2014 ASHRAE/IBPSA-USA
Building Simulation Conference
Atlanta, GA
September 10-12, 2014

A FRAMEWORK FOR MODELING OCCUPANCY SCHEDULES AND LOCAL TRIPS BASED ON HUMAN ACTIVITY PATTERNS IN THE CITY

ABSTRACT

This paper presents a methodology for the generation of building occupancy schedules and trips in urban areas based on regional activity-based travel surveys. Using eigendecomposition-based Principal Component Analysis (PCA) and the k-means clustering algorithm, the method identifies activities as well as city resident types. The outputs are typical occupancy schedules for building performance simulation, and activity profiles for mobility modeling in urban-scale simulation. As a proof of concept the framework is applied to the 2010/2011 Massachusetts Department of Transportation (MassDOT) travel survey. Four clusters of occupants were identified, and an example building occupancy schedule and complementary activity distributions were determined. A discussion about how such activity profiles can be used to generate a set of user behavior profiles for neighborhood and city models is presented.

INTRODUCTION

The current human world is, and will continue to be, mostly urbanized; with cities being planned, build and inhabited at an unprecedented rapid rate (WHO, 2013). In order for this urbanization process to be sustainable, architects and planners apply measures of energy efficiency and carbon emissions reduction on an urban scale. The employment of Building Performance Simulation (BPS) tools supports such methods by predicting the annual energy use of buildings using mathematical representation of thermal and luminous environments, within an acceptable margin of error (Fabi et al., 2013). These programs use complex models of occupant behavior within buildings to predict the control of various components, such as window opening (Fritsch et al, 1990), lighting controls and blind adjustments (Reinhart, 2004) and space heating/cooling demands (Hoes et al, 2009). However, examining individual buildings when modeling urban developments is insufficient, as there is a need to study how buildings interact with each other when modeling hundreds of buildings. In an urban context, occupants' movements between buildings during daily activities affects transportation energy. This makes the simulation of activities in urban areas critical to the assessment of trips made for local transportation and consequent mode choices (walking, biking, automobile, etc.) based on proximity.

Travel behavior and forecasting is a topic of interest for urban planners and transportation engineers, and links to the built environment have been established in the literature.The relationship between land use and travel behavior based on utility-based and activity-based theory of transportation was discussed (Maat et al, 2005), the impacts the built environment has on travel behavior have been presented (McCormack, 2001; Handy 2002), and microsimulation of activity-based travel patterns was developed for travel forecasting (Kitamura et al, 2000). However, BPS has not benefited from the relationship between urban mobility and occupants travel. This work argues that understanding travel behavior in urban areas can enhance simulation of energy flows in and around groups of buildings in BPS, and can provide architects, urban designers and planners with the tools to promote sustainable mobility in terms of promoting the walkability and bikeability of cities.

This paper presents a framework to generate building occupancy profiles based on analysis of human activity patterns in metropolitan areas. The method is developed for building and urban performance simulation, where the outputs are twofold: typical occupancy schedules for operational energy use simulation of buildings, and activity distributions for urban simulation. The paper is divided into four sections: the methodology details data mining and analysis procedures for typical activity-based travel databases, followed by the results of research focused on the Massachusetts Department of

Transportation (MassDOT) travel survey that illustrates BPS relevant occupancy schedules and activities for urban modeling simulations. A discussion is presented next, highlighting potentials and limitations, followed by a conclusion of the presented framework which suggests future work.

METHODOLOGY

Figure 1 demonstrates the developed framework. An activity based mobility database is analyzed in order to identify what kind of activities exists in the chosen urban area, as well as to derive occupant behavior archetypes. Schedules and behavior profiles for different occupants are then generated. The details of each stage are described next.

Input: Activity Based Mobility Database – Travel Survey

In order to understand complex activities that occur in urban areas and the dynamics of its inhabitants, an activity-based travel database is used. The use of cell phone traces (Calabrese et al, 2013), online social media check-ins (Noulas et al, 2011) and position tracking (Gonzalez et al, 2008) to comprehend human mobility and associated activities are emerging urban sensing techniques. The authors forsee that the future of activity-based mobility databases to be used in this workflow will be based on such technologies, but at the time of writing this paper, they are still being developed and are hindered by access to such data due to privacy restrictions and limitations of cell phone technologies. Traditional travel surveys are publically available and are the standard means of understanding mobility in cities. In this paper, we use the MassDOT Travel survey as the database that will be analyzed in order to understand trends and patterns in urban human mobility.

The activity survey data used included 15,000 households between June 2010 and November 2011, where participants were asked to identify where and how they traveled on a specific, designated travel day (24 hours). The sample was carefully chosen to represent the Massachusetts population by asking each participating houshold a series of detailed questions about their socioeconomic characteristics and access to transportation. Data mining techniques used to analyze this database are presented in the following section.

Note that the only survey travel days assigned were weekdays, so our analysis excludes weekends. Our analysis also excludes Friday reports, as Friday travel behavior substantially differs from that of other weekdays.

Identifying Urban Activities and Occupant Types

In order to generate occupant profiles from the survey data, we use the mathematical approach presented by Jiang et al for the the Chicago metropolitan area. An overview is provided here, but readers desiring a rigorous mathematical treatment (including specific methodological justifications) are encouraged to refer to that paper (Jiang et al, 2012).

The method's thrust is the clustering of survey respondents based on behavior patterns, and this grouping is accomplished by the application of the k-means clustering algorithm (Wu et al, 2008). This process organizes groups of points in some (potentially high-dimensional) space into clusters, and a way to apply it to categorical data, such as transportation activity participation, is non-obvious, since there is no clear metric for "closeness" between any two particular travel activities. The method devised by Jiang et al is similar to a proposal by Ralambondrainy, encoding the activities using binary attributes in a manner that creates an implicit and sensible distance metric so that k-means can then operate (Ralambondrainy, 1995).

MATLAB was used as the numerical computing environment. First, each respondent's day is broken into ninety-six 15-minute intervals, and each interval is assigned the activity that the respondent is performing at the beginning of it. (The original 25 activity choices were compressed into nine – see Table 1.) For this purpose, travel is assumed to be instantaneous, with each activity beginning as soon as the traveller leaves their previous activity.

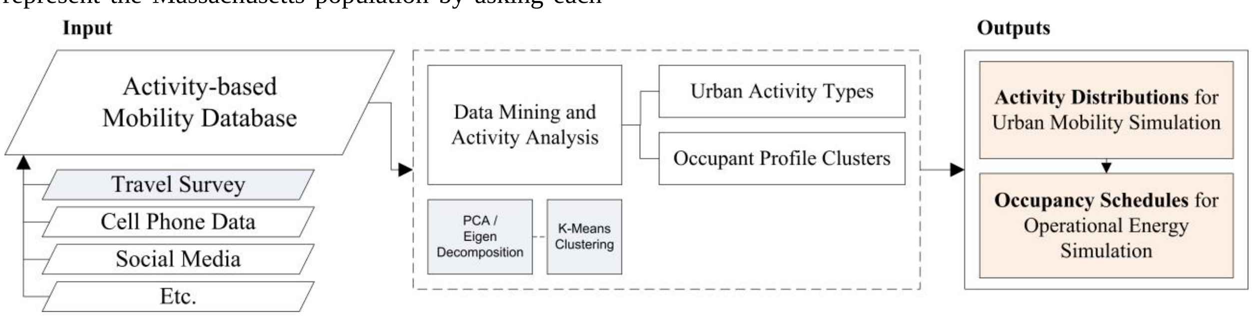

Figure 1 Presented framework

From this activity description, nine 96-element vectors are then generated, one corresponding to each activity. In each of these vectors, the *i*th element is set to 1 if the respondent is performing its associated activity during time interval *i*, and 0 otherwise. Finally, these nine vectors are concatenated, producing a single, 864-element vector of zeros and ones that describes the respondent's activity throughout the day.

Table 1 Aggregated activities types and original trip purposes

ACTIVITY	ORIGINAL PRIMARY TRIP PURPOSE
Home	- Working at home (for pay) - All other activities at home
Work	- Work/job - All other activities at work - Volunteer work/activities - Work business related
School	- Attending class - All other school activities
Transit	- Changed type of transportation - Drop off passenger from car - Pick up passenger from car - Loop trip
Shopping/ Errands	- Service private vehicle (gas, oil lube, etc.) - Routine shopping (groceries, clothing, convenience store, HH maintenance) - Shopping for major purchases or specialty items (appliance, electronics, new vehicle, major HH repairs) - Household errands (bank, dry cleaning)
Personal	- Personal business (visit government office, attorney, accountant) - Health care (doctor, dentist)
Recreation/ Entertainment	- Eat meal outside of home - Outdoor recreation/entertainment - Indoor recreation/entertainment - Visit friends/relatives
Civic/ Religious	- Civic/religious activities
Other	- While traveling – other - Other

As performing k-means clustering on these 864-dimensional activity vectors would be computationally expensive, we turn to eigendecomposition-based Principal Component Analysis (PCA) (Hastie et al, 2009) in order to dimensionally reduce the input data so that clustering can be performed. Some quantity of 864-dimensional *eigenactivities* is generated such that each daily activity vector can be represented as a linear combination of them. Each linear combination can be "reconstructed" into a daily activity vector by identifying the highest-valued activity of each time step, setting the corresponding reconstructed vector elements to 1, and setting all other vector elements to 0. As the number of eigenactivities increases, the error between reconstructed activity vectors and original activity vectors drops to zero. Following Jiang et al, we choose *h* eigenactivities such that the error is <1%. We can represent each activity vector using only the coefficients of its linear combination of eigenactivities, of which there are *h*, so we have reduced the dimension of our activity representation from 864 to *h*. As as long as *h* is small enough, k-means is now fast enough to be applicable.

Application of the k-means clustering algorithm requires a determination of an appropriate number of clusters. In order to define this, we examine Cluster Silhouettes (Rousseeuw, 1987), which are measures of a cluster's tightness and separation, and can be used to identify the optimal numbers of clusters.

Output: Occupant Schedules and Behavior Profiles

Two outputs are created through the previous data mining workflow. First, a typical at-work BPS occupancy schedule based on a "Worker" cluster, where an office building occupancy percentage of each timestep is the percentage of respondents in the cluster who are at work during that timestep. Second, an activity distribution profile for each of the clustered users. This sets the basis for modeling mobility in an urban area by identifying the type of activities people engage in and their expected activity-based travel within the city. We applied the framework in the state of Massachusetts, and results are detailed next.

RESULTS

Figure 2 shows time of day activity variations aggregated for all samples of the MassDOT travel survey. This is the first step to construct eigenbehaviors, cluster participants and produce schedules.

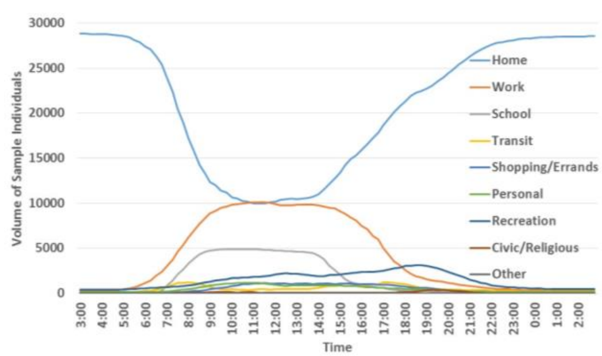

Figure 2 Temporal rhythm of activities on a weekday in Massachusetts

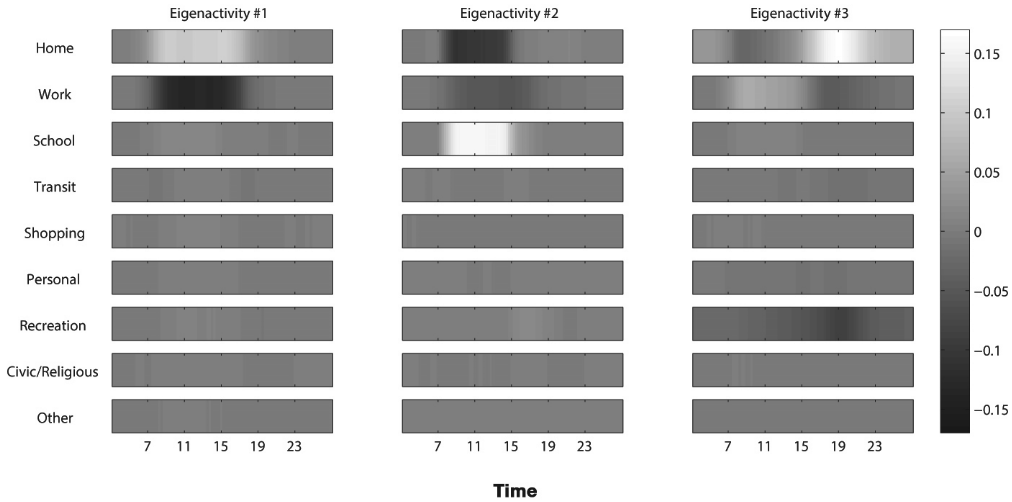

Figure 3 the first three eigenactivities of a weekday

Eigenactivities

In order to understand the structure of basic human activity patterns in the survey, eigenactivities were computed for all sample individuals. Twenty three eigenactivities were derived for the weekday, and Figure 3 demonstrates the first three as an example. The first eigenactivity (left) represents a high probability of staying home and not going to work between 7 AM and 5 PM. The second eigenactivity shows that between 8 AM and 3 PM there's a high probability of going to school and a low probability of staying at home, and the third eigenactivity displays relatively higher probablity of staying at home from 3 PM to 11 PM coupled with a low probability of going for recreation, and a high probability of working from 7 AM to 12 PM. In order to group individuals based on activity sequences during the weekday, the k-means clustering algorithm is used

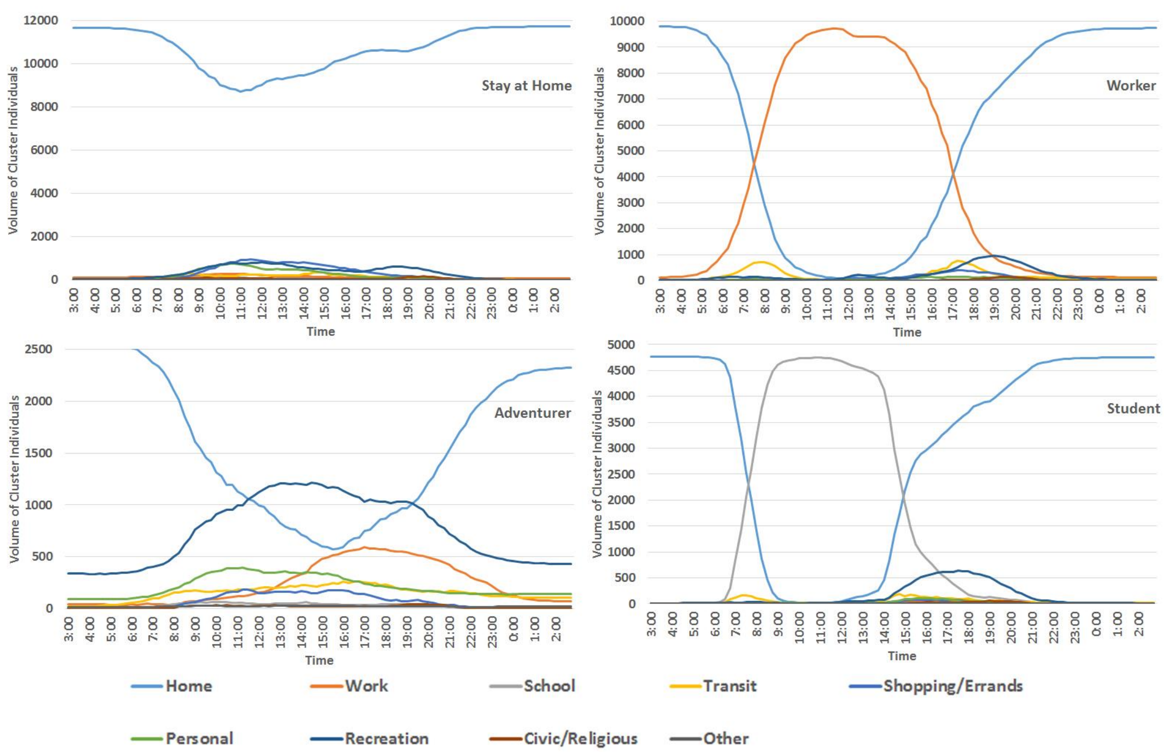

Figure 4: Four clusters of activity patterns

Clustering Users

The examination of cluster silhouettes for various cluster counts suggested four clusters for individuals, and they were observed as Stay at Home, Worker, Adventurer and Student as demonstrated by Figure 4.

As their names suggests, each cluster spends most its time of day engaging in the activity that its title proposes. However, it is important to note that other activities should not be disregarded, as specific patterns emerge with each cluster. For example, a student is likely to go for a recreational activity after school hours and up until 10 PM and an adventurer, who spends most of his/her time on recreation still goes to work later in the day.

Occupancy Schedules and Behavior Profiles

Through the presented analysis of activity patterns in the city, two outputs are extracted to be used in the field of BPS:

- Occupancy Schedules

Typical schedules used for energy simulation use diversity factors for weekday and weekend patterns in using buildings. Such diversity factors were extracted from the Worker cluster to demonstrate the pattern of office building types weekday according to analysis of the MassDOT survey, and are presented in Figure 5.

The schedule starts at 6 AM, where 10% of occupants are expected to arrive at their work place, with a linear increase until 11 AM, where 97% of the users are present, and an expected decrease around noon takes place for lunch time, but it is only a change of 5%. A steady decline then of workers leaving the work place takes place until 6 PM, where only 20% of the occupants remain and decrease to 5% 8 PM to reach 1% at 10 PM.

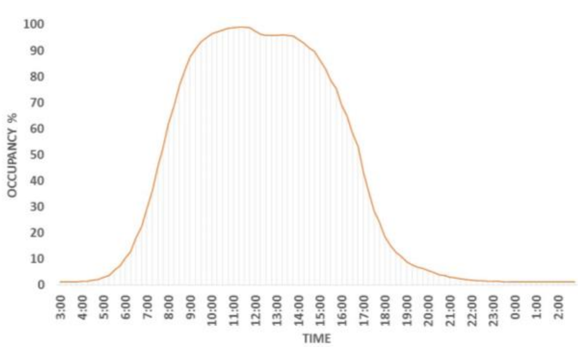

Figure 5 Office building occupancy schedule as extracted from the Worker cluster

- Behavior Profiles

For the simulation of activities taking place in urban areas, behavior distribution profiles are also extracted from the clusters. Figure 6 demonstrates the distribution of urban activities for the Worker cluster. While members of this cluster spend most of their time either at home or at work, other activities present themselves temporally, which gives a better understanding of the kind of activities this cluster would engage in a city.

The Worker cluster activity profile gradually replaces staying at home temporally with work, with the rise of time specific activities mainly represented in transit from home in the morning and to home in the afternoon. A significant portion of other activities goes to recreation starting 4 PM up to 10 PM.

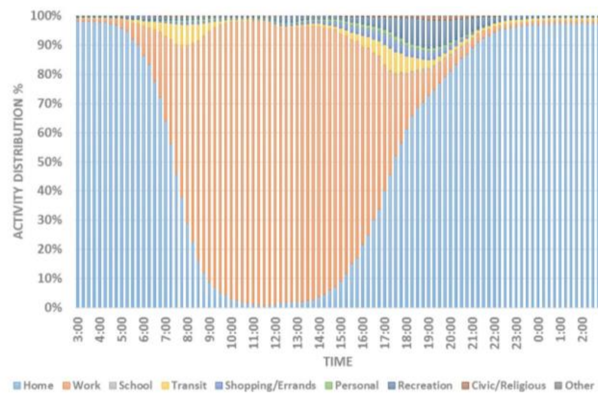

Figure 6 Activity patterns distribution in urban areas according to the Worker cluster

Such activity time series for individual clusters are translated into a BPS occupancy schedule by mixing clusters for different building types. Through such combinations occupancy patterns emerge that differ with change in building use. Figure 7 shows an example occupancy schedule for a single family house composed of a “Worker,” a “Stay at Home” and two “Students.”

Since this schedule focuses on “home” all other activities were disregarded, and the produced profile represents the basis for a residential building type schedule aggregated from a variety of clusters through their activity patterns. If one then assigns to which school and workplace three members of the household go, complementary building occupancy schedules for these buildings can also be generated.

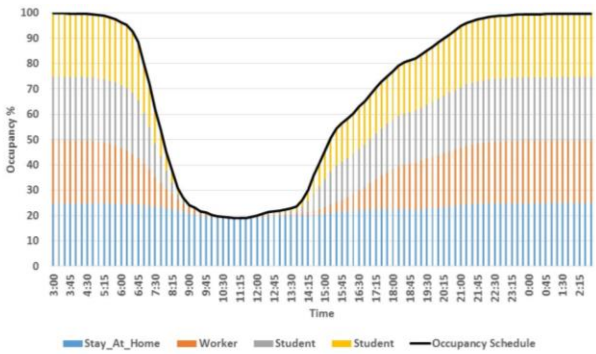

Figure 7 Single family house occupancy schedule as composed from three clusters and four occupants

DISCUSSION

This section will discuss the framework, potentials in generating activity-based occupancy schedules and will speculate on the use of complete behavior profiles for urban modeling.

Framework

The presented framework links two domains of knoweldge, BPS and activity-based transportation analysis. While the demonstrated workflow is not intended to replace any practice standards, it attempts to create common grounds for architects, planners and transportation engineers innovatively by utilizing activity-based travel surveys and databases to produce useful occupancy information for energy modelers. The field of urban modeling and simulation is currently growing, and that is why having an approach, such as the presented framework, that addresses people as both building occupants and as travelers is becoming essential.

Occupancy Schedules

Occupancy schedules are considered indispensable for whole energy use simulation and consequent potential for energy savings (Oldewurtel et al, 2013). Numerous models have been previously developed based on occupancy monitoring (Duarte et al, 2013), which is plausible when a building is erect, and simulation is being done for that structure specifically. However, the presented method aims to understand occupant behavior as part of a time-series of activities in the city. This makes the traditional occupancy schedule part of a bigger scope for urban simulations based on clusters of users. For example, An office building in an “innovation district” may have 50% workers and 50% adventurers. A more traditional office has a mix of 90% workers and 10% adventurers. By mixing the clusters for different spaces new occupancy patterns emerge, since simulation software don’t place emphasis on what type of users are in a space at a point in time, just the number of occupants.

Basis for Activity Centered Agent Based Modeling

Behavior profiles are an activity-based approach to visualize the type of activities clusters undergo temporally. This is a useful output for mobility-based simulations in urban modeling, as a previous challenge to this type of simulations was the unavailability of a workflow that addresses “trip chaining” in relevance to simulation on the building and neighborhood scale (Rakha et al, 2013). With the cluster-based profiles each activity is not considered a “trip” on its own, activities are now linked to each other temporally. However, it is presented in this framework as a deterministic output, while simulating occupant behavior and user activities in the city are better addressed using a stochastic process capable of producing probabilistic outcomes that are more in-line with real human behavior. (Wang, et al, 2010).

Figure 8 shows how this approach can be used as part of a Markov transition matrix, where each activity is linked to the previous not just temporally, but also probabilistically based on characteristics of the cluster and expected behavior. For example, as the work day comes to an end occupants leave their building with probabilities of various activities. While this is of little value to building scale simulations, as modelers are only interested in when occupants arrive and leave, for urban modeling this approach allows for realistic micro simulation of mobility behavior based on the analysis of real human activity patters. This sets the basis for a probabilistic workflow, where a stochastic process, such as Markov Chains, could be used as a generator for agent-based activities for urban mobility simulations.

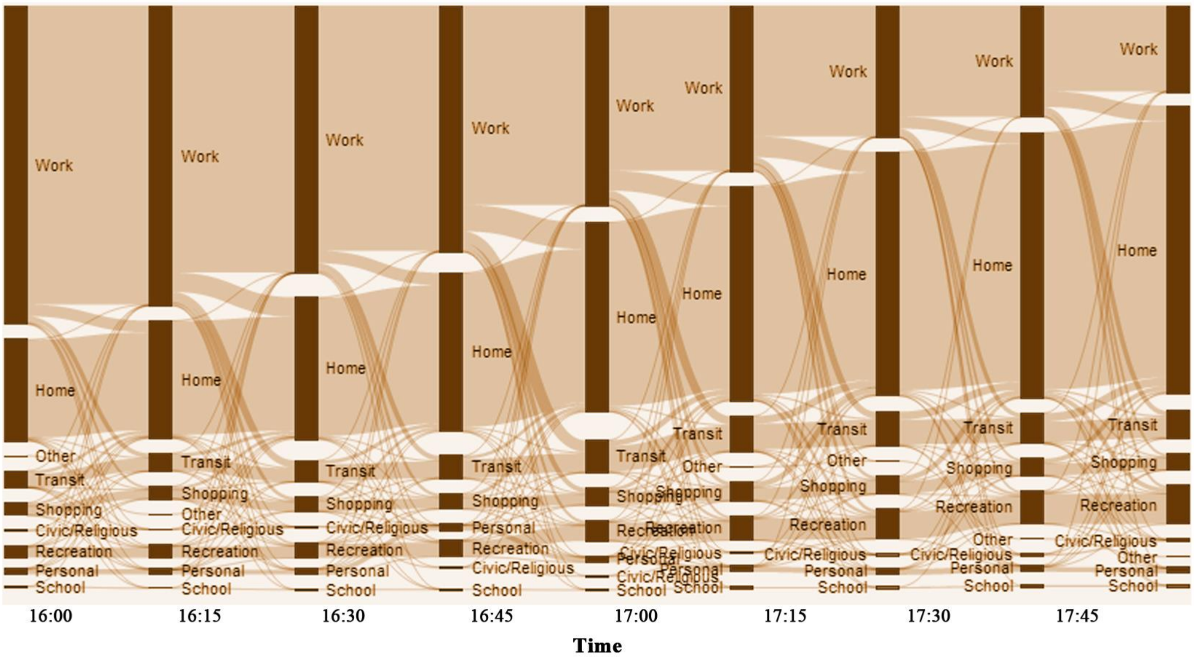

Figure 8 Sankey diagram for Worker cluster flow of activities at the end of the work day. At 4 PM Workers start to change their work activity gradually to being in transit to home or other activities. As 6 PM is reached, most workers have shifted their activity to being at home.

CONCLUSION

Urban performance simulation is an exciting interdisciplinary field on the rise. Current and future researchers and practitioners face numerous challenges, as urban landscapes change and transform rapidly. This motivates us to produce tools and workflows that bridge the gap between the various disciplines involved in the creation of sustainable and energy efficient built environments.

This paper presented a framework for the generation of occupancy profiles based on activity-based travel behavior analysis. The aim was to approach the issue of human behavior modeling for BPS based on patterns of activity in cities. While the technology needed to track real-time human behavior in order to design stochastic models for probabalstic simulation is under development, this workflow was explored through a traditional travel survey for Massachusetts as a proof of concept, and extracted results were shown to be useful for both BPS and urban modeling. Future research should address the shortcomings of using travel surveys such as weekday and weekend patterns, and progress in urban sensing technologies and its reflection on simulation workflows should be observed as a means to support such an interdisciplinary approach to urban modeling.

REFERENCES

Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira Jr, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. Transportation research part C: emerging technologies, 26, 301-313.

Duarte, C., Van Den Wymelenberg, K., & Rieger, C. (2013). Revealing occupancy patterns in an office building through the use of occupancy sensor data. Energy and Buildings, 67, 587-595.

Fabi, V., Andersen, R. V., Corgnati, S. P., & Olesen, B. W. 2013. A methodology for modelling energy-related human behaviour: Application to window opening behaviour in residential buildings. In Building Simulation (Vol. 6, No. 4, pp. 415-427). Springer Berlin Heidelberg.

Fritsch, R., Kohler, A., Nygård-Ferguson, M., & Scartezzini, J. L. (1990). A stochastic model of user behaviour regarding ventilation. Building and Environment, 25(2), 173-181.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. L. (2008). Understanding individual human mobility patterns. Nature, 453(7196), 779-782.

Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: views from urban planning. American journal of preventive medicine, 23(2), 64-73.

Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., & Tibshirani, R. (2009). The elements of statistical learning (Vol. 2, No. 1). New York: Springer.

Hoes, P., Hensen, J. L. M., Loomans, M. G. L. C., De Vries, B., & Bourgeois, D. (2009). User behavior in whole building simulation. Energy and Buildings, 41(3), 295-302.

Jiang, S., Ferreira, J., & González, M. C. (2012). Clustering daily patterns of human activities in the city. Data Mining and Knowledge Discovery, 25(3), 478-510.

Kitamura, R., Chen, C., Pendyala, R. M., & Narayanan, R. (2000). Micro-simulation of daily activity-travel patterns for travel demand forecasting. Transportation, 27(1), 25-51.

Maat, K., van Wee, B., & Stead, D. (2005). Land use and travel behaviour: expected effects from the perspective of utility theory and activity-based theories. Environment and Planning B: Planning and Design, 32(1), 33-46.

McCormack, E., Scott Rutherford, G., & Wilkinson, M. G. (2001). Travel impacts of mixed land use neighborhoods in Seattle, Washington. Transportation Research Record: Journal of the Transportation Research Board, 1780(1), 25-32.

Noulas, A., Scellato, S., Mascolo, C., & Pontil, M. (2011). An Empirical Study of Geographic User Activity Patterns in Foursquare. ICWSM, 11, 70-573.

Oldewurtel, F., Sturzenegger, D., & Morari, M. (2013). Importance of occupancy information for building climate control. Applied Energy, 101, 521-532.

Rakha, T. and Reinhart, C F. (2013). A carbon impact simulation-based framework for land use planning and non-motorized travel behavior interactions, in Proceedings of Building Simulation 2013, 25 -28 August, Chambery: France.

Ralambondrainy, H. (1995). A conceptual version of the *K*-means algorithm. Pattern Recognition Letters, 16(11), 1147–1157.

Reinhart, C F. (2004). Lightswitch-2002: a model for manual and automated control of electric lighting and blinds. Solar Energy, 77(1), 15-28.Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65.

Wang, C., Yan, D., & Jiang, Y. (2011). A novel approach for building occupancy simulation. In Building simulation (Vol. 4, No. 2, pp. 149-167). Tsinghua Press.

World Health Organization (WHO). "Urban population growth." Global Health Observatory. World Health Organization. Available: http://www. who. int/gho/urban_health/situation_trends/urban_population_growth_text/en/Accessed (2013).

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... & Steinberg, D. (2008). Top 10 algorithms in data mining. Knowledge and Information Systems, 14(1), 1-37.