Transportation Research Part B 58 (2013) 16–43

Contents lists available at ScienceDirect

Transportation Research Part B

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

Daily activity pattern recognition by using support vector machines with multiple classes

Mahdieh Allahviranloo *, Will Recker [1]

Department of Civil and Environmental Engineering, University of California, Irvine, 4000 Anteater Instruction and Research Building, Irvine, CA 92697-3600, USA

ARTICLE INFO

Article history:
Received 2 January 2013
Received in revised form 10 September 2013
Accepted 16 September 2013

Keywords:
Activity pattern recognition
Activity sequence
Support Vector Machines (SVMs)
Hidden Markov Models (HMMs)

ABSTRACT

The focus of this paper is to learn the daily activity engagement patterns of travelers using Support Vector Machines (SVMs), a modeling approach that is widely used in Artificial intelligence and Machine Learning. It is postulated that an individual's choice of activities depends not only on socio-demographic characteristics but also on previous activities of individual on the same day. In the paper, Markov Chain models are used to study the sequential choice of activities. The dependencies among activity type, activity sequence and socio-demographic data are captured by employing hidden Markov models. In order to learn model parameters, we use sequential multinomial logit models (MNL) and multi-class Support Vector Machines (K-SVM) with two different dependency structures. In the first dependency structure, it is assumed that type of activity at time '*t*' depends on the last previous activity and socio-demographic data, whereas in the second structure we assume that activity selection at time '*t*' depends on all of the individual's previous activity types on the same day and socio-demographic characteristics. The models are applied to data drawn from a set of California households and a comparison of the accuracy of estimation of activity types and their sequence in the agenda, indicates the superiority of K-SVM models over MNL. Additionally, we show that accuracy in estimating activity patterns increases using different sets of explanatory variables or tuning parameters of the kernel function in K-SVM.

1. Introduction

In recent years, activity based models have begun to be employed for travel demand forecasting purposes and in previous years numerous research efforts have been reported, including: studying characteristics of activity patterns by discrete choice utility maximization (see Arentze et al., 2011; Bhat and Srinivasan, 2005; Bhat, 2005; Bowman and Ben Akiva, 2000); mathematical programming models (Gan and Recker, 2013; Recker, 1995, 2001; Recker et al., 1986a,b; Recker and Parimi, 1999); modeling agenda formation and characteristics, ranging from need-based models (Arentze and Timmermans, 2009) and sequential agenda formation formulations (Kitamura et al., 1997; Pendyala et al., 1997.), to such highly complicated comprehensive models as CEMDAP (Bhat, 2004), ALBATROSS (Arentze and Timmermans, 2000) and TRANSIM (Smith et al., 1995). Bhat and his group of researchers use Multiple Discrete Continuous Extreme Value (MDCEV) methods to estimate the activity duration for household members for the residents of Southern California (Bhat et al., 2013).

Borrowing from the biological sciences, researchers have used sequential alignment techniques to compare strings of activity patterns. Similar to chromosomes comparison, the technique finds the minimum number of operations required

* Corresponding author.
E-mail addresses: mallahvi@uci.edu (M. Allahviranloo), wwrecker@uci.edu (W. Recker).
[1] Tel.: +1 949 824 5989; fax: +1 949 824 8385.

0191-2615/$ - see front matter
http://dx.doi.org/10.1016/j.trb.2013.09.008

to align two strings to each other. Joh and his colleagues (Joh et al., 2001, 2002) used a multidimensional sequential alignment that incorporates such different attributes of pattern as activity type, location, mode and accompanying person. A genetic algorithm was developed to overcome the combinatorial nature of the problem. More recently, Sammour et al. (2012) used sequential alignment to validate the performance of the ALBATROSS model.

Arguably, preceding activities of individuals are influential factors in the choice of subsequent activities in the agenda (Kitamura et al., 1997). Ostensibly, the human brain's inability to enumerate all possible combinations of activity patterns to generate a chain of activities simultaneously as well as any change in the psychological mood of the individual during different times of the day makes sequential activity pattern generation more reasonable and compatible with human decision making procedure (Kitamura et al., 1997) than does the simultaneity assumption that forms the basis of most models that have been applied widely in the study of various aspects of activity patterns.

Although relatively uncommon, there nonetheless have been some notable studies that have attempted to build upon sequential decision making in activity formation. Kitamura et al. (1997) used fully conditional probability models to generate sequential activity location, type and scheduling given the previous activities during the day. In analyzing activity selection as a sequential decision making process, Markov models have drawn some attention. For example, Liao et al. (2007) used hierarchical Markov models applied to GPS location data to predict users' activity selection. Their model predicts the location of the next stop, duration, and mode for the next trip. Inference is based on Rao–Blackwellized particle filtering method, where the posterior distribution of location depends on the previous GPS data. Leszczyc and Timmermans (2002) applied Markov transitions to estimate the choice and duration of the next activity as dependent on the duration and the type of the previous activity; they estimated activity durations by Hazard models. Goulias (1999), studied activity patterns using latent mixed Markov models applied to panel data; the trend of change in the daily and long-term activity patterns of individuals were analyzed by Markov transition probabilities.

In continuous Markov models the probability of any event at time '$t + s$' depends on the state of the chain at time 's' (Serfozo, 1979). As such, they are good candidates to model activity durations and the timing of their transitions to a following activity. However, because individuals likely don't share similar activity durations – even the time spent by the same person on the same activity is not constant – these models are not particularly useful in predicting the discrete changes in state (occurring at various times) that make up activity order. For this latter case, discrete Markov models are a more logical choice.

Despite being easy to estimate, Markov models have some pitfalls that create questions about their credibility for application. First, in Markov chains, it is assumed that the transition probability between different states is time-invariant. This assumption is certainly questionable for the activity selection/order problem; e.g., the transition probability from "physical exercise" to "work" activity in the morning likely is different from that probability in the afternoon. Second, the "memoryless" property of Markov chains, by which the choice in the next state depends only on the current state, probably is not compatible with the daily activity selection procedure, which typically takes place within a short period of time. Third, absent socio-demographic data, pure Markov models are descriptive models that describe the status and probability of transition between different states – they are not known as proper models for inference or for replicating observed patterns.

In this research, we use Support Vector Machine (SVM) techniques to estimate the sequence of activities in the daily agenda of individuals. By employing hidden Markov models and conditional random fields, we estimate the type of activities and their sequence in daily activity patterns of individuals. Unlike most previous applications that use Markov models to find the probability of the next state based only on the current state, we include socio-demographic characteristics of the individuals as exogenous variables in the model. It should be noted that in this study what we address are estimating activity types and their ordering – such parameters as activity location, duration, and mode are not considered.

Our prime contention is that replicating activity agendas of individuals based on their socio-demographic attributes is an important step in moving current activity based models closer to application. As an analogy, travel forecasts grounded in trip-based modeling methodologies (e.g., the traditional four-step modeling approach) typically begin with a "trip generation" step that, for want of any proven theoretical foundations, correlates the number (and, possibly, categories) of trips expected to be generated by any particular observation (region, household or individual) to socio-demographic data commonly available either from regional planning studies or census data. Subsequent to this phase, the remaining steps of modeling efforts (e.g., destination, mode and route choice) are the subjects of models that generally have theoretical causal relationships. By analogy, in activity-based modeling approaches, "step 0" requires the specification of the observation's activity agenda (the activity-based counterpart to the number of trips in trip-based models); subsequent to this step, the remaining steps in activity-based frameworks (either rule-based, computational, or utility-based, among others) are based on theoretical constructs that parallel (at least in spirit) the theories behind individual trip-making behavior. Thus far, the vast majority, if not all, of proposed activity-based models assume either that activity agendas are given, or can be extracted directly from activity diaries – as such, with few notable exceptions (e.g., CEMDAP, TRANSIMS) advances in such approaches have been limited to largely academic exercises; practical application to either replace or modify regional planning models with activity-based principles has remained a challenge.

In the approach presented here, we propose to "tease out" general tendencies of activity involvement and sequencing of those activities using Markov Chains with Conditional Random Fields-CRF (Lafferty et al., 2001). Fig. 1 represents the structure of the Markov Chain with Conditional Random Fields-CRF. CRFs are categorized as discriminative models and focus on estimating conditional distribution of $p(y|x)$ rather than joint distribution of $p(x, y)$. In this study we use directed graphs to represent the relation between X, explanatory variables, and y, activity patterns. Although general CRF models are undirected

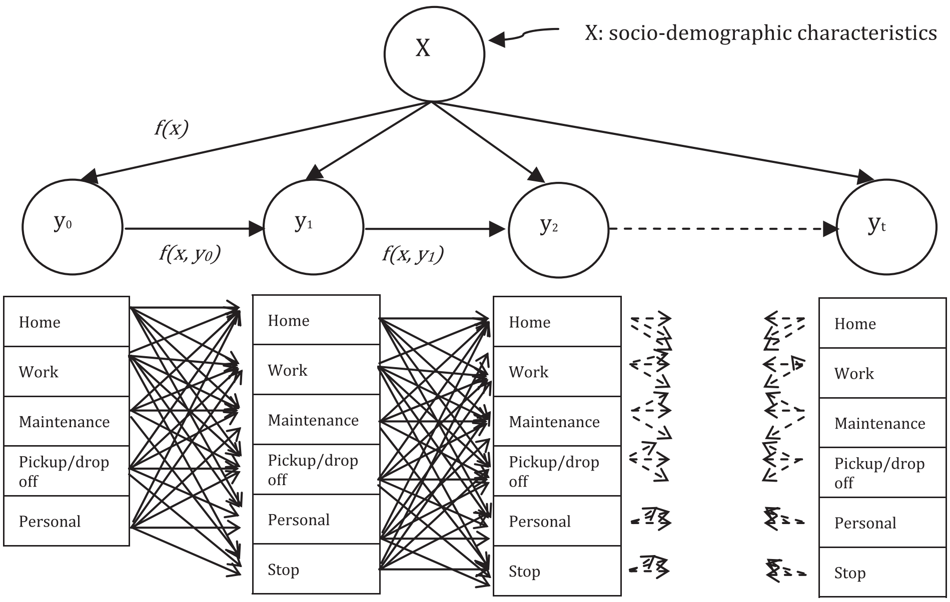

Fig. 1. Markov model structure for activity pattern sequence.

graphical models, because we are analyzing the dependency among predictions conditional on given explanatory variables, we introduce our model as hidden Markov models with conditional random fields. In this figure, 'X' represents socio-demographic data, and 'y' represents activity type at different time steps. 'y' is a multinomial response vector and the last activity of the individual is followed by 'stop'; as an example, if the individual has 5 activities in his/her agenda, then his/her choice for the states of '$t > 5$' would be 'stop'.

As noted above, one of the underlying assumptions for Markov models, which likely does not hold for daily activity selection, is that of time-invariant transition probabilities. We resolve this issue by using time-variant state transition variables and learning model parameters by using Support Vector Machines (SVM). Further, we evaluate the performance of different dependency structures among states. We develop two different model structures for SVM: structure 1: choice of activity in each state depends only on the activity type at the previous state and socio-demographic data; structure 2: choice of activity in each state depends on all previous states and socio-demographic data.

Socio-economic data used in this investigation include: household size, vehicle ownership, household income, age, gender, student status, and employment status. In selecting socio-economic variables to include in the model, the criteria were to keep number of variables small enough to avoid overfitting of the train data, and to use a set of distinctive variables that are commonly available in travel survey and census data. Statistical evidence reported in the paper verifies that a proper set of variables is included in the model.

We hypothesize that there is at least a weak correspondence among these characteristics and tendencies associated with the ordering (sequencing) of daily activities (which, of course, is heavily influenced by exogenous factors that are confoundedly hard to capture, much less predict). Our intent is to try to move activity-based modeling approaches closer toward their use in planning applications by casting some of the essential features that define activity/travel patterns in terms of commonly-available household-level statistical data. The hope is that the models provided here can be used in conjunction with other, more formal and more detailed, models of activity behavior to help refine the theoretical "behavioral" constructs that underlay the resultant formation and execution of the activity agendas in those models.

In Section 2 of the paper, we describe some preliminary concepts of SVM classifiers. In Section 3, the structure of multinomial logit models and support vector machines are provided. Section 4 presents the empirical results obtained from the model application, and Section 5 discusses some detailed analysis of the outcome followed by Section 6, which concludes the paper by presenting the findings of the research and its potential application.

2. Support Vector Machines (SVMs)

2.1. SVM with binary classes

Support vector machines are among the most widely applied supervised classification methods in the field of machine learning and artificial intelligence. The method is well known in computer science, and there have been some applications of the method in transportation – for example, Theja and Vanajakshi estimated short term traffic parameters including speed, headway and volume (Theja and Vanajakshi, 2010). However, application of this methodology in studying aspects of travel behavior has been extremely limited, and none of those applications use Markov chains as an underlying structure of sequence of activity selection. Yang et al. (2010) used support vector machines to study travel behavior of individuals using GPS data. Activity categories in their study are time-fixed work, flexible work, movie, eating out, shopping, KTV, exercise, and sightseeing. The spatial–temporal features used in their study include activity start time, end time and the distance between activity locations. In their study, learning is unsupervised – given the previous data they forecast the individual's next choice in terms of location, schedule and mode. Moons et al. (2007) used different estimation models to compare linear, semi-linear and nonlinear model performances relative to mode choice. They conclude that for skewed data, nonlinear models (SVM and CART) overfit the results; however, for balanced distributed data the nonlinear models perform well. Zhang and Xie (2008) used SVM to study mode choice behavior of travelers. A performance comparison of SVM and neural networks in Cheu et al.'s work indicates a higher accuracy is obtained from SVM for incident detection on San Francisco I-880 freeway data (Cheu et al., 2003).

As a brief introduction, a support vector machine with binary class is a supervised learning method that predicts the labels of points in a test dataset by learning the model for training dataset. Every data point i in the training set is represented by its feature vector, x_i, and label, l_i. Depending on which class C_i belongs to, its label takes the value of -1, or 1 as follows:

$$l_i = \begin{cases} -1, & \textit{if } x_i \in C_1 \\ 1, & \textit{if } x_i \in C_2 \end{cases}$$

In the case of a linearly separable training dataset as Fig. 2, using notation similar to (Bishop, 2006), we can draw any hyper-plane with weight vector of W and the intercept of b, $f(x) = W^Tx + b$, such that for any point x_i in test set, if $f(x_i) > 0$ then x_i will be classified in class 2, C_2, otherwise it belongs to class 1, C_1.

Among all possible hyper-planes, the optimum classifier – that with the lowest risk of misclassification – is the one that has the largest margins between the closest data points to the hyper-plane (Bishop, 2006). The perpendicular distance from any point to the hyper-plane is calculated as $|f(x_i)|/\|W\|$ (Bishop, 2006); then the objective function that maximizes the margins to the hyper-plane can be written as

$$\arg\max_{w,b}\left\{\frac{1}{\|w\|}\min_n[W^Tx_n + b]\right\}. \tag{1}$$

Eq. (1) is an over-parameterized model and needs to be restricted. By setting $W^Tx_n + b = 1$ and writing the objective function in quadratic form, the problem is converted to equation set (2).

$$\begin{aligned} &\min \quad \frac{1}{2}\cdot w^2 \\ &\text{s.t.}: \quad W^Tx_i - b \geqslant 1, \quad \forall l_i = 1, \\ &\qquad\quad W^Tx_i - b \leqslant -1, \quad \forall l_i = -1. \end{aligned} \tag{2}$$

Solution procedures to problem 2 mostly result in infeasible solutions, since in most cases perfect separation between classes is unattainable. However, the model can be relaxed by adding a penalty term $C\sum_i p_i$ to the objective function, where p_i is the distance of the point from the margin if it is classified in the wrong class, and C is a constant coefficient that may be varied to reflect the weight of the penalty. Rewriting the constraints of Eq. (2) as $l_i(W^Tx_i - b) \geqslant 1$, the model with penalty function can be written as

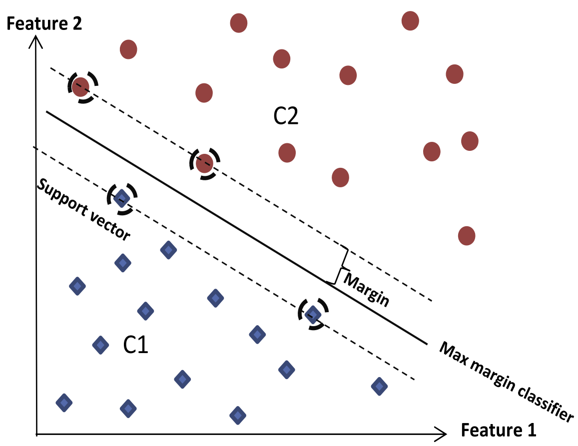

Fig. 2. Support vectors classifiers.

$$\begin{aligned} \min_{w,b,p} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} p_i \\ \text{s.t.}: \quad & l_i(w^T x_i - b) - 1 + p_i \geqslant 0 \quad \forall i, \\ & p_i \geqslant 0 \quad \forall i. \end{aligned} \tag{3}$$

The dual transformation of Eq. (3) can be written in the following form (interested readers can refer to the Appendix for more details on Eq. (4))

$$\begin{aligned} \max \quad & L_D = \sum_{i=1}^{N} \eta_i - \frac{1}{2}\sum_{ij} \eta_i \eta_j l_i l_j x_i^T x_j \\ \text{s.t.}: \quad & \sum_i \eta_i l_i = 0, \quad \forall i, \\ & 0 \leqslant \eta_i \leqslant C, \quad \forall i. \end{aligned} \tag{4}$$

In these equations η is the Lagrange multiplier. The dual function is kernelised, since we can write $x_i^T x_j \to k(x_i, x_j)$. Using the kernel function shifts the model to higher dimensional spaces, and non-separable data in lower spaces may be separable in higher dimensions. Fig. 3 illustrates how a non-separable data in 2 dimensional space can be separated if they are transformed to three dimension.

Having values of model parameters from equation set (4), the labels of data in the test set can be computed using Eq. (5) as

$$y_{test} = sign[w^T x_{test}] = sign\left[\sum_{n \in SV} \eta_n l_n K(x_{test}, x_n)\right]. \tag{5}$$

2.2. SVM with multiple classes

In this research, we have multinomial choices among different activity categories. Defining "A" as the set of activities, A = {work, maintenance, personal, pick-up/drop off}, at each state of activity sequence only one of the activities in "A" will take the value of 1. The choice of the next state is determined by using multiple class support vectors and every data point is assigned to the class with the largest margin.

Suppose the training set is given as $(x_1, l_1), (x_2, l_2), \ldots, (x_n, l_n)$, in which x is the feature set and l is the multinomial response model with K classes, $l_i \in \{1, 2, \ldots, K\}$. Using notation similar to Weston and Watkins (1998), the primal problem for SVM with multiple classes follows Eq. (6) given as:

$$\begin{aligned} \min_{w,b,p} \quad & \frac{1}{2}\sum_{m \in K} \|w_m\|^2 + C\sum_{i=1}^{N}\sum_{m \neq y_i}^{K} p_i^m \\ \text{s.t.}: \quad & w_{y_i}^T \cdot x_i + b_{l_i} - (w_m^T \cdot x_i + b_m) \geqslant 1 - p_i^m, \\ & p_i^m \geqslant 0, i = 1, \ldots, n; \; m \in \{1, \ldots, K\} \setminus l_i. \end{aligned} \tag{6}$$

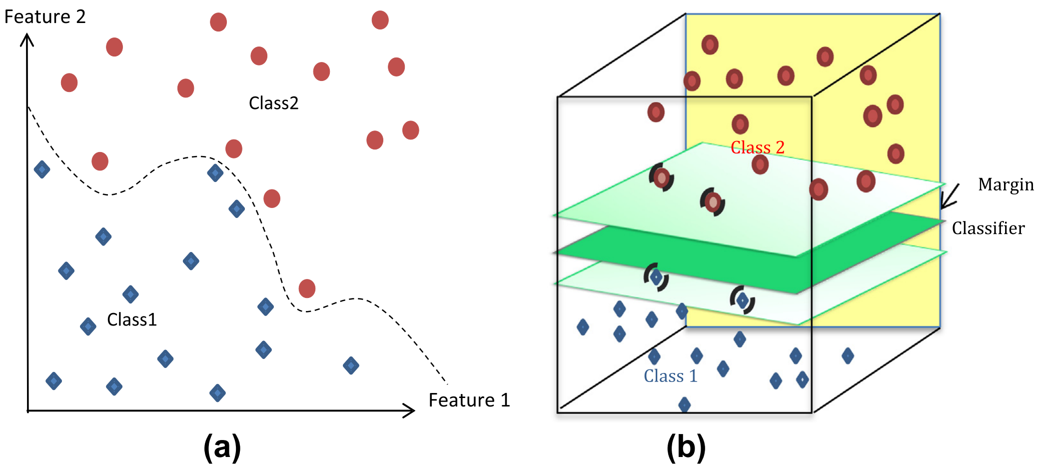

Fig. 3. Linearly inseparable data in lower feature (a) space can be separated in higher dimensional feature space (b).

Converting the problem to its dual function and solving the dual transformation of the model, gives the optimum solution to the problem in kernelized form as Eq. (7):

$$W^*(\eta) = 2\sum_{i,m}\eta_i^m + \sum_{i,j,m}\left[-\frac{1}{2}c_j^{l_i}A_iA_j + \eta_i^m\eta_i^{l_i} - \frac{1}{2}\eta_i^m\eta_i^m\right]x_ix_j, \tag{7}$$

$$A_i = \sum_{m=1}^{K}\eta_i^m, \quad c_i^m = \begin{cases} 1 & \text{if } l_i = m \\ 0 & \text{if } l_i \neq m \end{cases}.$$

The value of W^* obtained from Eq. (7) are the coefficients of the hyperplane that has the largest margin and separates different classes with the minimum misclassification error (Weston and Watkins, 1998). The decision rule for classifying test data in multiple classes is based on the following equation:

$$f(x) = \arg\max_m\left[\sum_{i=1}^{n}(c_i^m A_i - \eta_i^m)x_i.x_{test} + b_m\right]. \tag{8}$$

3. Model development

In this study, we use conditional random fields on Markov chains. Let T be the activity pattern of the individual, where $T = \{y_{t=0}, y_{t=1}, \ldots, y_{t=T}\}$. Then its corresponding probability for individual i can be written as

$$P_i(T) = P_i(y_{t=0}|x)\cdot P_i(y_{t=1}|y_{t=0},x)\cdot P_i(y_{t=2}|y_{t=1},x)\cdot\ldots\cdot P_i(y_{t=T}|y_{t=T-1},x). \tag{9}$$

In the following sections the details of the learning model parameters are presented.

3.1. Logistic regression with conditional random field

In this approach, activity selection at each state of the daily activity sequence is assumed to depend on socio-demographic features of individuals and only their most recent previous activity. Although in transportation applications logistic regression is normally used to estimate the probability of choosing one class with respect to other classes, in computer science applications it typically is used as a classifier, where each data point belongs to the class with the highest probability value. Logit models are based on the strong IIA (independent from irrelevant alternative) assumptions, for any two alternatives the ratio of their logit probability does not depend on the other alternatives, additionally in Logit models, off-diagonal terms of covariance are zero.

3.1.1. State $t = 0$

The probability of selecting a specific activity at state zero is computed based on a multinomial logit model. Assuming the distribution of error terms to be IID extreme value, for each individual the probability of choosing alternative j at $t = 0$ is calculated as:

$$P\left(y_{t=0}^{i,j} = 1|x\right) = \frac{\exp\left(\sum_p \theta_p^j x_p^i\right)}{\sum_l \exp\left(\sum_p \theta_p^l x_p^i\right)}, \tag{10}$$

where i is the index representing individual, j the index representing activity type, p the index representing explanatory variables, θ the coefficient of socio-demographic variables for first activity in the agenda, and x is the socio-demographic variable.

3.1.2. State $t > 0$

Selection of the activity among all other activities at later states depends on socio-demographic data, and the activity type at the previous state of the model, and is computed as:

$$y_{t>0}^{i,j} = \sum_p \gamma_{t,p}^j x_p^i + \sum_k \beta_{t-1=k}^j P_{t-1}^{i,k} + \varepsilon_t^j \tag{11}$$

where k is the index representing activity type at previous state, t the index representing time sequence, γ the time-variant coefficient of socio-demographic variables, and β is the time-invariant coefficient representing the dependency between activity type at previous state and the current state.

Eq. (11) has a recursive format and the latent variable at state $t, y_t^{i,j}$, is a function of probability of different activities at previous step, $P_{t-1}^{j,k}$. The probability of participation in activity j at time step $t > 0$ given previous activity type and socio-demographic data is computed as

$$P_i\left(y_{t>0}^{i,j}=1|y_{t-1}^{i,j},x\right)=\frac{\exp\left(\sum_p \gamma_{t,p}^j x_p^i+\sum_k \beta_k^j P_{t-1}^{i,k}\right)}{\sum_l \exp\left(\sum_i \gamma_{t,i}^l x_i+\sum_k \beta_k^l P_{t-1}^{i,k}\right)}. \tag{12}$$

3.1.3. Likelihood function

The corresponding likelihood function is

$$L=\prod_i\prod_j\left(\frac{\exp\left(\sum_p \theta_p^j x_p^i\right)}{\sum_l \exp\left(\sum_p \theta_p^l x_p^i\right)}\right)^{\delta\left(y_{t=0}^{i,j}=1\right)}\left(\prod_{t>0}\left\{\frac{\exp\left(\sum_p \gamma_{t,p}^j x_p^i+\sum_k \beta_k^j P_{t-1}^{i,k}\right)}{\sum_l \exp\left(\sum_i \gamma_{t,i}^l x_i+\sum_k \beta_k^l P_{t-1}^{i,k}\right)}\right\}^{\delta\left(y_{t>0}^{i,j}=1\right)}\right), \tag{13}$$

where δ is a binary variable takes the value of 1 if the function inside the parentheses holds.

The likelihood function in Eq. (13) is a non-convex problem; in order to fit the model parameters we use a differential evolution algorithm (Storn and Price, 1997). Differential evolution is similar to Genetic Algorithm, which explores the feasible region and generates new solutions by crossover and mutation and replaces the old solutions if new ones yield a better objective function. The differential evolution algorithm has been proven the most efficient evolutionary algorithm (Storn and Price, 1997). Stacking all unknown parameters as $\lambda=(\theta_1^1,\ldots,\theta_p^J,\gamma_{1,1}^1,\ldots,\gamma_{t,p}^J,\beta_1^1,\ldots,\beta_K^J)$, the steps in the algorithm are as follows:

I. Initialization:
  a. Generate initial set of solutions, S, by the Latin hypercube sampling (McKay et al., 1979) method; the number of rows of S equals to the size of λ and has 'n' columns, where 'n' is initial population size. The recommended value for n is 5–10 times the dimension (Storn and Price, 1997).
  b. Select random numbers $0\leqslant P_C\leqslant 1$ and $0\leqslant P_m\leqslant 1$, representing the acceptance threshold for crossover and mutation operations, respectively. In this study, these parameters are arbitrarily set to 0.8 and 0.1.
  c. Set K and F to some predetermined values such that $0\leqslant F\leqslant 1,\ 0\leqslant K\leqslant 1$; for the results reported in this study $K = 0.8$ and $F = 0.2$.

II. Crossover and mutation:
  a. Crossover each column in S, S_i, by three other columns (S_a, S_b, S_c), randomly selected from S, $NewColumn = S_i + K(S_i - S_a) + F(S_b - S_c)$.
  b. For every cell in *NewColumn*, generate a random number in the range of $[0,1]$ as u. If $u > P_C$, replace the value of cell in *NewColumn* with its corresponding value at S_i.
  c. For every cell of *NewColumn*, generate a random number r, if $r\geqslant P_m$, for that point of *NewColumn*, generate a new point using Latin hypercube sampling method.
  d. Evaluate the value of likelihood for *NewColumn*; if the value of likelihood is greater than the likelihood of S_i, replace S_i with *NewColumn*, otherwise keep S_i.

III. Check the convergence criteria:

If the difference between previous solution and the current solution is less than some predefined threshold accept the solution; otherwise go to the step 2.

3.2. Support vector machines with K classes (K-SVM)

A key assumption of Markov models is time-invariant transition probabilities between different activity dependencies at different states of the sequence. Although this assumption simplifies the problem, it has been criticized as an unrealistic representation for daily activity sequencing of individuals. Using K-SVM – SVM with multiple classes – transition probabilities between different activity categories do not need to be the same among different time states during the day. Additionally, in order to analyze the impacts of previous activities in the sequence on the current activity type, we train SVM with two different dependency assumptions as described in the following as Case I and Case II.

- Case I: represents the case where the activity at state 't' depends on socio-demographic characteristics and the activity choice at state '$t-1$' (for simplicity, in the remainder of the paper we refer to this case as K-SVM_St.1). Fig. 4 illustrates the feature set at different stages of K-SVM.
- Case II: as represented in Fig. 5, in this case it is assumed that activity at state t depends on socio-demographic characteristics and all previous activities of the individual in the same day (for simplicity, in the remainder of the paper we refer to this case as K-SVM_St.2).

As noted above, Eq. (7), which is the solution to the SVM with multiple classes, is a kernelized function. Based on a preliminary analysis on different kernel functions for the given data that revealed that the Gaussian kernel function had a higher

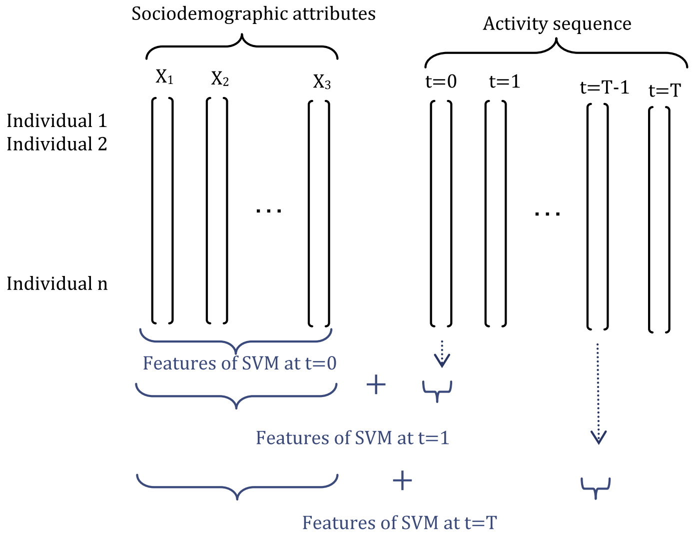

Fig. 4. Feature set representation of K-SVM_St.1.

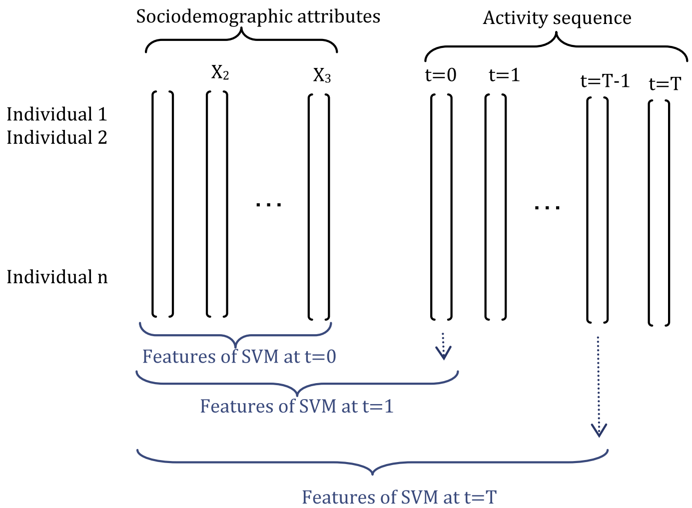

Fig. 5. Feature set representation of K-SVM_St.2.

Table 1
Data description used for the activity categories.

In home	Working at home, eating/preparing food, watching TV, shop by phone/online, exercise at home, other
Work	Work activities, regularly scheduled volunteer work, work related (calls, meetings, etc.)
Maintenance	Child care, buying gas, incidental shopping (grocery, house wares, medicine), major shopping (furniture, clothes, autos, etc.), post office, utilities, other household business, be with other household members at their activities
Personal activities	Fitness activities, voting, medical, ATM, banking. recreational (vacation, camping, etc.), entertainment (movie, dance club, bar, etc.), visit friends and relatives, community meeting, political or civic events, volunteer work, church, temple, religious meetings, out of home meals
Pick-up/drop-off	Pick someone up or get picked up, drop someone off or get dropped off

accuracy than the polynomial version, we used Gaussian kernel function in our study, $k(x_i, x_j) = \exp\left(\frac{-\|x_i - x_j\|^2}{2\sigma^2}\right), \sigma = 1$, to separate classes in higher dimensions.

Table 2
Data on observed activity sequences for 3671 individuals.

		Activity at state $t + 1$				
		In-home	Work	Maintenance	Personal	Pickup/drop off
Activity at state t	In-home	98	1081	763	247	464
	Work	928	149	258	45	104
	Maintenance	872	163	438	42	58
	Personal	196	26	98	20	15
	Pickup/drop off	337	122	112	28	105

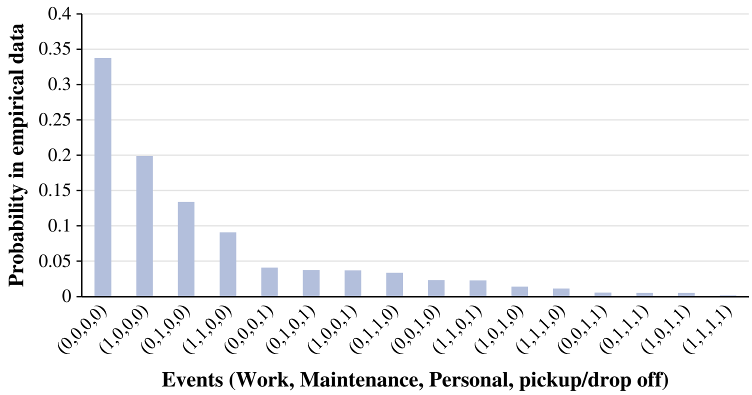

Fig. 6. Distribution of different combinations of agenda.

4. Empirical results

We applied the models to data associated with 3671 residents of Orange and San Diego Counties in California drawn from the 2001 California Personal Travel Survey (2002). In this study, eight socio-demographic characteristics of individuals, including income level, age, gender, household size, student status, number of vehicles in the household, employment status, and education level of individuals enter into the analysis; activity categories are "in-home", "work", "maintenance", "personal", and "pickup/drop off". Table 1 represents the details of activity categories used in this study. Trips of children younger than 6 years and incomplete data were excluded from dataset.

The successive transitions between these activity types for the sample are shown in Table 2.

Fig. 6 represents the distribution of different possible combination of agenda in the dataset. According to this data, about 20% of the individuals only participate in work activity, the highest share among out of home activities. In this figure, the events represent different possibilities of participation in activity type, 1 indicates participation and zero otherwise.

The distribution of statistical parameters of the dataset is shown in Fig. 7. Fig. 8 illustrates the distribution of the number of daily activities for individuals in our sample. The average number of activities for individuals is 2.2 activities per day. As is shown in the graph, less than 2% of individuals participate in more than 9 out-of-home activities during a day. Based on this, we assume 9 time states for daily activity engagements in the estimations. In our dataset, 51% of individuals are female and number of employed individuals is 2027.

In the estimation, 85% of data is used for training and the remaining 15% is used for testing model performance. In the following section, the results associated with both classification techniques – Multinomial Logit classifier and Support Vector Machines – are presented.

4.1. Multinomial logistic regression classifier

The coefficients of parameters for different activity purposes are illustrated in Figs. 9–13; the supportive statistical evidence representing the lower bound and upper bound for 95% of confidence interval is appended to the paper. In these figures, the *x*-axis represents different states of the Markov chain; these states indicate the daily stops of individuals. Based on the results obtained from MNL with conditional random fields, for the "In-Home" category all socio-demographic variables have positive coefficients at the first state – a result compatible with the first activity of the day *de facto* the "in-home" activity. Age has positive coefficients in early states of "maintenance," and "personal" activities; also its coefficient for "work" activity is positive for middle states of the activity sequence.

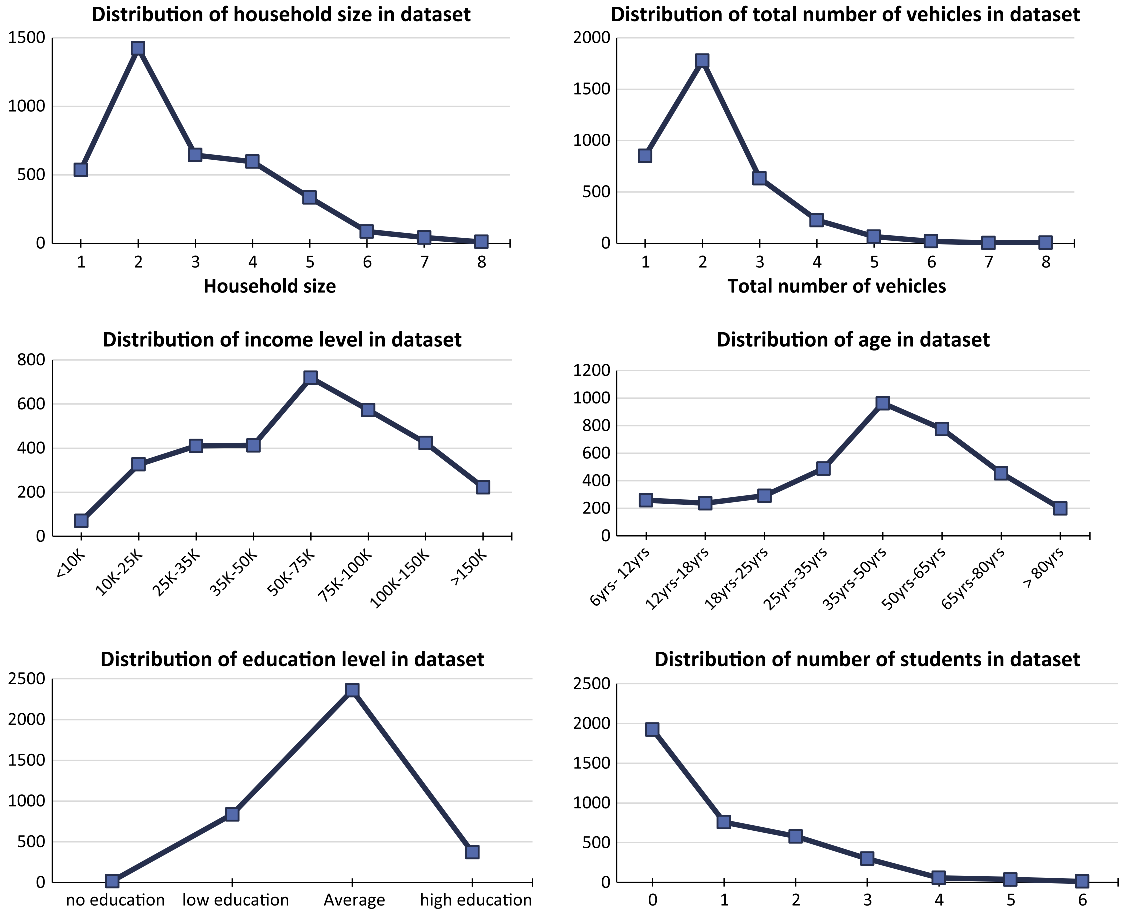

Fig. 7. Distribution of socio-demographic data.

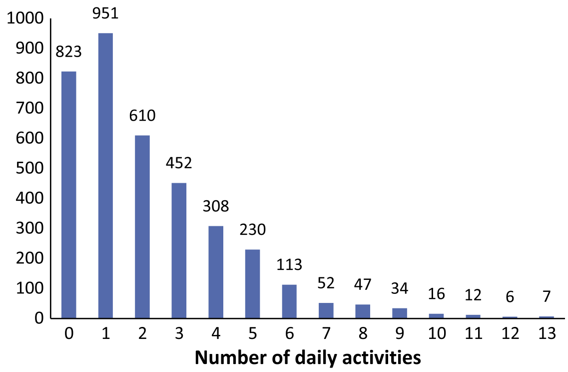

Fig. 8. Distribution of number of daily activities for individuals.

Females tend to participate in "pickup/drop off" activities in three states (2nd, 5th and 7th state) of the day, compatible with household roles in taking care of children. "Maintenance" activities appear in the middle states of females' agenda while a "work" activity is estimated to be in the very beginning and later states of their agenda.

Individuals with higher incomes tend to engage in "work" and "personal" activities in early states of the activity sequence; in the later states they tend to engage in "in-home" activities. Students tend to start their day with "maintenance" or "personal" activities and "pickup/drop-off" activities appear mostly in the second state of their agenda. However, the coefficient for student's activity decreases in later states for all activity types.

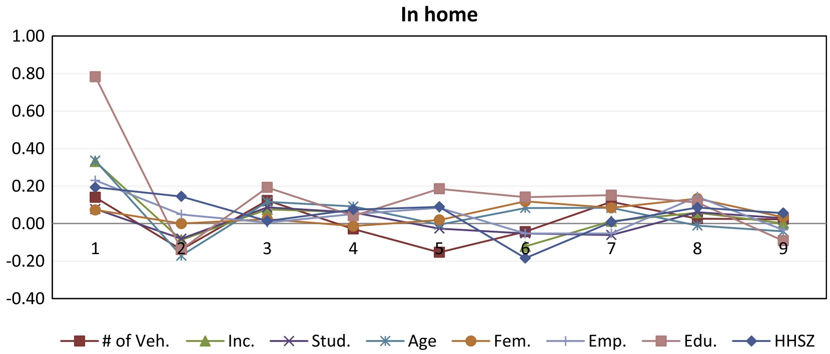

Fig. 9. Coefficients of socio-demographic variables at different states of in-home activity (obtained from MNL with CRF).

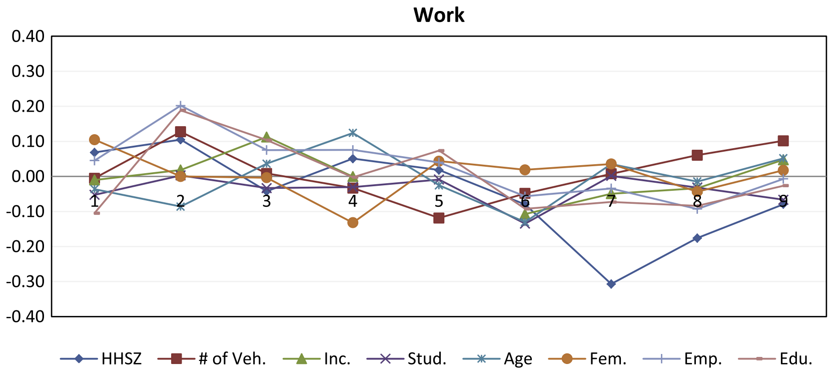

Fig. 10. Coefficients of socio-demographic variables at different states of work activity (obtained from MNL with CRF).

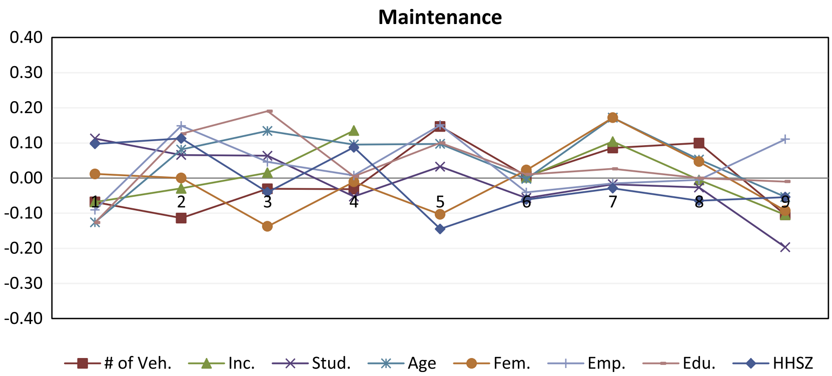

Fig. 11. Coefficients of socio-demographic variables at different states of maintenance activity (obtained from MNL with CRF).

Individuals with larger household size tend to attend "work" and "maintenance" activities in early states of the day and have "in-home" and "personal" activities in the later states of their agenda. As expected, "pickup/drop-off" activities appear in several stages of the agenda for individuals with larger household size.

In addition to the coefficients of the different socio-demographic variables, dependency coefficients between different activity categories were obtained from the MNL model. As shown in Table 3, except for the "work" activity – where the coefficient is almost zero – the coefficient for repeating two similar activities right after each other is negative; this is consistent with the concept of need-based models where individuals attend activities to fulfill their needs and when the need is met the probability to participate in the same activity during the same day decreases.

Based on the results shown in Table 3, individuals tend to leave home for "work." Also, "pickup/drop off" activities have a higher chance to be accomplished before "work" and "personal" activities. The tendency for participation in a "maintenance" activity is higher if the person has already participated in a "work" activity.

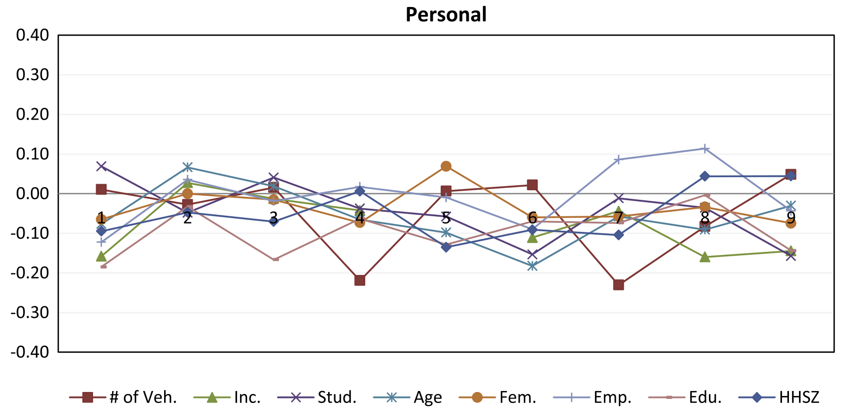

Fig. 12. Coefficients of socio-demographic variables at different states of personal activity (obtained from MNL with CRF).

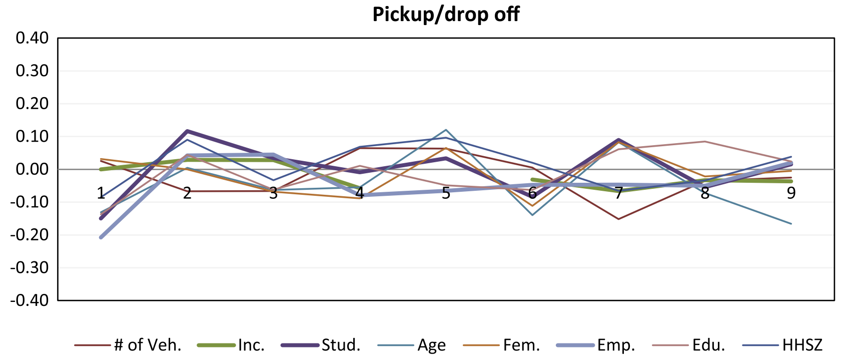

Fig. 13. Coefficients of socio-demographic variables at different states of pickup/delivery activity (obtained from MNL with CRF).

Table 3
Coefficients of inter-activity dependency obtained from MNL with conditional random fields.

		Activity at State '$t+1$'									
		In-home		Work		Maintenance		Personal		Pickup/drop off	
		Mean	%95 CI	Mean	%95 CI	Mean	%95 CI	Mean	%95 CI	Mean	%95 CI
Activity at state 't'	In-home	−0.109	[−0.12, −0.10]	0.128	[0.06, 0.196]	−0.088	[−0.10, −0.08]	−0.294	[−0.32, −0.27]	−0.175	[−0.18, −0.16]
	Work	−0.128	[−0.16, −0.10]	0.005	[−0.06, 0.06]	0.215	[0.09, 0.34]	−0.087	[−0.12, −0.05]	−0.067	[−0.070, −0.064]
	Maintenance	−0.001	[−0.04, 0.04]	−0.147	[−0.24, −0.05]	−0.046	[−0.92, 0]	−0.305	[−1.02, 0.42]	−0.147	[−0.34, 0.05]
	Personal	0.160	[0.14, 0.17]	−0.132	[−2.78, 2.50]	−0.078	[−2.01, 1.85]	−0.059	[−24.24, 24.12]	0.191	[−11.75, 12.12]
	Pickup/drop off	0.000	[−0.26, 0.26]	0.208	[−0.34, 0.76]	−0.197	[−0.56, 0.16]	0.010	[−3.80, 3.81]	−0.007	[−1.88, 1.86]

Table 4
Precision (%) of multinomial logit classifier for dataset.

State	1	2	3	4	5	6	7	8	9
Test dataset	99	34	38	44	38	66	65	74	73
Training dataset	98	36	68	47	71	71	81	87	89

Table 4 demonstrates the accuracy of the estimation with the MNL classifier for both the training and test sets. As noted, the pattern of accuracy for test and training sets is similar and the lowest accuracy rate belongs to the second state, whereas the highest is obtained for the first state which, for the majority of individuals is observed to be *de facto* "in-home" activity. A similar argument can be made for later states of the day in terms of returning home trips. Alternatively, the variation between activity choices in the middle states of the activity sequence is the highest, resulting in lower estimation accuracy.

4.2. Support vector machines with multiple classes

For the multiple-class case we used the SVM training package provided by Alain Rakotomamonjy (2003). As noted above, we trained SVM with two different dependency assumptions, K-SVM_St.1 and K-SVM_St.2, as shown in Figs. 4 and 5. K-SVM_St.1 has nine parameters comprising eight socio-demographic variables and one variable indicating the most recent activity, and for K-SVM_St.2, depending on activity state, the number of variables varies between 8 and 17, where the first activity of the day has eight variables (socio-demographic variables), second activity has 9 variables (8 socio-demographic variables, and one previous activity type), third activity has 10 variables (8 socio-demographic variables, and 2 previous activity types), etc. Using the kernel function in the dual formulation of SVM increases dimension of the problem, which results in computational complexity. In order to train the model within reasonable estimation duration, we randomly selected 10 subsets of data with a size of 2200 individuals and trained K-SVM on those subsets – obviously, including the entire data for simulation will increase prediction accuracy. The test set with the highest accuracy was chosen as the best model. Because the evaluations revealed that higher prediction accuracy is achieved for standardized data, the data were converted to standardized data with the mean of zero and variance of 1.

Fig. 14 compares the accuracy of the three methods. Comparing results of MNL and K-SVM_St.1, and K-SVM_St.2, we observe that K-SVM in both cases outperforms MNL model and in some cases the gap between accuracy estimation of MNL and k-SVM is 26% for test set, which belongs to estimation at state 5. Additionally we observe that using a more complex dependency structure in K-SVM_St.2 improves the accuracy of the model for the test dataset within 5% range comparing to K-SVM_St1.

In order to experiment with the performance of K-SVM_St.2 on different combinations of the dataset, we narrowed our sample to include only employed individuals and applied the model to "workers' activity patterns" in the sample. In addition, explanatory variables were reduced to: income level, age, gender, and household size.

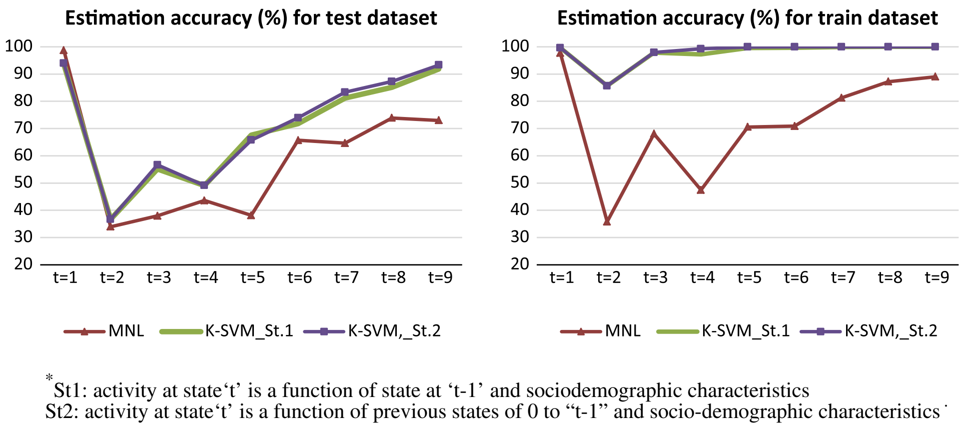

Fig. 14. Comparing the estimation accuracy of different methods for test dataset and training dataset.

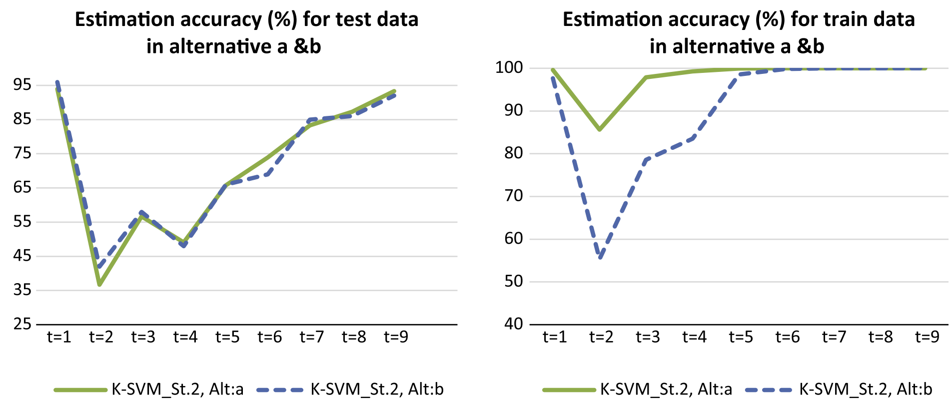

Fig. 15. Accuracy of K-SVM_St.1 for two alternatives: "a: all population" and "b: workers" activity pattern.

Table 5
Confusion matrixes for K-SVM_St1, K-SVM_St2 and MNL models.

	H[a]	W	M	P	PK	S
(a) K-SVM_St1 on test set						
H	*0.167*	*0.019*	*0.029*	*0.006*	*0.015*	*0.024*
W	*0.013*	*0.020*	*0.008*	*0.004*	*0.007*	*0.008*
M	*0.022*	*0.008*	*0.014*	*0.003*	*0.005*	*0.012*
P	*0.002*	*0.002*	*0.002*	*0.002*	*0.001*	*0.004*
PK	*0.006*	*0.003*	*0.003*	*0.001*	*0.005*	*0.005*
S	*0.023*	*0.012*	*0.015*	*0.005*	*0.008*	*0.521*
(b) K-SVM_St2 on test set						
H	*0.172*	*0.017*	*0.026*	*0.004*	*0.012*	*0.028*
W	*0.014*	*0.020*	*0.007*	*0.003*	*0.007*	*0.009*
M	*0.026*	*0.007*	*0.013*	*0.002*	*0.005*	*0.011*
P	*0.004*	*0.003*	*0.002*	*0.002*	*0.001*	*0.003*
PK	*0.007*	*0.003*	*0.003*	*0.001*	*0.006*	*0.004*
S	*0.025*	*0.012*	*0.015*	*0.002*	*0.006*	*0.522*
(c) MNL on test set						
H	0.163	0.012	0.002	0.000	0.000	0.093
W	0.015	0.036	0.000	0.000	0.000	0.029
M	0.025	0.014	0.001	0.000	0.000	0.037
P	0.003	0.006	0.001	0.000	0.000	0.006
PK	0.005	0.010	0.001	0.000	0.000	0.018
S	0.076	0.024	0.004	0.000	0.000	0.418
(d) K-SVM_St1 on train set						
H	*0.241*	*0.004*	*0.005*	*0.001*	*0.002*	*0.005*
W	*0.003*	*0.048*	*0.001*	*0.001*	*0.001*	*0.002*
M	*0.005*	*0.003*	*0.054*	*0.001*	*0.001*	*0.003*
P	*0.001*	*0.001*	*0.000*	*0.008*	*0.000*	*0.001*
PK	*0.001*	*0.001*	*0.001*	*0.000*	*0.025*	*0.001*
S	*0.004*	*0.004*	*0.003*	*0.001*	*0.002*	*0.565*
(e) K-SVM_St2 on train set						
H	*0.242*	*0.003*	*0.004*	*0.001*	*0.002*	*0.005*
W	*0.003*	*0.049*	*0.001*	*0.001*	*0.001*	*0.002*
M	*0.005*	*0.003*	*0.055*	*0.001*	*0.001*	*0.003*
P	*0.001*	*0.001*	*0.000*	*0.008*	*0.000*	*0.001*
PK	*0.001*	*0.001*	*0.000*	*0.000*	*0.025*	*0.001*
S	*0.004*	*0.004*	*0.003*	*0.001*	*0.001*	*0.566*
(f) MNL on train set						
H	0.201	0.011	0.003	0.000	0.000	0.071
W	0.021	0.029	0.000	0.000	0.000	0.016
M	0.037	0.009	0.001	0.000	0.000	0.029
P	0.005	0.003	0.000	0.000	0.000	0.006
PK	0.011	0.010	0.000	0.000	0.000	0.009
S	0.027	0.016	0.002	0.000	0.000	0.480
(g) K-SVM_St1 on entire population						
H	*0.161*	*0.007*	*0.012*	*0.001*	*0.005*	*0.070*
W	*0.013*	*0.003*	*0.002*	*0.000*	*0.001*	*0.036*
M	*0.020*	*0.003*	*0.005*	*0.001*	*0.001*	*0.035*
P	*0.003*	*0.001*	*0.000*	*0.000*	*0.000*	*0.007*
PK	*0.006*	*0.001*	*0.001*	*0.000*	*0.001*	*0.019*
S	*0.078*	*0.007*	*0.037*	*0.001*	*0.005*	*0.453*
(h) K-SVM_St2 on entire population						
H	*0.212*	*0.009*	*0.014*	*0.002*	*0.006*	*0.013*
W	*0.007*	*0.037*	*0.004*	*0.001*	*0.002*	*0.004*
M	*0.013*	*0.004*	*0.039*	*0.001*	*0.002*	*0.006*
P	*0.002*	*0.001*	*0.001*	*0.005*	*0.000*	*0.002*
PK	*0.005*	*0.002*	*0.002*	*0.001*	*0.017*	*0.002*
S	*0.013*	*0.007*	*0.008*	*0.002*	*0.004*	*0.549*
(i) MNL on entire population						
H	0.196	0.011	0.003	0.000	0.000	0.093
W	0.020	0.030	0.000	0.000	0.000	0.018
M	0.035	0.010	0.001	0.000	0.000	0.031
P	0.004	0.003	0.000	0.000	0.000	0.006
PK	0.010	0.010	0.000	0.000	0.000	0.011
S	0.035	0.017	0.002	0.000	0.000	0.451

[a] H: In home, W: Work, M: maintenance, P: Personal, PK: pick up/drop off, S: Stop.

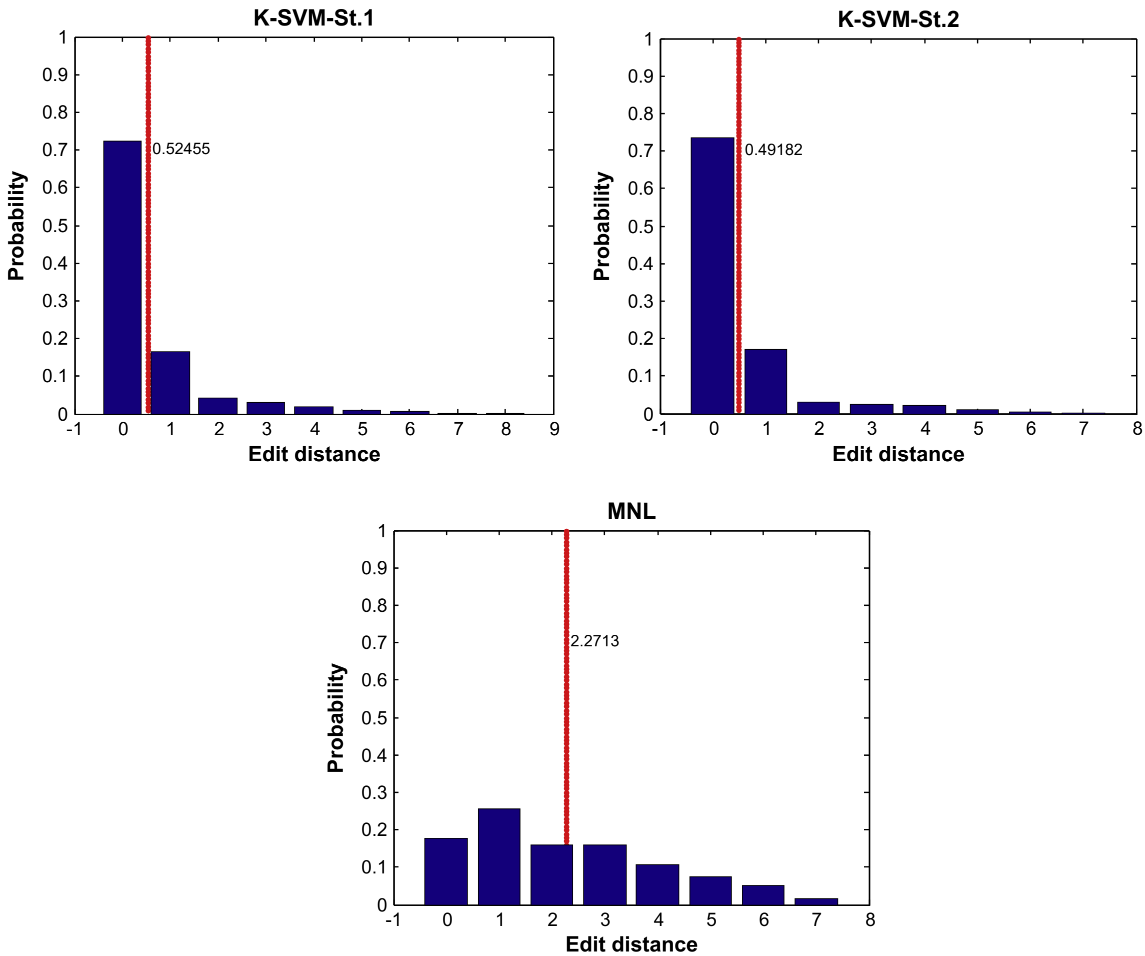

Fig. 16. Comparing the distribution of edit distance between observed and estimated activity patterns among different models.

Fig. 15 illustrates the accuracy of estimation by K-SVM_St.2 for two alternatives. In these figures, alternative "a" represents the dataset for the whole population with eight explanatory variables (income level, age, gender, household size, student status, number of vehicles, employment status, and education level) and alternative "b" represents "workers" activity pattern with only four explanatory variables (income level, age, gender, household size). According to the results provided in Fig. 15, the accuracy of K-SVM for the test data set is higher when we use clustered data with fewer explanatory variables, whereas the opposite holds for the training set – adding more explanatory variables results is overfitting.

5. Discussion

In addition to the results reported in Section 4, we here present a detailed analysis of the performance of proposed models.

5.1. Activity types and position in agenda

As was noted, replicating activity types in the agenda is an important outcome of the current study. Table 5 demonstrates the confusion matrix for each K-SVM structure and for the multinomial logit model. These matrices identify whether or not the respective model can successfully replicate both the presence of each activity type in the dataset; an indication of the estimation accuracy of the model is given by the sum of diagonal cells of confusion matrices. As an example, suppose the 3rd activity of person *i* is 'work activity', then in computing the confusion matrix, we compare only the 3rd observed activity of the same person with his 3rd activity in the estimation – regardless of whether or not his/her previous or succeeding activity was replicated correctly. According to the results presented in Table 5, estimation accuracy based on this criterion

Table 6
Sensitivity analysis of percentage of change in activity sequencing with change in socio-demographic characteristics.

Explanatory variable Variable	State	Age +10 Years	Income +$10,000/year	Household size +1 Person	Number of vehicles +1 Vehicle
In home	1	−0.68	−0.44	0.57	−3.76
	2	−6.43	−1.01	−4.36	−0.84
	3	−7.44	−1.85	−4.88	1.53
	4	1.47	2.81	−2.40	−0.49
	5	−1.44	−3.46	−0.27	1.14
	6	−2.21	0.05	−4.74	−1.01
	7	−0.41	2.97	−3.57	−0.84
	8	−0.14	−0.98	1.14	0.65
	9	−0.57	0.11	0.30	0.16
Work	1	0.87	1.04	0.60	1.85
	2	−4.39	−4.20	−4.36	−3.30
	3	3.13	1.61	−0.98	0.33
	4	0.74	2.04	0.60	5.15
	5	1.34	1.58	−1.72	0.22
	6	−0.90	0.30	0.33	−0.19
	7	0.93	−0.95	2.02	1.66
	8	−0.05	−0.30	0.27	−0.14
	9	0.38	−0.11	0.11	0.33
Maintenance	1	−0.16	−0.54	−0.87	1.74
	2	0.16	−0.68	3.68	2.86
	3	6.70	1.23	2.97	−0.49
	4	1.34	−0.65	−1.23	−3.19
	5	1.39	4.50	0.25	2.37
	6	4.17	1.36	2.10	0.68
	7	0.38	−0.03	0.95	−0.79
	8	1.14	0.38	−0.27	−0.54
	9	−0.44	0.05	0.05	−0.11
Personal	1	−0.05	0.00	−0.08	0.05
	2	3.43	0.35	−0.65	2.15
	3	−0.60	0.79	−0.79	0.65
	4	−1.39	0.41	2.21	−0.30
	5	1.01	0.68	0.87	0.14
	6	0.63	−0.30	0.54	0.33
	7	0.03	−0.27	0.05	0.16
	8	−0.03	−0.03	−0.05	−0.05
	9	0.00	0.00	0.00	0.11
Pick up/drop off	1	0.03	−0.05	−0.22	0.11
	2	4.74	0.74	6.62	0.30
	3	−2.51	0.46	0.79	−0.68
	4	−0.63	−1.14	−0.76	1.83
	5	2.83	−0.03	4.22	−0.49
	6	2.10	0.35	2.02	−0.79
	7	0.38	0.19	1.36	−0.16
	8	−0.46	0.46	−0.16	−0.16
	9	0.00	−0.11	−0.05	−0.16
Stop	1	0.00	0.00	0.00	0.00
	2	2.48	4.80	−0.93	−1.17
	3	0.71	−2.23	2.89	−1.34
	4	−1.53	−3.46	1.58	−3.00
	5	−5.12	−3.27	−3.35	−3.38
	6	−3.79	−1.77	−0.25	0.98
	7	−1.31	−1.91	−0.82	−0.03
	8	−0.46	0.46	−0.93	0.25
	9	0.63	0.05	−0.41	−0.33

of K-SVM_St2 for the test set is 73.5%, for the training set is 94.5% and for the entire population is 85.8%; the corresponding values for K-SVM_St1 are 72.8%, 94.2% and 62.3%; and for MNL are 61.8%, 71.2% and 67.8%. Although K-SVM_St2 has a relatively higher accuracy for standardized data, it comes with a larger computation cost, which is due to higher-dimensional feature sets as we move to later states of the activity sequence.

In terms of replicating the content of the observed activity agendas, regardless of activity order or sequence, the MNL model replicates 45% of the individual's agenda correctly, whereas the accuracies of K-SVM_St1 and K-SVM_St2 equal to 80% and 84% for the entire population.

Table 7
Comparing the estimation precision (%) of SVM and MNL Classifiers.

	State 1	State 2	State 3	State 4	State 5	State 6	State 7	State 8	State 9
(a) Test set									
MNL	*98.73*	*33.94*	*37.93*	*43.56*	*38.11*	*65.70*	*64.61*	*73.87*	*73.00*
K-SVM_St.1	*93.64*	*36.67*	*55.15*	*49.09*	*67.58*	*71.82*	*81.21*	*85.15*	*91.82*
K-SVM,_St.2	*93.94*	*36.67*	*56.67*	*49.09*	*65.76*	*73.94*	*83.33*	*87.27*	*93.33*
(b) Train set									
MNL	*97.69*	*35.77*	*68.04*	*47.44*	*70.58*	*70.87*	*81.28*	*87.24*	*89.00*
K-SVM_St.1	*99.68*	*85.72*	*97.86*	*97.27*	*99.52*	*99.63*	*99.95*	*100.00*	*100.00*
K-SVM,_St.2	*99.63*	*85.67*	*97.91*	*99.25*	*99.95*	*100.00*	*100.00*	*100.00*	*100.00*

5.2. Activity sequencing

In this section, the order of activities in the observed data and the outcomes of the proposed models are compared using sequential alignment techniques (Needleman and Wunsch, 1970). We measure distance between activity patterns based on the number of steps required to align two sequences of activities (independent of activity duration or start times), where smaller distance between strings indicates higher similarity. The distribution of distance between observed and replicated patterns using MNL model, K-SVM-St.1, and K-SVM-St.2 are illustrated in Fig. 16. The red[2] line indicates the average distance between the observed and replicated patterns in entire dataset. According to the results, the minimum average distance is achieved for K-SVM-St.2, which is 0.49, and it is followed by K-SVM-St.1, 0.52. The average distance for MNL model equals to 2.27, which is substantially larger than are the distances for K-SVM models. Additionally, both K-SVM structures successfully estimate the activity order of more than 70% of the entire population, whereas the accuracy of MNL in replicating activity sequencing falls far below K-SVM models.

5.3. Sensitivity analysis

In order to evaluate the impacts of explanatory variables on SVM estimations, each variable is perturbed, while the remaining variables are fixed, and its impact on the results of SVM is tested. Table 6 summarizes the percentage of change in activity participation at different states of the day for every unit of change in each explanatory variable.

6. Conclusions

In this paper, activity sequencing of individuals was studied using Conditional Random Fields on Markov Models. Apart from adapting new techniques in modeling daily activity sequences, we applied three modifications to the existing studies in the field: first, we employed socio-demographic characteristics as exogenous variables in hidden Markov models, converting these models from descriptive to inference models. Second, we used time-variant transition probabilities between different states of the activity sequences, more compatible with an activity selection procedure within a short period of planning. Third, we tested different dependency structures between future states and previous states of activity sequence to investigate the effect of model structure on performance. Two classification techniques were used: sequential multinomial logistic regression model (MNL) and sequential support vector machines for multiple classes (K-SVM).

We fitted the parameters of the sequential MNL model using a differential evolution algorithm and analyzed the impacts of socio-demographic variables (income level, age, gender, household size, student status, number of vehicles, employment status, and education level) and previous activity types on individuals' activity sequences.

The K-SVM was trained with two different assumptions regarding the dependency structure of the model. In the first structure, K-SVM_St.1, activity type at time 't' depends on most recent activity, at state '$t-1$', and socio-demographic explanatory variables; in the second structure, K-SVM_St.2, activity type at time 't' depends on all previous activities "states 0 to $t-1$" and socio-demographic characteristics.

Based on the results, while all three models have relatively low precision for the second state of activity sequence in the test dataset – a result attributed principally to the large variation in the type of activities during the early stages of sequences in the observed data, as well as the stated expected weak linkage of these outcomes to merely socio-demographic variables – the models show relatively good predictive power for the ordering of activities performed later in more complex strings of activities. And, as shown in Table 7, K-SVM models generally have higher accuracy than MNL for both the training and test sets, particularly in cases involving complex patterns.

In order to further evaluate K-SVM_St.2's performance, we applied the model to a filtered sample that included only the trips of workers, and with a smaller set of socio-demographic variables. For such filtered data, K-SVM_St.2 has higher precision for the test set and lower accuracy for the training dataset.

[2] For interpretation of color in Fig. 16, the reader is referred to the web version of this article.

Using K-SVM, we replicate the activity sequences of the entire population with 86% accuracy, and its accuracy in replicating agenda is 84%. Small value of edit distance between replicated patterns and observed patterns, computed by sequential alignment techniques, supports the superiority of SVM models over multinomial logit formulation.

Although K-SVM is a very powerful classifying machine with compelling results, the computational cost diminishes its attractiveness. Also similar to other inference model, researchers need to be careful about the overfitting properties, normally it is caused by using large set of features, cross validation is a must do to prevent overfitting. Additionally, unlike the MNL regression model, which provides interpretable coefficients, SVM machine is a "black box" that is trained to estimate the labels of the data and its output is a set of support data-points and their corresponding weights. However, with relatively easy steps, we can evaluate its outcome sensitivity with respect to changes in different variables using support vector weights.

In this study we focused on introducing K-SVM as a classifying tool to study the ordering of activities within patterns with very limited information about the individuals – only commonly-available socio-demographic data that can be gleaned from census data. That SVM can be used as a tool to provide the type of the activities in the agenda and their order in the sequence of activities is a very valuable finding that can be employed to specify the likely agenda and activity order for input to activity-based models. For example, in the rigorously-defined HAPP model and its variations, the optimization of activity/travel linkages is specified only in terms of satisfying spatio-temporal feasibility constraints in a manner that optimizes some specified objective function. So, for example, if the objective is specified in terms of minimizing travel time, the optimal solution is indifferent between dinner followed by a movie and movie followed by dinner, as long as both are completed within the allowable time window – a model such as that developed here could be used to steer the preference toward dinner followed by a movie. In this way, known tendencies/preferences could be incorporated into the mathematical programming framework.

In the current study, in order to prevent overfitting, only a limited number of socio-demographic variables were used to replicate activity patterns; obviously, using standard statistical techniques, feature selection can be improved and extended to include more variables. The goal of this research is to capture the underlying sequential activity selection framework as hidden Markov models with conditional random fields and then to demonstrate that replication of the pattern using such advanced techniques not commonly used in transportation research as SVM can be achieved with accuracy superior to the more conventional MNL.

As stated in the introduction, we believe that relating activity agenda formulation – both in terms of content and sequence – to socio-demographic characteristics is an important step in the extension to the practical application of many activity-based models. Although we believe that this research has provided evidence of the potential for SVM to play a role in this quest, along with numerous other potential applications in transportation, there is undoubtedly substantial room for model improvement in terms of model structure.

Finally, it should be noted that in this study we address only the type and ordering of activities – such parameters as activity location, duration, and mode are not considered. However, adding average travel distance as a spatial indicator to the bins of multinomial choice, or activity duration, would be an interesting extension to the work and can be explored as future extensions of this research.

Acknowledgements

This research was supported, in part, by grants from the University of California Transportation Center and the ITS Multi-Campus Research Program and Initiative on Sustainable Transportation. Their support is gratefully acknowledged. Also, we would like to acknowledge Robert Regue for coding the Differential Evolution algorithm.

Appendix A

Given the primal of SVM to be as following function:

$$\begin{aligned} \min_{w,b,p} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} p_i \\ \text{s.t.}: \quad & l_i(w^T\phi(x_i) - b) - 1 + p_i \geqslant 0 \quad \forall i, \\ & p_i \geqslant 0 \quad \forall i. \end{aligned} \tag{A.1}$$

The Lagrangian of the primal can be written as:

$$L(W, b, p, \eta, \mu) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N} p_i - \sum_{i=1}^{N} \eta_i(l_i(w^T\phi(x_i) - b) - 1 + p_i) - \sum_{i=1}^{N} \mu_i p_i \tag{A.2}$$

Karush–Kuhn–Tucker conditions for the Lagrangian are given as following set of equations:

Table A1
Representing mean value and the lower bound and upper bounds of 95% interval for ‘household size’.

Household size	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$	0.193	0.117	0.269	0.000	−0.138	0.275	0.000	−0.103	0.298	0.000	−0.897	0.707	0.000	−0.666	0.496	0.000	−2.160	1.905
$t = 2$	0.145	0.136	0.153	0.105	0.104	0.105	0.113	0.112	0.114	−0.047	−0.049	−0.045	0.090	0.090	0.091	−0.024	−0.025	−0.024
$t = 3$	0.014	0.014	0.015	−0.042	−0.043	−0.041	−0.040	−0.041	−0.040	−0.070	−0.079	−0.062	−0.033	−0.038	−0.028	−0.020	−0.033	−0.006
$t = 4$	0.075	0.074	0.075	0.051	0.050	0.052	0.088	0.087	0.088	0.000	−0.002	0.014	0.069	0.067	0.071	0.091	0.090	0.091
$t = 5$	0.089	0.089	0.090	0.018	0.016	0.020	−0.145	−0.146	−0.144	−0.135	−0.156	−0.115	0.096	0.092	0.100	0.062	0.060	0.063
$t = 6$	−0.184	−0.185	−0.183	−0.080	−0.088	−0.073	−0.061	−0.063	−0.060	−0.091	−0.101	−0.081	0.020	0.017	0.023	−0.100	−0.101	−0.099
$t = 7$	0.008	0.006	0.009	−0.307	−0.322	−0.293	−0.029	−0.031	−0.027	−0.104	−0.125	−0.084	−0.066	−0.071	−0.060	−0.026	−0.027	−0.024
$t = 8$																		
$t = 9$																0.056	0.016	0.096

Table A2
Representing mean value and the lower bound and upper bounds of 95% interval for model parameters for ‘number of vehicles’.

Total number of vehicles in household	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$																		
$t = 2$	−0.138	−0.198	−0.078	0.127	0.125	0.130	−0.114	−0.118	−0.110	−0.027	−0.038	−0.017	−0.067	−0.071	−0.062	−0.038	−0.041	−0.035
$t = 3$	0.123	0.120	0.126	0.008	0.003	0.013	−0.031	−0.035	−0.027		−0.039	0.070	−0.066	−0.099	−0.033		−0.115	0.056
$t = 4$	−0.029	−0.033	−0.025	−0.033	−0.040	−0.027	−0.031	−0.036	−0.027	−0.219	−0.266	−0.173	0.065	0.051	0.078	0.050	0.046	0.053
$t = 5$	−0.153	−0.157	−0.149	−0.119	−0.130	−0.108	0.147	0.140	0.153		−0.101	0.115	0.064	0.040	0.087	−0.090	−0.096	−0.083
$t = 6$	−0.043	−0.050	−0.037	−0.049	−0.092	−0.005										−0.035	−0.041	−0.030
$t = 7$	0.116	0.108	0.124		−0.057	0.072	0.086	0.072	0.099	−0.230	−0.338	−0.122	−0.152	−0.179	−0.125		−0.008	0.008
$t = 8$	0.026	0.016	0.037		−0.056	0.178	0.099	0.069	0.130							0.019	0.008	0.030
$t = 9$																		

Table A3
Representing mean value and the lower bound and upper bounds of 95% interval for ‘income level’.

Income level	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$	0.331	0.149	0.513															
$t = 2$	−0.087	−0.102	−0.072	0.019	0.018	0.019	−0.029	−0.030	−0.028	0.028	0.025	0.030	0.029	0.028	0.030	−0.043	−0.044	−0.042
$t = 3$	0.075	0.074	0.076	0.113	0.111	0.114	0.015	0.014	0.016	0.000	−0.026	0.001	0.029	0.020	0.037	−0.196	−0.220	−0.171
$t = 4$	0.061	0.060	0.062	0.000	−0.002	0.001	0.135	0.134	0.136	−0.042	−0.055	−0.030	−0.057	−0.060	−0.054	0.116	0.115	0.117
$t = 5$				0.000	−0.003	0.003	0.000	−0.001	0.001									
$t = 6$	−0.122	−0.124	−0.120	−0.108	−0.119	−0.097	0.002	0.000	0.004	−0.111	−0.125	−0.097	−0.031	−0.036	−0.026	−0.002	−0.004	−0.001
$t = 7$	0.012	0.010	0.014	−0.049	−0.067	−0.032	0.104	0.101	0.107	−0.044	−0.074	−0.014	−0.066	−0.074	−0.058	0.100	0.098	0.102
$t = 8$	0.058	0.055	0.060	−0.034	−0.065	−0.003		−0.013	0.003	−0.160	−0.193	−0.126	−0.033	−0.045	−0.022	0.082	0.079	0.085
$t = 9$																0.185	0.129	0.242

Table A4
Representing mean value and the lower bound and upper bounds of 95% interval for ‘student status’.

Students	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$																		
$t = 2$	−0.081	−0.102	−0.060	0.003	0.002	0.004	0.066	0.064	0.068	−0.048	−0.055	−0.041	0.116	0.114	0.119	−0.112	−0.114	−0.110
$t = 3$	0.086	0.085	0.088	−0.033	−0.036	−0.031	0.064	0.062	0.066	0.041	0.020	0.062	0.035	0.023	0.046	0.048	0.020	0.076
$t = 4$	0.059	0.058	0.061	−0.031	−0.034	−0.027	−0.053	−0.055	−0.050	−0.038	−0.058	−0.017	−0.009	−0.014	−0.003	−0.057	−0.059	−0.056
$t = 5$	−0.027	−0.029	−0.024	−0.009	−0.015	−0.003	0.033	0.029	0.036	−0.057	−0.112	−0.003	0.034	0.023	0.045	−0.010	−0.013	−0.007
$t = 6$	−0.052	−0.056	−0.048	−0.135	−0.154	−0.115	−0.057	−0.061	−0.053	−0.153	−0.178	−0.128	−0.084	−0.092	−0.076	0.004	0.001	0.007
$t = 7$	−0.061	−0.066	−0.057				−0.018	−0.024	−0.011				0.090	0.077	0.103	−0.074	−0.079	−0.069
$t = 8$	0.061	0.055	0.066				−0.027	−0.045	−0.008				−0.054	−0.077	−0.031	−0.074	−0.080	−0.067
$t = 9$																−0.112	−0.210	−0.015

Table A5
Representing mean value and the lower bound and upper bounds of 95% interval for 'age'.

Age	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$	0.336	0.255	0.416															
$t = 2$	−0.171	−0.177	−0.165	−0.086	−0.086	−0.086	0.081	0.081	0.081	0.067	0.066	0.068	0.005	0.004	0.005	0.062	0.061	0.062
$t = 3$	0.116	0.116	0.116	0.037	0.036	0.037	0.134	0.134	0.135	0.019	0.013	0.024	−0.063	−0.066	−0.060	−0.070	−0.081	−0.059
$t = 4$	0.091	0.091	0.091	0.124	0.123	0.125	0.095	0.095	0.096	−0.065	−0.070	−0.060	−0.055	−0.056	−0.053	0.007	0.007	0.007
$t = 5$	−0.005	−0.006	−0.005	−0.025	−0.027	−0.024	0.097	0.097	0.098	−0.098	−0.110	−0.086	0.121	0.118	0.124	−0.082	−0.082	−0.081
$t = 6$	0.083	0.082	0.084	−0.131	−0.135	−0.126	−0.002	−0.003	−0.002	−0.182	−0.188	−0.176	−0.140	−0.142	−0.138	−0.008	−0.009	−0.008
$t = 7$	0.084	0.083	0.085	0.036	0.029	0.042	0.172	0.171	0.173	−0.058	−0.069	−0.046	0.083	0.080	0.086	0.113	0.112	0.113
$t = 8$	−0.011	−0.012	−0.010	−0.015	−0.027	−0.003	0.053	0.050	0.056	−0.091	−0.104	−0.077	−0.072	−0.076	−0.068	0.080	0.079	0.081
$t = 9$				0.052	0.005	0.099							−0.166	−0.280	−0.052	0.123	0.099	0.147

Table A6
Representing mean value and the lower bound and upper bounds of 95% interval for ‘gender’.

Gender	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$																		
$t = 2$																		
$t = 3$	0.024	0.024	0.024	−0.004	−0.004	−0.003	−0.138	−0.138	−0.137	−0.015	−0.024	−0.007	−0.068	−0.073	−0.063	0.062	0.046	0.078
$t = 4$	−0.014	−0.014	−0.013	−0.132	−0.133	−0.131	−0.010	−0.011	−0.010	−0.073	−0.080	−0.067	−0.089	−0.090	−0.087	0.074	0.073	0.074
$t = 5$	0.019	0.019	0.020	0.044	0.042	0.045	−0.103	−0.104	−0.103	0.069	0.053	0.085	0.065	0.062	0.068	−0.078	−0.079	−0.077
$t = 6$	0.118	0.117	0.119	0.019	0.012	0.025	0.023	0.022	0.024	−0.060	−0.068	−0.051	−0.111	−0.114	−0.108	0.012	0.012	0.013
$t = 7$	0.084	0.083	0.085	0.036	0.029	0.043	0.172	0.170	0.174	−0.058	−0.071	−0.044	0.083	0.079	0.086	0.113	0.112	0.114
$t = 8$	0.133	0.131	0.134	−0.044	−0.058	−0.030	0.047	0.043	0.050	−0.034	−0.049	−0.018	−0.021	−0.027	−0.016	−0.014	−0.016	−0.013
$t = 9$				0.000	−0.037	0.073				0.000	−0.563	0.414						

Table A7
Representing mean value and the lower bound and upper bounds of 95% interval for ‘employment status’.

Employment status	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
t = 1																		
t = 2				0.202	0.200	0.204	0.149	0.145	0.152	0.035	0.025	0.046	0.042	0.037	0.047	0.011	0.009	0.014
t = 3	0.005	0.003	0.007	0.075	0.071	0.079	0.046	0.043	0.049				0.045	0.011	0.079		−0.201	0.032
t = 4	0.051	0.047	0.054	0.076	0.070	0.081	0.007	0.004	0.011				−0.079	−0.090	−0.067	−0.159	−0.162	−0.156
t = 5	0.084	0.080	0.087	0.040	0.029	0.051	0.151	0.146	0.156				−0.066	−0.092	−0.040	−0.078	−0.084	−0.072
t = 6	−0.053	−0.059	−0.047	−0.057	−0.099	−0.015	−0.041	−0.047	−0.034	−0.089	−0.144	−0.035	−0.047	−0.066	−0.029	−0.085	−0.090	−0.080
t = 7	−0.052	−0.059	−0.045				−0.016	−0.027	−0.004				−0.047	−0.071	−0.022	−0.213	−0.220	−0.205
t = 8	0.144	0.134	0.153							0.114	0.006	0.222	−0.050	−0.088	−0.013	0.026	0.016	0.036
t = 9																		

Table A8
Representing mean value and the lower bound and upper bounds of 95% interval for ‘education level’.

Education	Home			Work			Main			Personal			Pickup			Stop		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
$t = 1$	0.784	0.705	0.862															
$t = 2$	−0.138	−0.143	−0.132	0.189	0.180	0.188	0.126	0.124	0.128				0.044	0.044	0.045	0.203	0.202	0.203
$t = 3$	0.193	0.192	0.191	0.104	0.103	0.106	0.190	0.189	0.191	−0.166	−0.172	−0.161	−0.062	−0.065	−0.059	−0.136	−0.145	−0.127
$t = 4$	0.041	0.040	0.042	−0.002	−0.003	−0.002	0.005	0.003	0.007	−0.063	−0.067	−0.059	0.011	0.010	0.012	0.069	0.069	0.069
$t = 5$	0.185	0.183	0.187	0.073	0.072	0.074	0.100	0.100	0.101	−0.128	−0.138	−0.118	−0.049	−0.051	−0.047	0.197	0.196	0.197
$t = 6$	0.141	0.139	0.142	−0.093	−0.099	−0.086	0.010	0.08	0.012	−0.070	−0.075	−0.065	−0.063	−0.065	−0.061	0.110	0.115	0.120
$t = 7$	0.151	0.150	0.152	−0.073	−0.078	−0.068	0.026	0.025	0.027	−0.074	−0.082	−0.065	0.061	0.059	0.063	0.082	0.081	0.083
$t = 8$	0.113	0.112	0.114	−0.084	−0.094	−0.075							0.085	0.082	0.088	0.116	0.112	0.1120
$t = 9$																0.334	0.315	0.352

$$I.\ \frac{\partial L}{\partial W}=0,\ w-\sum_{i=1}^{N}\eta_i l_i\phi(x_i)=0,\ W^*=\sum_{i=1}^{N}\eta_i l_i\phi(x_i)$$

$$II.\ \frac{\partial L}{\partial b}=0,\ \sum_{i=1}^{n}\eta_i l_i=0$$

$$III.\ \frac{\partial L}{\partial p}=0,\ C-\eta_i-\mu_i=0$$

$$IV.\ l_i(W^T\phi(x_i)-b)-1+p_i\geqslant 0$$

$$V.\ p_i\geqslant 0$$

$$VI.\ \mu_i\geqslant 0$$

$$VII.\ \eta_i[l_i(W^T\phi(x_i)-b)-1+p_i]=0$$

$$VIII.\ \mu_i p_i=0$$

According to condition VI, μ_i is either zero or it is a positive number.
Case I: $\mu_i = 0$

$$\eta_i+\mu_i=C,\ \text{then}\ \eta_i=C, p_i>0$$

$$L_D=\frac{1}{2}W^2+C\sum_{i=1}^{N}p_i-\sum_{i=1}^{N}\eta_i\left[l_i(W^T\phi(x_i)-b)-1+p_i\right]-\sum_{i=1}^{N}\mu_i p_i \tag{A.3}$$

Inserting W as $W^*=\sum_{i=1}^{N}\eta_i l_i\phi(x_i)$ from equation I in A.3, the Lagrangian can be simplified as A.4.

$$L_D=\frac{1}{2}\sum_{ij}^{N}\eta_i l_i\eta_j l_j\phi^T(x_i)\phi(x_j)+\sum_{i=1}^{N}\eta_i p_i-\sum_{ij}^{N}\eta_i\left[l_i(\eta_j l_j\phi^T(x_i)\phi(x_j)-b)-1+p_i\right] \tag{A.4}$$

By simplifying A.4, the dual for the binary SVM function can be written as A.5:

$$\begin{aligned}\max\quad & L=-\frac{1}{2}\sum_{ij}^{N}\eta_i l_i\eta_j l_j\phi^T(x_i)\phi(x_j)+\sum_{ij}^{N}\eta_i\\ \text{s.t.}:\quad & \sum_i\eta_i l_i=0,\quad\forall i,\\ & 0\leqslant\eta_i\leqslant C,\quad\forall i.\end{aligned} \tag{A.5}$$

Case II: $\mu_i > 0$ then $p_i = 0$ (KKT condition VIII), the Lagrangian can be written as:

$$L_D=\frac{1}{2}\sum_{ij}^{N}\eta_i l_i\eta_j l_j\phi^T(x_i)\phi(x_j)-\sum_{ij}^{N}\eta_i\left[l_i(\eta_j l_j\phi^T(x_i)\phi(x_j)-b)-1\right] \tag{A.6}$$

Based on KKT condition II:
$\sum_{ij}^{N}\eta_i y_i b=b.\sum_{ij}^{N}\eta_i y_i=0$, then the dual for binary SVM is:

$$\begin{aligned}\max\quad & L=-\frac{1}{2}\sum_{ij}^{N}\eta_i l_i\eta_j l_j\phi^T(x_i)\phi(x_j)+\sum_{ij}^{N}\eta_i\\ \text{s.t.}:\quad & \sum_i\eta_i l_i=0,\quad\forall i,\\ & 0\leqslant\eta_i\leqslant C,\quad\forall i.\end{aligned} \tag{A.7}$$

See Tables A1–A8.

References

Arentze, T.A., Ettema, D., Timmermans, H., 2011. Estimating a model of dynamic activity generation based on one-day observations: method and results. Transportation Research Part B 45 (2), 447–460.
Arentze, T.A., Timmermans, H., 2009. A need-based model of multi-day, multi-person activity generation. Transportation Research Part B 43 (2), 251–265.
Arentze, T.A., Timmermans, H., 2000. Albatross: A Learning-based Transportation Oriented Simulation System. European Institute of Retailing and Services Studies, Eindhoven University, Eindhoven.
Bhat, C.R., 2004. Comprehensive econometric microsimulator for daily activity-travel patterns. Transportation Research Record: Journal of the Transportation Research Board, 57–66 (No. 1894, Transportation Research Board of the National Academies, Washington, D.C.).
Bhat, C.R., Srinivasan, S., 2005. A multidimensional mixed ordered-response model for analyzing weekend activity participation. Transportation Research Part B 39 (3), 255–278.
Bhat, C.R., 2005. A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. Transportation Research Part B 39 (8), 679–707.
Bhat, C.R., Goulias, K.G., Pendyala, R.M., Paleti, R., Sidharthan, R., Schmitt, L., Hu, H., 2013. A household-level activity pattern generation model with an application for southern California. Transportation 40 (5), 1063–1086.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, New York.
Bowman, J.L., Ben Akiva, M.E., 2000. Activity-based disaggregate travel demand model system with activity schedules. Transportation Research Part A 35 (1), 1–28.
California Department of Transportation, 2002. California Statewide Household Travel Survey.
Cheu, R.L., Srinivasan, D., Teh, E.T., 2003. Support vector machine models for freeway incident detection. In: The Proceedings of Intelligent Transportation Systems. IEEE, pp. 238–243.
Gan, L.P., Recker, W.W., 2013. Stochastic pre-planned household activity pattern problem with uncertain activity participation (SHAPP). Transportation Science. 47 (3), 439–454.
Goulias, K.G., 1999. Longitudinal analysis of activity and travel pattern dynamics using generalized mixed Markov latent class models. Transportation Research Part B 33 (8), 535–558.
Joh, C.H., Arentze, T., Timmermans, H., 2001. Multidimensional sequence alignment methods for activity-travel pattern analysis: a comparison of dynamic programming and genetic algorithms. Geographical Analysis 33 (3), 247–270.
Joh, C.H., Arentze, T., Hofmanb, F., Timmermans, H., 2002. Activity pattern similarity: a multidimensional sequence alignment method. Transportation Research Part B 36 (5), 385–403.
Kitamura, R., Chen, C., Pendyala, R., 1997. Generation of synthetic daily activity-travel patterns. Transportation Research Record: Journal of the Transportation Research Board, 154–163 (No. 1607, Transportation Research Board of the National Academies, Washington, D.C.).
Lafferty, J., Mccallum, A., Pereira, F.C.N., 2001. Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of International Conference on Machine Learning, pp. 282–289.
Leszczyc, P.T.L.P., Timmermans, H., 2002. Unconditional and conditional competing risk models of activity duration and activity sequencing decisions: an empirical comparison. Journal of Geographical Systems 4 (2), 157–170.
Liao, L., Patterson, D.J., Fox, D., Kautz, H., 2007. Learning and inferring transportation routines. Artificial Intelligence 171 (5–6), 311–331.
McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239–245.
Moons, E., Wets, G., Aerts, M., 2007. Nonlinear models for determining mode choice (accuracy is not always the optimal goal). In: Proceedings of the Artificial Intelligence 13th Portuguese Conference on Progress in Artificial Intelligence, Guimarães, Portugal, pp. 183–194.
Needleman, S.B., Wunsch, C.D., 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. Journal of Molecular Biology 48 (3), 443–453.
Pendyala, R., Kitamura, R., Chen, C., Pas, E., 1997. An activity-based micro-simulation analysis of transportation control measures. Transport Policy 4 (3), 183–192.
Rakotomamonjy, A., 2003. Variable Selection Using SVM-based Criteria. The Journal of Machine Learning Research 3, 1357–1370.
Recker, W.W., 2001. A bridge between travel demand modeling and activity based travel analysis. Transportation Research Part B 35 (5), 481–506.
Recker, W.W., Mcnally, M.G., Root, G.S., 1986a. A model of complex travel behavior: Part I – theoretical development. Transportation Research Part A 20 (4), 307–318.
Recker, W.W., Mcnally, M.G., Root, G.S., 1986b. A model of complex travel behavior: Part II – an operational model. Transportation Research Part A 20 (4), 319–330.
Recker, W.W., 1995. The household activity travel pattern problem: general formulation and solution. Transportation Research Part B 29 (1), 61–77.
Recker, W.W., Parimi, A., 1999. Development of a microscopic activity-based framework for analyzing the potential impacts of transportation control measures on vehicle emissions. Transportation Research Part D 4 (6), 357–378.
Sammour, G., Bellemans, T., Vanhoof, K., Janssens, D., Kochan, B., Wets, G., 2012. The usefulness of the sequence alignment methods in validating rule-based activity-based forecasting models. Transportation 39 (4), 773–789.
Serfozo, R.F., 1979. An equivalence between continuous and discrete time Markov decision processes. Operations Research 27 (3), 616–620.
Smith, L., Beckman, R., Anson, D., Williams, M., 1995. TRANSIMS: TRansportation ANalysis and SIMulation System. Los Alamos National Laboratory Unclassified Report.
Storn, R., Price, K., 1997. Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11 (4), 341–359.
Theja, P.V.V.K., Vanajakshi, L., 2010. Short term prediction of traffic parameters using support vector machines technique. In: proceeding of 3rd International Conference on Emerging Trends in Engineering and Technology (ICETET), pp. 70–75.
Weston, J., Watkins, D., 1998. Multi-class Support Vector Machines. Department of Computer Science, Royal Holloway, University of London.
Yang, Y., Yaom, E., Yue, H., Liu, Y., 2010. Trip chain's activity type recognition based on support vector machine. Journal of Transportation Systems Engineering and Information Technology 10 (6), 70–75.
Zhang, Y., Xie, Y., 2008. Travel mode choice modeling with support vector machines. Transportation Research Record: Journal of the Transportation Research Board, 141–150 (No. 2076, Transportation Research Board of the National Academies, Washington, D.C.).