INFERRING ACTIVITY SELECTION AND SCHEDULING BEHAVIOR OF POPULATION COHORTS FOR TRAVEL DEMAND MODELING

Mohammad Hesam Hafezi, Dalhousie University
Hugh Millward, Saint Mary's University
Lei Liu, Dalhousie University
Proceedings of the 53nd Annual Conference
Canadian Transportation Research Forum

Introduction

Transportation can be considered as one of the main and essential human activities, that involves nearly everyone on a daily basis. To date, numerous travel demand models have been developed, using both aggregated and disaggregated approaches, for modeling short-term and long-term choices of travelers, such as activity participation, timing, transport mode, activity location, route choice, work/residential location, and vehicle ownership (Oppenheim, 1995; Ortuzar & Willumsen, 2011). Complexities in individual travel behavior have increased with continued urban development and rapid technological progress. Trip chaining and multimode transport, flexible working hours, self-employment, and online shopping have become far more common in recent years (Goran, 2001). As travel behavior becomes more complex, travel demand forecasting requires more detailed information. From a disaggregated modeling point of view, there are significant associations between trips and the activity participation of travelers (Kitamura, Chen, & Pendyala, 1997). Furthermore, travelers with varying socio-demographic and socio-economic characteristics in the region have divergent time-use activity patterns. This paper presents a new disaggregated travel demand microsimulation model framework that is sensitive to the mix of variables connected to travelers' decisions. New pattern recognition and inference models are developed to identify population clusters with homogeneous time-use daily activity patterns, and to predict activity selection and scheduling behavior of these population cohorts. The representative behavior within each cluster is then used as an information guide for modeling the 24-hour activity schedule and the travel linked to it. The proposed model is applied to data from the large Halifax STAR household travel diary survey. The proposed modeling framework has much higher reproducibility and shorter computational time compared to alternative modeling frameworks.

Data

This study utilizes time-diary and GPS geo-coordinate data as the primary data source, from the Space-Time Activity Research (STAR) survey undertaken in Halifax, Canada. The STAR survey represented the world's first large-scale application of global positioning system (GPS) technology for a household activity survey. The unique and rich Halifax STAR project produced a wide variety of data, including: (1) household roster data, (2) main file, (3) vehicle data, (4) time diary (episode and summary data file), (5) activity diary (episode data file), (6) land use database, (7) business hours survey data, (8) places and locations (PAL) directory data, and (9) global positioning systems (GPS) data. Full descriptions of the survey design and the socio-demographic features of respondents can be found in (Millward & Spinney, 2011; TURP, 2008).

The Halifax STAR project produced survey data from 1,971 randomly designated households in Halifax Regional Municipality (HRM) between April 2007 and May 2008. A primary respondent over age 15 was randomly selected in each household, completed a 2-day time diary, and carried a GPS unit (Hewlett Packard iPAQ hw6955) during all out-of-home activity. The respondent then completed a "day-after" debriefing through a Computer-Assisted Telephone Interview (CATI), to verify activities, times, and locations, supplemented and verified through GPS tracking The original STAR dataset comprised 188 activity sub categories defined under ten major activity classes. The activity codes employed were similar to those utilized in Statistics Canada's General Social Survey (GSS) time-use surveys, and relate to the prime purpose of the activity.

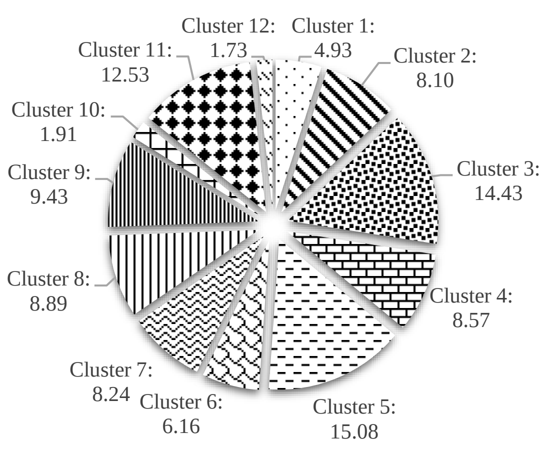
Methods

We provide here only a brief overview of the methods, and interested readers are referred to (Hafezi, Liu, & Millward, 2017, 2018a, 2018b) for more details. Initially, time-use activity patterns were used to derive population clusters displaying homogeneous activity, and the activity patterns for each identified cluster are modeled using a series of behaviorally realistic advanced econometric and machine learning micro-behavioral modules. In the next stage, the type and frequency of activities in the schedule, and their sequential arrangement, are predicted for modeling agenda formation. Subsequently, temporal information associated with the traveler's daily activity schedule are modeled. Finally, the representative activity pattern in each cluster is applied to a rule-based heuristic algorithm to schedule the activities of individuals with varying characteristics and behavior, based on their priority importance and empirical guide information.

The advanced machine learning based micro-behavioral modules developed are novel, time-efficient, and of practical use. A unique feature of the developed model that makes it different from other existing techniques is its degree of efficiency both in computational time and in minimizing exogenous errors. Furthermore, the proposed pattern recognition model improves on traditional models by using socio-demographic variables to classify the population into homogeneous clusters, based on time-use mobility patterns. Another advantage of the new proposed model is that, unlike previous approaches, the algorithm can recognize groups of people who typically tend to avoid travel in peak traffic periods. Furthermore, the inference model predicts both frequent and infrequent activities in the traveler's agenda.

Results

The proposed modeling framework was applied to 2,778 activity days drawn from the large Halifax STAR survey data. Individuals with homogeneous activity patterns were identified through the machine learning process and grouped into twelve clusters, as shown in Figure 1. In general, the algorithm identified six clusters with heterogeneous activity patterns for worker groups. For the non-worker groups, the algorithm recognized four clusters with dissimilar activity patterns. Lastly, two single clusters were identified for students and for people who mostly spent their time at home.



Cluster #1: Extended work-day workers
Cluster #2: Non-worker
Cluster #3: 8-4 workers
Cluster #4: Non-worker
Cluster #5: Stay-at-homes
Cluster #6: Shorter work-day workers
Cluster #7: 7-3 workers
Cluster #8: Non-worker, morning shopping
Cluster #9: Non-worker, afternoon shopping
Cluster #10: Evening workers
Cluster #11: 9-5 workers
Cluster #12: Students

Figure 1. Twelve Identified Clusters (with percentages of respondents).

The duration misclassification error in the scheduled activity for each cluster was computed and is shown in Table 1. The total error for each population cluster was estimated based on the summation of all misclassification errors over 24-hours of scheduled activity. The highest error percentage was found for the student cluster (no. 12), at 36.71%, followed by the evening worker cluster (no. 10), at 25.50%. Compared to other clusters, students and evening workers had the lowest sample size in our model. Our full results also show that the highest misclassification error in each cluster is for those activities with a shorter duration in the traveler's daily activity patterns. Empirical results therefore reveal that the proposed model can predict response variables with more precision when trained with a larger dataset. The mean estimation error for all twelve clusters in our model is 18.38% in the 24-hour period.

Table 1. Scheduling Estimation Accuracy (Duration of Misclassification)

	Mean estimation error (minutes)											
Population Cluster Number	**#1**	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**	**#8**	**#9**	**#10**	**#11**	**#12**
Total error in 24-h	285.4	225.2	330.4	234.9	212.9	221.7	279.2	172.7	172.4	367.2	189.4	528.6
Estimation accuracy (%)	80.2	84.4	77.1	83.7	85.2	84.6	80.6	88.0	88.0	74.5	86.9	63.3

Conclusion

Implementation of the proposed model shows that it can assemble the traveler's schedule with an average of 81.62% accuracy in the 24-hour period. In summary, the new microsimulation modeling framework proposed in this study offers a straightforward and easy-to-implement tool for transport modelers to model time-use activity patterns for different population cohorts in the region. Moreover, the proposed modeling framework can be used to advance transportation demand management for different cohorts of the urban population, as well as to analyze environmental mitigation and transport policy scenarios. The results of this study are expected to be implemented within the activity-based travel demand model, Scheduler for Activities, Locations, and Travel (SALT) for Halifax, Nova Scotia. The modeling framework can be readily adapted to the modeling of transportation demands for other major urban centers in North America.

Acknowledgments

This research was supported, in part, by grants from Dalhousie University. Their support is gratefully acknowledged. The authors would also like to thank the Natural Sciences and Engineering Research Council (NSERC) for their contribution in supporting the research. We would also like to thank the Dalhousie Transportation and Environmental Simulation Studies (TESS) group members for their valuable suggestions. Data for this research were provided by the Halifax STAR Project, supported through the Atlantic Innovation Fund from the Atlantic Canada Opportunities Agency, Project No.181930.

References

Goran, J. (2001). *Activity based travel demand modeling-a literature study*. Technical report, Danmarks Transport-Forskning.

Hafezi, M. H., Liu, L., & Millward, H. (2017). A Time-Use Activity-Pattern Recognition Model for Activity-Based Travel Demand Modeling. *Transportation*, 1-26. doi:10.1007/s11116-017-9840-9

Hafezi, M. H., Liu, L., & Millward, H. (2018a). Learning Daily Activity Sequences of Population Groups Using Random Forest Theory. *Transportation Research Record: Journal of the Transportation Research Board, Transportation Research Board of the National Academies, Washington, D.C.*

Hafezi, M. H., Liu, L., & Millward, H. (2018). *Modeling Activity Scheduling Behavior of Travelers for Activity-Based Travel Demand Model*. Paper presented at the 97th Annual Meeting of the Transportation Research Board,. Washington, D.C.

Kitamura, R., Chen, C., & Pendyala, R. (1997). Generation of Synthetic Daily Activity-Travel Patterns, *Transportation Research Record, No. 1607*, 154-163. doi:10.3141/1607-21

Millward, H., & Spinney, J. (2011). Time use, travel behavior, and the rural–urban continuum: results from the Halifax STAR project. *Journal of Transport Geography, 19*(1), 51-58. doi:http://dx.doi.org/10.1016/j.jtrangeo.2009.12.005

Oppenheim, N. (1995). *Urban Travel Demand Modeling: From Individual Choices to General Equilibrium*: Wiley.

Ortuzar, J. D. D., & Willumsen, L. G. (2011). *Modeling Transport*: Wiley.

TURP. (2008). *TURP (Time Use Research Program)*. Retrieved from Halifax regional space time activity research (STAR) survey: A GPS-assisted household time-use survey, survey methods. Halifax: Saint Mary's University.