 of them who accepted all the first 10 promotions or rejected all. These users are also excluded from our analysis.

5.1. Results of the logit and LDT models

Table 8 shows the results of the population-level logit model, MLM, and the LDT model. From the perspective of classification accuracy, MLM and the LDT model are similar. The LDT model slightly outperforms MLM. But in terms of interpretability, our LDT method is much better. For example, based on MLM there were 38.84% of the users who do not want or even dislike rewards (i.e., $\hat{\beta}_r < 0$). This is counter-intuitive and implies that more rewards only discourage users in accepting promotions. Equally puzzled is that based on MLM there were 21.25% of the users who do not like travel time saving (i.e., $\hat{\beta}_{TTS}$<0). The LDT model also shows that some users had this counter-intuitive signs, but the percentage is much smaller, and in the following Section 5.2 we will dive in this paradox and identify the peculiar data characteristics of this dataset that is probably responsible for why even LDT also showed a slight percentage of counter-intuitive result. Last but not least, the result shows that the standard deviations of the estimated parameters based on the LDT model are much smaller than those of MLM. This is consistent with our observation in Fig. 4 that the coefficients estimated by MLM are more dispersed.

Note here that the metric SER (sign error rate) does not necessarily mean that the sign of the estimated coefficient is wrong, because the underlying true preferences are unknown. It indicates results that are counter-intuitive, and points out directions for more investigations.

5.2. Discover behavior patterns

To show why for some users even the LDT model shows counter-intuitive signs, it is worthy of examining their raw data. Tables 9 and 10 are data from two users.

For User 625, we notice that $\hat{\alpha}_{SDE} < 0$, indicating that the user preferred departing earlier. It is easy to see why the LDT model learned a negative value of $\hat{\alpha}_{SDE}$, i.e., this user only rejected promotion No.2 and accepted all the others. By comparing this rejected promotion with choices No.1 and No.8, we can see that they all offered the same time saving and reward points, and the only difference was the change of departure time. The respondent rejected No.2 that asked to departure later but accepted the other two which required departing earlier. Therefore, it is reasonable that $\hat{\alpha}_{SDE}$ is negative in the estimated LDT model, although it is unknown if the user behaved rationally when making these decisions. This is one of the obstacles of small sample size that the estimation may largely dependent on only one or few observations.

Similarly, User 2078 accepted only one promotion which is No.2. It is interesting to notice that the same user rejected No.9 which was the same as No.2 except that No.9 had more time saving. This is the cause of the counter-intuitive value of $\hat{\alpha}_{TTS}$, i.e., which

Table 10
The answers of User ID 2078 (negative $\hat{\alpha}_{SDL}$ and positive $\hat{\alpha}_{TTS}$).

Question No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Delay early (min)	30	0	30	10	10	10	30	30	0	0	0	10	30
Delay late (min)	0	30	0	0	0	0	0	0	30	10	10	0	0
Time save (min)	5	5	5	5	5	10	10	10	10	5	5	5	2
Reward (points)	20	20	30	20	30	70	20	30	20	20	20	80	30
Decision	−1	1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1

Table 11
The answers of User ID 2361 (the only user with all three coefficients counter-intuitive).

Question No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Depart early (min)	30	0	30	30	0	30	30	30	0	30	60	60	30
Depart late (min)	0	30	0	0	10	0	0	0	10	0	0	0	0
Time save (min)	10	10	10	10	10	10	20	10	20	10	10	10	10
Reward (points)	10	10	20	30	10	40	10	50	10	60	40	50	20
Decision	1	1	1	1	−1	1	−1	1	−1	1	−1	1	1

turned out to be negative. Further, if we compare No.2 with No.10, we can see that the user rejected the one requiring less schedule delay late and this might be the cause of a negative $\hat{\alpha}_{SDL}$. Having seen these patterns, we conclude that these unexpected signs were largely caused by the irrational behavior of the users.

Among all the 654 investigated users, User 2361 is the only one whose coefficients are all counter-intuitive based on the LDT model. Table 11 shows the data of this user. It can be seen that this user clearly followed a pattern, that was to accept the choice when the given reward is higher than the choice before, i.e., for No.3 and No.4, he/she accepted them because they offered higher reward points than the first two rounds. Then for No.5, however, even though it was the same as the second promotion except for less departure time change, he/she rejected it because he/she had seen higher reward points (20 and 30). Similar behavior can be observed in later decisions. The user accepted the ones offering 40, 50 and 60 reward points sequentially and rejected those with only 10 points. Thus, the learned LDT model, while showing counter-intuitive results, captured what was in the data.

6. Conclusion and discussion

The rapid proliferation of smart, personal technologies has given birth to smart Transportation Demand Management (TDM) systems that can give personalized incentives to users. This personalization capacity builds on accurate modeling of user behavior; however, the widely-used discrete choice models often assume that the data collection is independent with the user behavior. On the contrary, when users interact with the apps, the real potential and also the real challenge for modeling user behavior is that the apps not merely are data collection tools, but also change users' behaviors. We propose the Latent Decision Threshold (LDT) model that aims to provide a fair characterization of the user–app interaction process and the decision-making behavior by the user in this new environment. The LDT model is a novel graphical model, and a further analysis of its structure leads us to develop an efficient learning algorithm based on an interesting connection between the graphical model with max-margin learning. Extensive simulation studies and a real-world application show that the LDT model outperforms the logit models in both model estimation and interpretability.

There are a few directions that we may further improve the LDT model. For example, constrained by the limited sample size in many real applications, the proposed LDT model considers only linear relation between the attributes and the latent decision threshold, which could be an oversimplification of user's behavior. Also, the individual-level LDT models in this work are constructed independently for each user and assumes that each user's preferences are stable. These assumptions could be further relaxed, and the extensions of our current model could be developed accordingly. Although this work focuses on binary choice cases, it can be extended to multiple choice problems as well, i.e., one strategy is to learn from existing literature that had established techniques to convert a multiple-choice problem in to formulations with binary outcomes (Allwein et al., 2000; Galar et al., 2011; Rocha and Goldenstein, 2013). In the future, we will also incorporate intra-individual relations such as time-varying preferences or inter-individual relations such as latent class models or personalized learning models to enrich our LDT model.

CRediT authorship contribution statement

Jingshuo Feng: Conceptualization, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **Shuai Huang:** Validation, Supervision, Writing - review & editing. **Cynthia Chen:** Data curation, Investigation, Writing - review & editing.

Appendix. Proof of Lemma 1

We provide the proof of Lemma 1 in this section.

Lemma 1. *In logit models, when the reward r is linearly related to the attributes $\boldsymbol{x}$, there will be an infinite set of estimated coefficients $\hat{\boldsymbol{\beta}}$ that can lead to same utilities, and therefore, all maximize the log-likelihood of the model. In some of these solutions, $\hat{\beta}_r$ may be negative.*

Proof. To prove this, we first rewrite the coefficient $\boldsymbol{\beta}$ as $\boldsymbol{\beta} = [\boldsymbol{\beta}_{\boldsymbol{x}}^\top, \beta_r]^\top$, so the systematic utility $V = \beta_0 + \boldsymbol{\beta}_{\boldsymbol{x}}^\top \boldsymbol{x} + \beta_r r$. For the logit models like the one in Eq. (2), the log-likelihood to be maximized is:

$$\max_{\boldsymbol{\beta}_{\boldsymbol{x}}, \beta_r} \quad \sum_t \log\left[1 + \exp(\beta_0 + \boldsymbol{\beta}_{\boldsymbol{x}}^\top \boldsymbol{x}_t + \beta_r r_t)\right] - y_i(\beta_0 + \boldsymbol{\beta}_{\boldsymbol{x}}^\top \boldsymbol{x}_t + \beta_r r_t), \tag{12}$$

where t indicates the data points. Suppose that $\{\beta_0^*, \boldsymbol{\beta}_{\boldsymbol{x}}^*, \beta_r^*\}$ is one optimal solution of the log-likelihood function. When r_t is linearly related to $\boldsymbol{x}_t$, which we can denote as $r_t = \gamma_0 + \boldsymbol{\gamma}^\top \boldsymbol{x}_t$, for any real number $\delta \neq 1$, we will have:

$$r_t = \frac{\gamma_0}{1-\delta} + \frac{\boldsymbol{\gamma}^\top}{1-\delta}\boldsymbol{x}_t - \frac{\delta}{1-\delta} r_t. \tag{13}$$

Thus, we can define an infinite set of optimal solutions by noticing that

$$\beta_0^* + \boldsymbol{\beta}_{\boldsymbol{x}}^{*\top} \boldsymbol{x}_t + \beta_r^* r_t = (\beta_0^* + \frac{\beta_r^* \gamma_0}{1-\delta}) + (\boldsymbol{\beta}_{\boldsymbol{x}}^* + \frac{\beta_r^* \boldsymbol{\gamma}}{1-\delta})^\top \boldsymbol{x}_t + \frac{-\beta_r^* \delta}{1-\delta} r_t, \tag{14}$$

i.e., here, $\{\beta_0^{**} = (\beta_0^* + \frac{\beta_r^* \gamma_0}{1-\delta}), \boldsymbol{\beta}_{\boldsymbol{x}}^{**} = (\boldsymbol{\beta}_{\boldsymbol{x}}^* + \frac{\beta_r^* \boldsymbol{\gamma}}{1-\delta})^\top, \beta_r^{**} = \frac{-\beta_r^* \delta}{1-\delta}\}$, is a different optimal solution that achieves the same optimality as $\{\beta_0^*, \boldsymbol{\beta}_{\boldsymbol{x}}^*, \beta_r^*\}$.

Because of the arbitrariness of the optimal solutions, it is possible that $\hat{\beta}_r = \frac{-\beta_r^* \delta}{1-\delta} < 0$ (e.g., when $0 < \delta < 1$ and $\beta_r^* > 0$). The estimated coefficients for other attributes $\hat{\boldsymbol{\beta}}_{\boldsymbol{x}}$ will also be shifted by the amount of $\frac{\beta_r^* \boldsymbol{\gamma}}{1-\delta}$, and cause some unexpected sign errors. □

References

Allwein, E.L., Schapire, R.E., Singer, Y., 2000. Reducing multiclass to binary: A unifying approach for margin classifiers. J. Mach. Learn. Res. 1 (Dec), 113–141.
Arian, A., Ermagun, A., Zhu, X., Chiu, Y.C., 2018. An empirical investigation of the reward incentive and trip purposes on departure time behavior change. In: Advances in Transport Policy and Planning, Vol. 1. Elsevier, pp. 145–167.
Bamberg, S., Rölle, D., Weber, C., 2003. Does habitual car use not lead to more resistance to change of travel mode? Transportation 30 (1), 97–108.
Bastin, F., Cirillo, C., Toint, P.L., 2010. Estimating nonparametric random utility models with an application to the value of time in heterogeneous populations. Transp. Sci. 44 (4), 537–549.
Ben-Akiva, M., de Palma, A., McFadden, D., Abou-Zeid, M., Chiappori, P.A., de Lapparent, M., Durlauf, S.N., Fosgerau, M., Fukuda, D., Hess, S., et al., 2012. Process and context in choice models. Mark. Lett. 23 (2), 439–456.
Ben-Akiva, M.E., Lerman, S.R., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, Vol. 9. MIT Press.
Ben-Elia, E., Ettema, D., 2011a. Changing commuters' behavior using rewards: A study of rush-hour avoidance. Transp. Res. F: Traffic Psychol. Behav. 14 (5), 354–368.
Ben-Elia, E., Ettema, D., 2011b. Rewarding rush-hour avoidance: A study of commuters' travel behavior. Transp. Res. A 45 (7), 567–582.
Chaptini, B.H., 2005. Use of Discrete Choice Models with Recommender Systems (Ph.D. thesis). Massachusetts Institute of Technology.
Chechik, G., Heitz, G., Elidan, G., Abbeel, P., Koller, D., 2008. Max-margin classification of data with absent features. J. Mach. Learn. Res. 9 (Jan), 1–21.
Chen, C., Ma, J., Susilo, Y., Liu, Y., Wang, M., 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. Transp. Res. C 68, 285–299.
Chen, D.R., Wu, Q., Ying, Y., Zhou, D.X., 2004. Support vector machine soft margin classifiers: error analysis. J. Mach. Learn. Res. 5 (Sep), 1143–1175.
Chorus, C.G., Rose, J.M., Hensher, D.A., 2013. Regret minimization or utility maximization: it depends on the attribute. Environ. Plan. B: Plann. Des. 40 (1), 154–169.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297.
Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. Ser. B Stat. Methodol. 39 (1), 1–22.
Elias, W., Shiftan, Y., 2012. The influence of individual's risk perception and attitudes on travel behavior. Transp. Res. A 46 (8), 1241–1251.
Floares, A., Ferisgan, M., Onita, D., Ciuparu, A., Calin, G., Manolache, F., 2017. The smallest sample size for the desired diagnosis accuracy. Int. J. Oncol. Cancer Ther..
Galar, M., Fernández, A., Barrenechea, E., Bustince, H., Herrera, F., 2011. An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. Pattern Recognit. 44 (8), 1761–1776.
Giuliano, G., 1992. Transportation demand management: promise or panacea? J. Am. Plan. Assoc. 58 (3), 327–335.
Grant, M.C., Boyd, S.P., 2008. Graph implementations for nonsmooth convex programs. In: Recent Advances in Learning and Control. Springer, pp. 95–110.
Grant, M., Boyd, S., 2014. CVX: Matlab software for disciplined convex programming, version 2.1.
Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transp. Res. B 37 (8), 681–698.
Guevara, C.A., Ben-Akiva, M., 2010. Addressing endogeneity in discrete choice models: Assessing control-function and latent-variable methods. In: Choice Modelling: the State-Of-The Art and the State-Of-Practice. pp. 353–371.
Guevara, C.A., Hess, S., 2019. A control-function approach to correct for endogeneity in discrete choice models estimated on SP-off-RP data and contrasts with an earlier FIML approach by Train & Wilson. Transp. Res. B 123, 224–239.
Habibian, M., Kermanshah, M., 2013. Coping with congestion: Understanding the role of simultaneous transportation demand management policies on commuters. Transp. Policy 30, 229–237.
Hancock, T.O., Hess, S., Choudhury, C.F., 2018. Decision field theory: Improvements to current methodology and comparisons with standard choice modelling techniques. Transp. Res. B 107, 18–40.
Hensher, D.A., 1994. Stated preference analysis of travel choices: the state of practice. Transportation 21 (2), 107–133.
Hensher, D.A., Greene, W.H., 2003. The mixed logit model: the state of practice. Transportation 30 (2), 133–176.

Hess, S., Daly, A., Batley, R., 2018. Revisiting consistency with random utility maximisation: theory and implications for practical work. Theory and Decision 84 (2), 181–204.
Jiang, H., Qi, X., Sun, H., 2014. Choice-based recommender systems: a unified approach to achieving relevancy and diversity. Oper. Res. 62 (5), 973–993.
Katzev, R., 2003. Car sharing: A new approach to urban transportation problems. Anal. Soc. Issues Public Policy 3 (1), 65–86.
Koller, D., Friedman, N., 2009. Probabilistic Graphical Models: Principles and Techniques. MIT Press.
Lee, S.Y., Shi, J.Q., 2001. Maximum likelihood estimation of two-level latent variable models with mixed continuous and polytomous data. Biometrics 57 (3), 787–794.
Mayeres, I., Ochelen, S., Proost, S., 1996. The marginal external costs of urban transport. Transp. Res. D 1 (2), 111–130.
McFadden, D., et al., 1973. Conditional Logit Analysis of Qualitative Choice Behavior. Institute of Urban and Regional Development, University of California ....
McFadden, D., et al., 1975. The revealed preferences of a government bureaucracy: Theory. Bell J. Econ. 6 (2), 401–416.
Meyer, M.D., 1999. Demand management as an element of transportation policy: using carrots and sticks to influence travel behavior. Transp. Res. A 33 (7–8), 575–599.
Milton, J.C., Shankar, V.N., Mannering, F.L., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accid. Anal. Prev. 40 (1), 260–266.
Möser, G., Bamberg, S., 2008. The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence. J. Environ. Psychol. 28 (1), 10–26.
Nijkamp, P., Shefer, D., 1998. Urban transport externalities and pigouvian taxes: A network approach. In: Road Pricing, Traffic Congestion and the Environment. Edgar Elgar, pp. 171–189.
Rocha, A., Goldenstein, S.K., 2013. Multiclass from binary: expanding one-versus-all, one-versus-one and ECOC-based approaches. IEEE Trans. Neural Netw. Learn. Syst. 25 (2), 289–302.
Rothengatter, T., 1982. The effects of police surveillance and law enforcement on driver behaviour. Curr. Psychol. Rev. 2 (3), 349–358.
Rouwendal, J., Verhoef, E.T., 2006. Basic economic principles of road pricing: From theory to applications. Transp. Policy 13 (2), 106–114.
Roweis, S., Ghahramani, Z., 1999. A unifying review of linear Gaussian models. Neural Comput. 11 (2), 305–345.
Schuitema, G., Steg, L., 2008. The role of revenue use in the acceptability of transport pricing policies. Transp. Res. F: Traffic Psychol. Behav. 11 (3), 221–231.
Stopher, P.R., 2004. Reducing road congestion: a reality check. Transp. Policy 11 (2), 117–131.
Train, K.E., 2009. Discrete Choice Methods with Simulation. Cambridge University Press.
Wei, G.C., Tanner, M.A., 1990. A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. J. Am. Statist. Assoc. 85 (411), 699–704.
Wold, S., Esbensen, K., Geladi, P., 1987. Principal component analysis. Chemometr. Intell. Lab. Syst. 2 (1–3), 37–52.
Zhu, X., Wang, F., Chen, C., Reed, D.D., 2019. Personalized incentives for promoting sustainable travel behaviors. Transp. Res. C.