Enhancing Traveler Context through Transferable Activity Patterns

Chad A. Williams[1], Abolfazl Mohammadian[2], Joshua Auld[3], and Sean T. Doherty[4]

[1] Department of Computer Science
Central Connecticut State University, New Britain CT, USA
[2] Dept. of Civil and Materials Engineering
University of Illinois at Chicago, Chicago IL, USA
[3] Argonne National Laboratory, Chicago IL, USA
[4] Dept. of Geography and Environmental Studies
Wilfrid Laurier University, Waterloo, Ontario, Canada

Abstract. Developing a model of the needs of a mobile traveler is critical to good personalization. Transportation planners have been modeling these needs for years, but these models have not been used to date due to two outstanding questions: 1) are they applicable to individual travelers 2) are they useful beyond the studied region. This study demonstrates these studies can directly enhance the model of mobile users, and be done in a practical way through the transference of activity patterns across cities. This work then demonstrates how these studies can be combined with patterns of an individual mobile user successfully.

Keywords: travel prediction, activity prediction, traveler context.

1 Introduction

With the increase in hand held computers and GPS devices, there has been an increasing demand for predicting an individual's future travel plans for devices such as smart traveler's assistants. Most existing studies of projecting the patterns of individuals have examined learning traveler behavior by concentrating on the travel itself or location attributes such as Krumm and Horvitz [1]. Other approaches have examined trying to determine the activities of the person based on location such as that by Liao et al. [2]. More recent work by Zheng et al. has tried to combine these approaches by using GPS traces for both activity and location recommendation [3]. These works and others have shown that a considerable amount of information can be inferred about a traveler based solely on analysis of their GPS trace. This study takes a different approach by instead looking at the decision factors and reasons behind the travel.

Examining the reason behind travel has been studied by transportation planners through detailed, time-consuming surveys conducted to capture not just travel, but the planning and decision processes that go into travel. Unfortunately these data collection methods are too burdensome on users to be practical for

D. Uhler, K. Mehta, and J.L. Wong (Eds.): MobiCase 2012, LNICST 110, pp. 412–423, 2013.

mobile applications. In this work we examine using information gathered in these data-rich surveys to enhance the model of mobile users with limited user input. Specifically the aim of this work is to use data that can be collected from a mobile user unobtrusively such as GPS analysis and augment that data without burdening the traveler.

In this work, we address two questions. First, can travel surveys be used to enhance the patterns of mobile users in locations other than the study area? While it is somewhat intuitive that this might work if the survey area is the same as that of the traveler; being restricted to the areas where a survey was conducted would greatly reduce the applicability of this approach. To address this question this study examines how well the patterns of one city might be used in a different metropolitan area. Second, we examine how to augment the observed travel characteristics with patterns observed in these surveys.

One of the goals of this study is to use an understandable model such that the factors influencing a particular outcome can be readily understood. The reason for this criterion is that an easily understood model is likely to make it easier to create personalization that could influence the actions of the traveler. Prior work has shown that sequential associative rules are particularly effective for traveler context prediction using survey data [4]. However for the problem described above there is significantly more missing data compared to the survey prediction problem. As a result, the second aspect we examine is a comparison of traditional sequential rules with attribute constrained rules (ACR) which were specifically created to address sequential mining with missing data [5]. As this work will empirically demonstrate, ACRs can be used to learn the patterns of an individual and augment them effectively with general patterns outside of the original survey area despite large amounts of missing data.

Below we begin by introducing the problem and motivation for this study, followed by a formal description of the problem and the methods used. Next, an empirical evaluation of transference and the appropriateness of the ACR learning method are demonstrated. Finally, a discussion of the impact of this work and future areas of study are presented.

2 Background and Motivation

Transportation planners have studied travel behavior extensively over the years. More recently, their focus has shifted from looking at travel alone to understanding why a trip is made and when this decision was made [6]. One of the approaches used for this has been examining the activity needs of the person as the reason the travel is made. This is important because by understanding the activity priorities of travelers, not only can a model be built to project what activities a person is likely to do over a fixed period of time; a model of what influences their behavior can also be obtained. We propose this type of information is more important than travel and location information alone in understanding what information is useful to the traveler before they *plan* their trip.

The inputs for these models are collected through activity-based travel surveys. These surveys consist of a person meticulously recording all of their trips, important locations, activities at every stop, and reasons these were made. Some go a step further and have participants record planning information for when an activity, location, and schedule were determined. Finally, planners take the information gathered from the surveys to build activity models which are then used to project travel patterns for people in the area of study. In creating these models, a key assumption is made that activity models can represent the basic patterns of people within a similar area. The validity of this approach has been demonstrated when applied to the **same** metropolitan area [7].

While this type of approach has been verified within the same metropolitan area at an aggregate level, the ability to use activity patterns within one city to project the activity patterns in another metropolitan area at an individual level has been largely unexplored. We assert that since activity needs are the factor driving travel, similar activity patterns will be seen regardless of the environment. Thus, we propose patterns from one metropolitan area can be transferred to other metropolitan areas. This is important because without this transferability, only cities where an activity survey has been conducted would benefit from this information. While many cities do conduct some form of activity surveys, the information collected can vary greatly making it impractical to apply location specific studies to every mobile user. If patterns from one city could be shown to be transferable to other cities, this would mean a general model could be applied to wherever a user might be.

The basic concept behind activity-based analysis is that all travel is driven by the activity needs of the individual. Research has shown the planning of these activities happens in a fluid manner based on both personal flexibility and activity flexibility, making these aspects critical for planning decisions. As a result, the goal of traveler prediction is not activity or location prediction, but traveler context prediction which includes these flexibility and planning factors.

The contribution of this paper is demonstrating that the micro patterns of one city can be used to augment the observations of travelers within a different city for more complete modeling of the trip characteristics of an individual traveler. In the past the amount of data entry collected from the user to accomplish this would not be practical; but with GPS enabled smart phones a significant amount of detail about the trip can be captured passively making the task of inferring additional information much more tractable. While this work does not explore if this relationship holds across significant cultural differences, it does show that these patterns can be transferred across varying transportation networks and urban structures within a selection of cities within the United States and Canada.

3 Partially Labeled Sequence Completion

In our analysis, we break the trips into discrete steps of travel and activities. For each of these discrete steps there are a number of different characteristics associated with each segment such as trip time, mode of travel, activity, and location.

Here we address adding to those characteristics information such as personal and activity flexibility information, as well as when various aspects of these were decided. When put together these form a sequence of sets of activity and trip characteristics describing that user's travel.

Algorithms for rule mining of sequential patterns have been a major source of interest since they were first introduced in Agrawal and Srikant [8]. The problem we examine is given a sequence where there are a known set of attributes that describe an event within the sequence, infer any missing values of the attributes for a target set. This problem addresses the situation of knowing a set of attributes through methods such as a GPS trace analysis, but needing to infer missing flexibility information about a traveler.

Problem Statement. For partially labeled sequence completion, let

$$H = \{H_1, H_2, \cdots, H_n\}$$

be a database of sequences, and let:

$$H_i =< S_1, S_2, \cdots, S_n >$$

be the sequence of sets of observations in a sequence i; where each observation set S_j is composed of 1 to m attributes $\{a_{j,1}, a_{j,2}, \cdots, a_{j,m}\}$. Each attribute $a_{j,k}$ has a discrete set of values for the k^{th} position that is shared across all observation sets S. Intuitively the sequence H_i can be thought of as a series of traveler contexts (sets) with discrete measures (the attributes), where at each event j all measurements are relevant, but only a portion of these measures may actually be recorded in the set S_j. Given a sequence H_{target} of length l and a target set S_t to be completed where $1 \leq t \leq l$ and between 1 to m arbitrary attributes are missing values. Determine the values of all missing attributes $\{a_{t,1}, a_{t,2}, \cdots, a_{t,m}\}$ in S_t. Thus our goal is to use the surrounding sequence information in H_{target} to populate any missing values to complete the set S_t. In other words, given some information that can be determined about a traveler from GPS analysis, can we infer the missing traveler context information.

3.1 Sequential Set and Attribute Constrained Rules

In this section we discuss how the associative sequence mining and prediction problem relates to the traveler context prediction problem. This is followed by an illustrative example of how ACRs introduced in Williams et al. are better suited for this problem since missing data is an issue [5]. This approach stems from the general apriori associative sequence prediction problem, but is a specialized form of it more suited to the specific traits of this problem. A comparison of the results of the two will be discussed in the results section.

One characteristic of the problem being examined is that unlike the general form of sequential set prediction, the travel context prediction problem is constrained to always contain the same set of attributes describing the travel. This sub-problem of sequences of sets referred to as partially labeled sequences, applies to sequences that have sets of a fixed group of attributes where some attribute values (labels) may be unknown. Due to this structure it is also possible to tell which attributes are missing rather than being a negative example that the attribute does not have the specified value. For this sub-problem ACRs were introduced to use these types of constraints to better estimate the benefit of a sequential rule when missing data is an issue [5]. Specifically ACRs work by identifying templates of sequential patterns to mine rules that specifically target completions of particular attribute values when missing values are an issue.

3.2 Illustrative Example

Here we provide an illustrative example to demonstrate the benefit of ACRs when missing data is an issue. Throughout these examples refer to Table 1 as the sequence database. Below we use the standard definitions of *support* and *confidence* defined as: The **support** of the sequential rule $X \rightarrow Y$ is the fraction of sequences in the database that contain Y. The **confidence** of a sequential rule $X \rightarrow Y$ is the fraction of sequences in the database that contain X that also contain Y.

Table 1. Example sequence database

H_1	$< \{a_1\}\{b_1\}\{a_2, b_2\} >$
H_2	$< \{a_1\}\{a_2, b_1\}\{a_2, b_2\} >$
H_3	$< \{a_1\}\{b_1\}\{a_2\}\{b_2, c_2\} >$
H_4	$< \{a_1\}\{a_2, c_1\}\{b_1\} >$

For an example of how constrained rules can better represent the applicable confidence, consider the following scenario: $H_{target} = \langle\{a_1, b_1\}\,\{b_1\}\,\{a_2,\ \mathbf{?}\}\rangle$, where the last set, $S_3 = \{a_2,\ \mathbf{?}\}$, is the target set and we are interested in completing the set with the value of attribute b. The following traditional sequence associative rule would be applicable:

$$\langle\{a_1\}\,\{b_1\}\,\{a_2\}\rangle \rightarrow \langle\{a_1\}\,\{b_1\}\,\{a_2, \mathrm{b}_2\}\rangle$$
$$[\mathrm{sup} = 2/4, \mathrm{conf} = 2/3]$$

Where S_3 can be completed $\{a_2, b_2\}$ with a confidence of 2/3.

With ACR this idea is extended to constrain pattern matches to particular attribute values of interest. In our example, since we are specifically interested in the value of attribute b, the ACR version of the same rule would be:

$$\langle\{a_1\}\,\{b_1\}\,\{a_2, \mathbf{b{:}*}\}\rangle \rightarrow \langle\{a_1\}\,\{b_1\}\,\{a_2, \boldsymbol{b}_2\}\rangle$$
$$[\mathrm{sup} = 2/4, \mathbf{conf} = \mathbf{2/2}]$$

where $\{a_2, \mathbf{b{:}*}\}$ is a template such that matching examples in the history must contain a_2 and **some** value for b, in other words the attribute value can't be missing. This results in the calculation of rule confidence only considering sequences that could either support or negate the consequence rather than being counted as a negative example because a value is missing. Thus, it is able to discount sequence H_3 included by sequential rules as it does not match the attribute constrained pattern. This advantage in accurately evaluating the value of the constrained sequence rule is the reason we examine ACR.

4 Evaluation

The basic idea behind activity pattern transferability is that the patterns from one metropolitan area can be used to predict the patterns for another area. The main purpose of these experiments is to compare how survey data from one city can be used to predict the activity behavior in a completely different city. To evaluate this hypothesis, a series of experiments was conducted to examine multiple aspects of the transferability of activity patterns across a selection of cities of various sizes and locations. First, we demonstrate that the patterns of one city can be transferred to another city with reasonable performance as compared to using the patterns of the original city. Next, we examine using this idea to improve learning the patterns of an individual outside of where the patterns were collected to demonstrate that this idea can be applied in general to new locations. The details of the experiment design and results are given below.

4.1 Data Selection

For evaluating this problem we chose a selection of sets of activity and planning data that were collected by transportation planning activity based surveys. The reason for this choice was in part due to the proven record of this type of data being effective at modeling traveler behavior [7]. The second reason these activity surveys were chosen was because they focus on a mix of fields that require active collection as well as fields that may be determined via passive means.

The differentiating factor between these two types of data is that *actively* collected data requires a user to enter the information to be determined. For example **why** they chose a particular location. *Passively* collected data, on the other hand, is data that may be determined without user interaction. An example of this would be the user's location since this can be determined via the GPS device. The importance of selecting surveys that combine these two factors is that while some of the actively collected data would be very useful for personalization it would likely be too burdensome to ask a mobile user repeatedly. However as this study will show by combining this data with passive means it is possible to match the patterns of the passive data and infer the active data. This would thus allow a mobile application to infer a preference profile of a user that would otherwise be impractical to collect due to the burden on the user.

Data. To make an evaluation of how well activity patterns transfer across cities, activity surveys from a variety of areas were selected. Specifically the data from the Toronto CHASE survey [9], the 2001 Atlanta Household Travel Survey [10], the 2002 Anchorage Household Travel Survey [11], and the Chicago UTRACS survey [12] were used. These data sources were selected as they provide a good mix of metropolitan populations and differences in urban structure.

For this study the following fields were compared: location, activity type (22 categories), mode, observed sequence of the activity, arrival time, departure time, duration, individuals involved (related to involvement of others in trip), whether the activity was mandatory, when planned, when decided on the location, when decided who would participate, when decided timing of activity, when travel mode selected, why travel mode selected, why route was chosen, duration flexibility, time flexibility, spatial flexibility, and when the activity was planned.

Data Summary. Table 2 summarizes the different characteristics of the data sources listed above. As this table shows, the more time consuming planning surveys (CHASE and UTRACS) tend to be considerably smaller than the basic activity surveys, but they also offer longer survey periods providing better examples of multi-day activity patterns.

Table 2. Data set characteristics

Data set	Activity Information	Planning Information	GPS Trace	Time Period	Households
CHASE 2002-03	X	X		7 days	271
Atlanta 2001-02	X			2 days	8,069
Anchorage 2002	X			1 day	1,293
UTRACS 2009-10	X	X	X	14 days	100

4.2 Methods

Within these experiments, the primary goal was to determine if activity patterns from one region can be used in a different region compared to a model built on the same city. The approach for verifying this transferability is similar to that used in Arentze et al. where the data of one city is used as training data for a second city and the results of that test being compared to the results of a city being trained on its own data [13]. For their study the evaluation was made at an aggregate level to compare distributions, by contrast here we analyze the transferability of patterns at an individual level.

For these experiments the results of training and testing for the same city were performed using a ten times cross-folding methodology. The cross-folding methodology was executed such that each fold had the same number of travelers although the number of activities per fold could be different depending on the number of activities for the users selected for a particular fold. For all other tests of training based on one city and testing on a second city were carried out by using the entire survey set of the training and target cities.

The method for executing the tests was to train the classifier and then for the test data set evaluate each traveler's sequence separately. The training data was built by taking each user sequence in the training set and splitting the sequence by home-based tour. To capture a new tour's start being dependent on the activity last completed before the current home-based tour began, each tour after the first would start with the last two elements of the previous sequence. Thus, each training user's sequence was split such that each new sequence would begin with an out of home activity, followed by a home-based activity, followed by all activities up to and including the next home-based activity. This method was used to break up long sequences such as those in the multi-day CHASE and UTRACS data where an activity sequence could contain over 80 activity sets.

The evaluation was done so that in evaluating the test user, the activity sequence of the user was evaluated in a stepwise fashion. In other words, for the first prediction of the user there was no history to base the prediction on, but for the second prediction there would now be one set of history in the sequence. Similar to the training approach, the user's history used in testing was the current home-based tour and the previous home-based tour. This was continued through the end of the sequence where each prediction of the n^{th} step of the tour was based on $n-1$ sets in the user's current tour plus the previous home-based tour. This process was repeated for each user in the test set and the average of all of these predictions is presented.

One of the challenges of learning the activity patterns of these data sets is that all four of these contain numerous missing values from survey participant's not answering all questions. This is important as it is similar to only collecting data that can be derived from passive data such as GPS traces, yet attempting to resolve data only observed by active data entry such as that in the survey data. To verify the benefits of the ACR technique in this environment, the experiments were conducted with both standard apriori sequential mining and ACR. The results showed that ACR outperformed standard sequential rules across all experiments in terms of both precision and recall. The results of standard sequential rules are not shown here for brevity.

4.3 Experimental Evaluation

As described above, the problem of partially labeled set completion involves taking a sequence and trying to fill in or predict items within a single target set within a sequence. Since the problem of partially labeled set completion can take the form of predicting anywhere from a single item in a target set to all items in the target set, the results below reflect the average of all possible combinations of the target pattern in all possible positions for the target set. Where target pattern means: the set of attribute values in the target set that are being evaluated. Thus in the experiments below, for the target set any attribute value that is not specifically of interest as specified by the target pattern retains its original

Table 3. Activity pattern transferability

F-measure

		Test			
		UTRACS	Atlanta	Anchorage	CHASE
Training	UTRACS	**0.461126**	0.447509	0.427923	0.575487
	Atlanta	0.333476	**0.443616**	0.441377	0.60563
	Anchorage	0.41161	0.461426	**0.546549**	0.57918
	CHASE	0.276572	0.417372	0.426867	**0.625694**

value for determining matching rules. For example if the target pattern included attributes a and c ($S_T = \{a_T c_T\}$). In testing the sequence:

$$\langle\{a_1 b_2 c_1\}\,\{a_2 b_2 c_2\}\,\{a_1 b_1 c_2\}\rangle$$

If the target set was S_3 for the sequence, the test sequence would thus be:

$$H_{target} = \langle\{a_1 b_2 c_1\}\,\{a_2 b_2 c_2\}\,\{a_T b_1 c_T\}\rangle$$

In making the predictions, once all rules were identified the ranking scheme discussed in Williams et al. was applied [5]. Specifically, the rules were ranked in order by confidence, number of target productions, support, and antecedent length. The productions of the top ranked rule were then applied to the target set, the remaining matching rules were then re-ranked by finding the next highest rule whose consequent did not contradict the values previously completed. This was repeated until either all rules had been exhausted or all fields in the set being tested were completed. The values shown in the tables are based on selecting the parameters of support and confidence thresholds that yielded the maximum F-measure. The precision and recall metrics discussed are based on this configuration for the maximum F-measure.

The experiments were executed so that every city was treated as a training and test set for all other cities. Table 3 contains the F-measure results of these tests. The rows represent how well the survey of a particular city was at predicting the patterns in other cities. Likewise the columns represent how well various cities performed at predicting the patterns of that particular city. Across the diagonal marked in **bold** are the results of training and testing on the same city.

As these results in Table 3 demonstrate, it is possible in nearly all of these cases to find a secondary city which can be used to train activity patterns for the first city nearly as well as the city itself. For Atlanta there appear to be two cities that can actually slightly improve the performance over using Atlanta alone, however the difference is not statistically significant. The one exception to this was Anchorage for which there was a drop off in transferability. As will be discussed further in the recall results, part of the reason for this, discovered in further data analysis, is that the Anchorage study captured a more diverse set of observations.

In examining the precision and recall results (not shown), it was interesting to note that the UTRACS survey showed the highest precision across all other test cities. However it also had the worst recall. This is likely due to the lengthier study being better able to capture the underlying patterns, but the small size of the study hurting the number of patterns observed. This assertion about why the recall is low is further supported by the CHASE data set which also suffered from lower recall and it is also one of the smaller studies. An interesting observation from these results is that the CHASE data set has the poorest transferability. This could warrant further study to better determine if this is potentially related to differences between preferences in the United States and Canada. Another interesting result is when the model was built from the Anchorage dataset, it resulted in the highest recall across all datasets. As noted above this is likely due to this dataset having the highest diversity in observations (activity sequences, flexibility differences), while also having a relatively large mix of missing data across all attributes. Thus when the missing data is accounted for, the diversity of the observations make it well suited for covering a wide variety of environments.

As shown above, while generally transferring patterns from one city to another produced good results, the selection of the training data could significantly impact overall performance as well as the precision versus recall mix. Even so, the implication of this is that even if a user is in a city that other histories are not available, survey data from a second city could be used instead yielding good results despite the lack of city specific data.

4.4 Combined Knowledge

In the next set of experiments a comparison was made of using a model built just on a user's individual history against one based on just the general patterns as well as a hybrid model that combines the two approaches. Since one of the goals of this work is to demonstrate this approach can be used in areas outside of these survey areas; this experiment was conducted using the Anchorage data set for the general patterns and the test set was the UTRACS data set. The UTRACS data was selected as it is the longest survey thus allowing the cumulative learning to be more apparent. The points charted are the average score of predicting all fields of the activity context for the day averaged across all users for that day

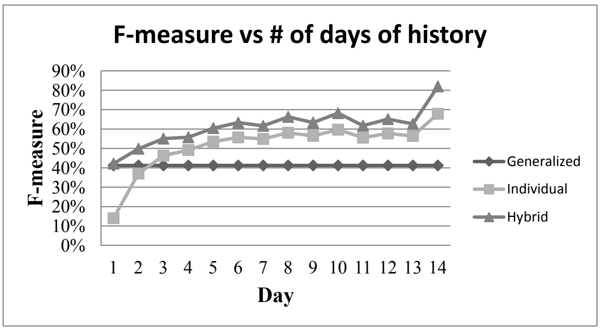

Fig. 1. Comparison of F-measure vs. patterns used

against their individual histories. In other words the points at day 3 represent the average across all users for just the patterns observed on day 3.

Using the approach described, 3 results were recorded. Figure 1 shows the comparison of the F-measure of these three different approaches day by day for the 14 day period of the UTRACS survey. First, the accuracy of predicting the individual histories using only the generalized patterns from Anchorage were recorded which is depicted as the *generalized* patterns in Figure 1. The line shown for the generalized patterns is the average across all of the days. The second, is the accuracy of predicting the patterns of the individual based solely on their own history without any generalized patterns, which is depicted as *individual* patterns. As the results show, the generalized patterns initially perform better before enough individual data is collected, which occurred at day 3. After day 3 the model based solely on the individual data performed better.

The final results are based on a combination of these two data sets depicted as *hybrid* patterns. The intuition behind creating the hybrid is to create a model of the individual augmented with general patterns to increase recall of patterns not previously seen in the user's history. The hybrid was created by using two separate training sets: the generalized patterns, and the patterns of the individual. For making predictions the predictions would be generated as complete as possible using only the individualized patterns. Once the individual patterns/rules were exhausted the generalized patterns were then used to fill in any remaining gaps. As the results show, when only that limited amount of information is provided it is difficult to learn a model of that individual quickly. This problem referred to as the "cold start" problem occurs due to limited training examples being available until enough history is gathered. As the results demonstrate not only does the hybrid benefit from the initial general patterns to avoid a cold start, it also immediately benefits from the individual patterns. By examining the results we see that the hybrid offers better performance than either method alone throughout the 14 day study period.

5 Discussion

In this paper we have examined the idea of using transportation planning activity survey data to augment personal history for a richer model of a mobile user. The goal of this work is to allow the limited data that can be collected from a mobile user through passive means such as a GPS trace to be enhanced with features such as constraints, planning and flexibility information that cannot be collected practically through unobtrusive methods. This idea of traveler context that is made up of not just current and next step activities but accessibility options, timing constraints, and scheduling behavior represents a significant increase in the likely behavior of a traveler without adding an additional burden on the mobile user. An additional benefit of the ACR technique is that the predictive model is much more transparent than techniques such as neural networks, allowing additional insight to be gained from the models created.

The goals of the majority of studies in this area have been focused on modeling what the traveler is doing or going to be doing. We propose that the end goal of the research in this area should be to build a more in depth model which includes additional detail such as planning and timing flexibility. Capturing/inferring this information provides great opportunities for personalizing travel applications before events are *planned* rather than just before the trip is made.

References

1. Krumm, J., Horvitz, E.: Predestination: Inferring Destinations from Partial Trajectories. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 243–260. Springer, Heidelberg (2006)
2. Liao, L., Fox, D., Kautz, H.: Location-based activity recognition using relational Markov networks. In: Proceedings of the International Joint Conference on Artificial Intelligence, IJCAI (2005)
3. Zheng, V.W., Zheng, Y., Xie, X., Yang, Q.: Collaborative location and activity recommendations with GPS history data. In: Proceedings of the 19th International Conference on World Wide Web (WWW 2010), pp. 1029–1038. ACM (2010)
4. Williams, C.A., Mohammadian, A., Nelson, P.C., Doherty, S.T.: Mining Sequential Association Rules for Traveler Context Prediction. In: Proceedings of the First International Workshop on Computational Transportation Science held at The International Conference on Mobile and Ubiquitous Systems: Networks and Services, MOBIQUITOUS 2008 (2008)
5. Williams, C.A., Nelson, P.C., Mohammadian, A(K.): Attribute Constrained Rules for Partially Labeled Sequence Completion. In: Perner, P. (ed.) ICDM 2009. LNCS, vol. 5633, pp. 338–352. Springer, Heidelberg (2009)
6. Timmermans, H.: Progress in Activity-Based Analysis. Elsevier (2005)
7. Doherty, S.T., Mohammadian, A.: The Validity of Using Activity Type to Structure Tour-based Scheduling Models. In: Proc. of 86th Annual Meeting of the Transportation Research Board, Washington D.C. (January 2007)
8. Agrawal, R., Srikant, R.: Mining sequential patterns. In: Yu, P.S., Chen, A.S.P. (eds.) Eleventh International Conference on Data Engineering, Taipei, Taiwan, pp. 3–14. IEEE Computer Society Press (1995)
9. Doherty, S., Nemeth, E., Roorda, M., Miller, E.: Design and Assessment of the Toronto Area Computerized Household Activity Scheduling Survey. Journal of the Transportation Research Board 1894, 140–149 (2004)
10. NuStats: 2001 Atlanta Household Travel Survey: Final Report. Technical report, Atlanta Regional Commision (April 2003)
11. NuStats: 2002 Anchorage Household Travel Survey Technical Report. Technical report, Municipality of Anchorage (September 2002)
12. Frignani, M.Z., Auld, J., Mohammadian, A., Williams, C., Nelson, P.: Urban Travel Route and Activity Choice Survey (UTRACS): An internet-based prompted recall activity travel survey using GPS data. In: Proceedings of 89th Annual Meeting of the Transportation Research Board, Washington D.C. (January 2010)
13. Arentze, T., Hofman, F., van Mourik, H., Timmermans, H.: Spatial Transferability of the Albatross Model System: Empirical Evidence from Two Case Studies. Transportation Research Record: Journal of the Transportation Research Board 1805, 1–7 (2002)