Optimizing copious activity type classes based on classification accuracy and entropy retention

Wim Ectors[a,*], Sofie Reumers[a], Won Do Lee[b], Bruno Kochan[a], Davy Janssens[a], Tom Bellemans[a], Geert Wets[a]

[a] *UHasselt - Hasselt University, Transportation Research Institute (IMOB), Agoralaan, 3590 Diepenbeek, Belgium*
[b] *Manchester Metropolitan University, Crime and Well-being Big Data Centre, All saints, M15 6BH Manchester, England*

Abstract

Despite the advantages, big transport data are characterized by a considerable disadvantage as well. Personal and activity-travel information are often lacking, making it necessary to deduce this information with data mining techniques.

However, some studies predict many unique activity type classes (ATCs), while others merge multiple activity types into larger ATCs. This action enhances the activity inference estimation, but destroys important activity information. Previous studies do not provide a strong justification for this practice. An objectively optimized set of ATCs, balancing model prediction accuracy and preserving activity information from the original data, becomes essential.

Previous research developed a classification methodology in which the optimal set of ATCs was identified by analyzing all possible ATC combinations. However, this approach is practically impossible in a finite amount of time for e.g. the US National Household Travel Survey (NHTS) 2009 data set, which comprises 36 ATCs (home activity excluded), since there would be $3.82 \cdot 10^{30}$ unique combinations (an exponential increase).

The aim of this paper is to optimize which original ATCs should be grouped into a new class, and this for data sets for which it is impossible or impractical to simply calculate all ATC combinations. The proposed method defines an

*Corresponding author. Tel.: +32-11-269114.
Email address: `wim.ectors@uhasselt.be` (Wim Ectors)

optimization parameter U (based on classification accuracy and information retention) which is maximized in an iterative local search algorithm. The optimal set of ATCs for the NHTS 2009 data set was determined. A comparison finds that this optimum is considerably better than many expert opinion activity type classification systems. Convergence was confirmed and large performance gains were found.
Keywords: Activity type classification, (Big) transport data annotation, optimal set of activity types, local search algorithm, classification accuracy, entropy indices

1. Introduction

These days, big data sets are collected continuously and in real time, making large amounts of data that are temporally and spatially referenced available to researchers [1]. Due to the availability of spatio-temporal information, big transport data are very effective in exploring individual mobility patterns. Despite the advantages, however, big transport data are characterized by a considerable disadvantage as well: personal and activity-travel information are often lacking [2], making it necessary to deduce this information from the available travel patterns.

In order to overcome this shortcoming, behavioral data mining techniques are frequently used to infer activity types (sometimes otherwise denoted as trip or travel purposes, activity classes, activity categories or activity encoding) from behavioral attributes, such as temporal attributes and spatial information (e.g. [3, 4, 5]). However, in these researches different classifications of activity types exist. Some studies infer many activity classes, while others aggregate or group several activity types, limiting the number of activity type classes (ATCs) [6]. As argued in [6], in none of such studies a strong justification is established. The activity type classification in the majority of researches merely relies on the travel survey design, due to a lack of clear standards for ATCs which are grounded by a theoretical background [7].

The ATCs (and the size of this set of classes) strongly affect the classification accuracy. Often, activity types are aggregated (or *grouped*) in order to enhance the activity inference estimation. However, by aggregating activity types, and thus enhancing the activity inference estimation, important activity information is lost. Therefore, the need for a *standardized method* for activity categorization arises. An optimal set of activity types is an essential prerequisite for a robust and sound transport data annotation in a particular study area.

Previous research [6] developed a classification methodology using a rule-based heuristic algorithm in which the optimal grouping of ATCs was identified. This methodology is an objective alternative to the subjective choice of ATCs based on intuition or expert-opinion. The optimization method searches for an optimal balance between improving model accuracy and preserving activity information from the original data set. The method was applied to two household travel surveys (HTSs), i.e. the Seoul HTS and the Flanders (Belgium) HTS called OVG.

The optimization method in [6], however, might not be appropriate when the initial data set contains too many unique ATCs. The optimization strategy comprises three stages, where in the first stage all possible combinations of ATCs are generated. Considering 10 distinct activity types in the OVG and Seoul HTS, this brute-force approach calculated approximately 117,000 unique sets of combinations of classes. However, for the US National Household Travel Survey (NHTS) 2009 data set [8] which comprises 36 ATCs (home activity excluded, see section 3), calculating classifiers for all possible grouping combinations is impossible since the increase in distinct combinations is exponential. In other words, a large number of initial activity types which are considered for aggregation will result in an extremely large set of grouping combinations that needs to be processed as shown in Table 1. Subsection 4.1 will discuss how the number of possible ATCs in Table 1 can be calculated. The computation time of the second stage of the optimization method would rise up to $1.13 \cdot 10^{23}$ years for the US NHTS data set using the same setup as in [6]. Note that the age of the universe is only $13.8 \cdot 10^{9}$ years [9]. Because of this reason, the earlier proposed

Table 1: Number of possible activity type class (ATC) grouping combinations N as a function of the number of activity types n

n	N	n	N	n	N	n	N
1	1	11	6.786E+05	21	4.749E+14	31	1.029E+25
2	2	12	4.214E+06	22	4.507E+15	32	1.281E+26
3	5	13	2.764E+07	23	4.415E+16	33	1.630E+27
4	15	14	1.909E+08	24	4.460E+17	34	2.120E+28
5	52	15	1.383E+09	25	4.639E+18	35	2.816E+29
6	203	16	1.048E+10	26	4.963E+19	36	3.820E+30
7	877	17	8.286E+10	27	5.457E+20	37	5.287E+31
8	4,140	18	6.821E+11	28	6.161E+21	38	7.463E+32
9	21,147	19	5.833E+12	29	7.134E+22	39	1.074E+34
10	115,975	20	5.172E+13	30	8.467E+23	40	1.575E+35

method cannot be used for cases with copious activity types.

Considering that the US NHTS is not the only data set that includes a large number of activity types, this computation time issue will also surface for other travel data sets. In the UK HTS data set [10], for example, 22 distinct activity types are employed. In Table 2, several travel data sets are listed, together with the number of activity types that are considered in each case. In most cases the number of activity types is much larger than 10, preventing the use of the method of [6].

To overcome this process time issue, the research in this paper proposes an update of the optimization methodology using a 'local search' algorithm. The local search algorithm starts from a predefined ATC grouping combination and iteratively tries to optimize this group by applying random changes, hereby reducing the required computation process time. Subsequently, the algorithm is used to determine an optimal set of ATCs for the US NHTS as, to the knowledge of the authors, this HTS has the most copious activity type variable.

The remainder of this paper is structured as follows. Next section provides

Table 2: Examples of household travel survey data sets with their number of distinct activity type classes (ATCs)

Data set	**Country (or region) of origin**	**Number of person days surveyed**	**Number of ATCs***
AUS VISTA 2007 & 2009 [11, 12]	Australia	67,060	12
BEL Beldam 2010 [13]	Belgium	11,279	11
BEL OVG 3.0-4.5 [14]	Belgium (Flanders)	13,522	10
CHE Thurgau 2003 [15]	Switzerland	8,522	25
DEU Mobidrive 1999 [16]	Germany	13,244	22
FIN HLT 2010-2011 [17]	Finland	10,137	19
FRA ENTD 2008 [18]	France	17,996	31
GBR NTS 2009-2014 [10]	United Kingdom	551,234	22
IRL NTS 2009 [19]	Ireland	5,023	9
KOR Seoul HTS 2010 [20, 21]	Republic of Korea	219,269	10
NLD OViN 2013 [22]	The Netherlands	34,710	13
SVN Ljubljana 2013 [23]	Slovenia	3,426	12
SWE RVU 2011-2014 [24]	Sweden	31,457	25
USA NHTS 2009 [8]	United States of America	257,586	36

* 'Home' activity excluded, see also section 3

the results of a literature review. It is followed by a description of the data and afterwards the methodology. Subsequently, the results of the convergence of the local search algorithm are presented, followed by a presentation of the optimal ATCs for annotation. Finally, a conclusion is formulated.

2. Literature review

Advancements in information and communication technologies (ICT) and the improvement of location-aware technologies facilitate the collection of transport data, e.g. daily trajectories. The new transport data-collection methods support researchers with refined, detailed data sets of real-time data. Social media can be a rich source for the understanding of urban activity patterns [25]. These large collections of spatio-temporal information offer research opportunities, i.e. they enable a better investigation and understanding of human travel behavior.

Wolf et al. [26] presented a successful proof of concept with their early work on the trip purpose detection from GPS logs combined with land use information. A significant number of contributions followed in this domain of activity type inference, e.g. [27, 28, 29, 30, 31], yet none discussed a thorough justification for a particular set of inferred ATCs. However, some activity inference research recognized in their studies that there exists a need to determine an optimal set of activity type categories [32], or that ATC inference accuracy could (inappropriately) be increased by reducing the number of predicted ATCs [33]. To the authors' best knowledge, these concerns were not comprehensively addressed in literature yet [6].

The problem is augmented by the fact that each HTS data set is different, and considers different sets of ATCs. This observation shows that there is no uniformity in the way ATCs are selected for a study [7]. Table 3 illustrates the rather problematic differences in ATCs as how they might occur in different HTSs (here for the Seoul HTS and the US NHTS which are both used in this research). Each ATC may be defined according to some set of constraints, which allow to find corresponding ATCs in the different HTSs (based on the class definitions used in these HTSs). One notices the different levels of refinement of ATCs, the absence of categories in some HTSs, different definitions etc. For example, 'work' in the Seoul HTS corresponds to 'Go to work' in the NHTS, whilst such a definition would be interpreted as the act of traveling itself in the Seoul HTS (and for which it has different ATCs).

Previous research by the authors [6] developed a classification methodology in which the optimal set of ATCs was identified by analyzing all possible ATC combinations. The optimization strategy comprises three stages, where in the first stage all possible combinations of ATCs are generated. This brute-force approach calculated approximately 117,000 unique sets of combinations of classes for both HTSs that were used in the study (both HTS data sets considered 10 distinct activity types, after removing the 'Being at home' activity). In the second stage of the optimization strategy, classifiers are trained and tested on the data that were transformed according to the ATC combinations of the first stage. Finally, the optimal set of ATCs is defined in the third stage of the optimization method. On a server equipped with two intel Xeon EQ-2643 v2 processors (running at approximately 80% capacity, i.e. 20 threads) estimating 117,000 classifiers took roughly 30 hours of computation time.

As detailed in the introduction, running the above algorithm for the 36 ATCs in the US NHTS would be impossible given the exponential increase of combinations in stage one. Therefore, the proposed method of the current study will define an optimization parameter U (based on classification accuracy and information retention) which is maximized in an *iterative local search* algorithm.

In recent studies on activity-travel data mining, different inference techniques and data sources are investigated. The methods used can be classified into probability and rule-based heuristic approaches. In studies regarding the former, the naïve Bayes classifier is usually adopted to generate the probability of an alternative (e.g. [37, 38]), while the rule-based heuristic approach studies consider machine learning algorithms (e.g. [39]). Examples are decision trees (DTs), random forests, and support vector machines. This study employs DTs

Table 3: Activity type classes and potential classification definitions in different HTS data sets

Urgency constraints [34, 35]	**Work & flexibility constraints**	**Temporal constraints**	**Spatial constraints**	**Activity types in HTS Seoul [36]**	**Activity types in NHTS 2009 [8]**
Mandatory	**Work**	**Fixed start & end time**	**Fixed destination**		Work*
				Work	Go to work
					Return to work
				Business	Attend business meeting/trip
					Other work related
	Non-work fixed				School/religious activity
					Go to school as student
				School	Go to library - school related
				Education service	
		Opening hours			Day care
Maintenance	**Non-work flexible**	**Fixed start & end time**	**Fixed destination**	Bring/get	Transport someone
					Pick up someone;
					Take and wait
					Drop someone off
		Opening hours	**Available facilities**		Shopping/errands
				Shopping	Buy goods
					Buy services
					Use professional services
					Use personal services: grooming, haircut, nails
		Unconstrained			Buy gas
		Unconstrained / Opening hours			Pet care
	Non-work fixed	**Fixed start & end time**			Family personal business/obligations
			Fixed destination		Attend funeral/wedding
		Opening hours			Medical/dental services
Discretionary	**Non-work flexible**	**Fixed start & end time**	**Available facilities**	Personal & religious activities	Rest or relaxation/vacation
		Opening hours			Go to gym/exercise/ play sports
					Meals
					Get/eat meal
					Coffee/ice cream/snacks
		Unconstrained			Other reason
			Every zone	Leisure/recreation/ communication	Social/recreational
					Social event
					Go out/hang out
					Visit public place
			Fixed destination		Visit friends/relatives
	Non-work fixed	**Fixed start & end time**	**Available facilities**		Go to religious activity
			Fixed destination		Attend meeting: PTA/ home owners association/
Home		**Unconstrained**	**Home location**	Being at home	Home
Travelling**	**Work**	**Fixed start & end time**	**Fixed destination**	Back to office	
	Non-work flexible	**Unconstrained**	**Home location**	Back home	

* The 'Work' activity (NHTS-code 10) is slightly peculiar and differs from the expected interpretation. Subsection 5.2 and Table 5 provide more details.

** The act of traveling itself.

and therefore integrates rule-based concepts. However, the type of classifier can be varied (see also subsubsection 4.2.2).

3. Data description

Two HTSs were used in this research. The first HTS, the Seoul HTS, was conducted in the Seoul Metropolitan Area (SMA), Republic of Korea, in 2010. This data set consists of self-reported daily household activity-travel data from approximately 76,000 individuals. As reported in Table 2, this data set contains 11 distinct trip motives (or activity types), of which the 'home' activity will be excluded. The home activity is excluded from the experiments because this activity type is quite easy to classify and is mostly predicted with a very high accuracy (e.g. [40]). Additionally, due to a large share of home activities in the data set, its good classification capability obscures the sub-optimal or bad classifications of out-of-home activities. The Seoul HTS was included in this study to confirm the correct convergence of the proposed search algorithm to the optimum which was found in [6], and to benchmark the algorithm's performance gains. The convergence on this data set will be discussed and the performance of the algorithm will be compared to the approach in [6], justifying the need and benefits of the iterative search approach. The optimum set of ATCs of this data set will however not be discussed here. Interested readers may find a thorough analysis in [6].

The second HTS used in this study is the NHTS 2009 from the USA. It contains surveyed information from 308,901 individuals. This massive data set contains detailed trip information of approximately $1.17 \cdot 10^6$ trips, of which the trip purpose is encoded in 37 distinct classes. After excluding trips having the 'home' trip purpose, approximately 768,000 records remain to train activity type classifiers. The copious activity types in this data set are the reason for the development of the proposed methodology, as explained in the introduction. To the author's best knowledge, this is the richest activity type encoding in a HTS (not considering time-use surveys); see also Table 2. It is therefore a challenge to find the optimal set of activity types, which may be used in any activity type inference or annotation research. Additionally, this data set is employed in many studies to train their models. Finding and using an optimal set of activity types may enable the seamless consolidation of multiple research outcomes.

Only temporal variables such as activity start time and duration are used to train classifiers in this study. All other variables in the data are disregarded. This choice was made in order to present a generic categorization method which is applicable to as many big data sources and study areas as possible. Additionally, many applications start from e.g. GPS recordings, smart card data etc. for which classification based on temporal variables gives already good results [37]. Other types of attributes, e.g. spatial information, can however also be used in the proposed method.

The data was split in a train set (75%) and test set (25%). According to common practice, the train set is used to train a classifier, whilst the test set is used to evaluate its prediction accuracy on 'new' data.

4. Methodology

4.1. Grouping of activity types

This section discusses the combinatorial challenge of grouping or aggregating activity types into new classes. For example, in the set of ATCs [[1], [2], [3], [4], [5], [6], [7], [8], [9], [10]], activity types 3 and 6 may be merged into a new class as such: [[1], [2], [3, 6], [4], [5], [7], [8], [9], [10]]. The number of possible ATCs grows exponentially with the number of distinct activity types: n. This is the result of all the permutations of activity types across possible groups and the different combinations of possible group sizes. The order of activity types within a group, and the order of the groups among themselves does not matter. The possible group size combinations for a given n may be obtained by computing the integer partitions $p_{n,i}$. For example, for $n = 4$ the integer partitions $p_{4,i}\big|_{i=1..5}$ are $\{1{+}1{+}1{+}1,\ 2{+}1{+}1,\ 2{+}2,\ 3{+}1,\ 4\}$. Each element in these partitions represents a group's size. The first partition represents the case where no activity types are merged, the final one represents the case where all 4 activity types are grouped in one group of size 4. For each partition $p_{n,i} : g_1 + g_2 + ... + g_j = n$ where g_j is an element in partition $p_{n,i}$ representing a group's size, there are $x_{n,i}$ number of ways to distribute n activity types across the groups:

$$x_{n,i} = \frac{n!}{\prod_j \left(g_j! \cdot \sqrt[\frac{1}{f_j}]{f_j!} \right)} \tag{1}$$

where f_j represents the frequency of a particular element in the partition (which represents a group's size). For example, in the partition 2+1+1, element '2' has a frequency of 1. In partition 2+2, element 2 has a frequency of 2. The factor $\prod_j \left(\sqrt[\frac{1}{f_j}]{f_j!} \right)$ corrects $x_{n,i}$ for the permutations of equal-sized groups as the order of these equal-sized groups is unimportant, and should not increase $x_{n,i}$ (that is, $2_a + 2_b = 2_b + 2_a$). The total number of possible ATC combinations $N(n)$ is the sum of all $x_{n,i}$ for a given n: $N(n) = \sum_i x_{n,i}$. These values are listed in Table 1. One observes how the increase of possible combinations increases exponentially, hereby strengthening the justification for the need of the proposed methodology.

4.2. Optimization through local search

In order to optimize the ATCs, the proposed method combines some of the original activity types into a new class, and subsequently calculates the classification accuracy and entropy of the activity type variable. The classification accuracy represents the performance of predicting an ATC, and the entropy represents the amount of information such a prediction is giving. The entropy (or embedded information) is greatest when no activity types are merged into a new class, yet the classification accuracy increases when activity types are merged into new classes (as there are fewer classes to predict). Grouping or aggregating activity types into a new class will destroy some of the information entailed in the data.

For example, if two activity types 'Go to gym/exercising/play sports' and 'Go out/hang out: entertainment...' are merged into a new class 'Recreational', one can no longer make any distinction on the type of recreational activity. However, if one attempts to infer the activity type of a detected stop in a GPS trace, the prediction accuracy will be higher in case of the new class 'Recreational' compared to when using the two original classes. As a rule of thumb, classification accuracy increases with decreasing number of classes to predict.

4.2.1. The optimization parameter U

The aim of this paper is to optimize which original activity types should be grouped into a new class, and this for data sets for which it is impractical or impossible to simply calculate all ATC combinations (due to an extremely large amount of combinations). The proposed method defines an optimization parameter U which is maximized in an iterative search algorithm.

At the heart of the optimization strategy in [6] is the optimization parameter which may be calculated using Equation 2:

$$U = \frac{A_i - A_0}{R_A} - a\frac{E_0 - E_i}{R_E} \tag{2}$$

where A_i is the test set accuracy and E_i the activity type entropy of a particular combination i of ATCs. A_0 and E_0 are, respectively, the test set accuracy and activity type entropy of the reference case of no activity type aggregation into new classes. $R_A = A_{\max} - A_{\min}$ is the range in test set accuracy improvement and $R_E = E_{\max} - E_{\min}$ is the range in entropy reduction, observed within the set of results of all ATC combinations. Parameter a can be used to give a relative weight to either the classification accuracy improvement or to the entropy retention if there exists such an intrinsic bias for one of these indices. A sensitivity analysis of this parameter is described in [6]. The entropy may be calculated with Equation 3:

$$E = -\sum_j p_j \log_2(p_j) \tag{3}$$

where p_j is the probability on class j. However, Equation 2 can only be used when the results from all ATC combinations are known, as R_E depends on the minimum entropy $E_{\min}$, and R_A requires the maximum classification accuracy

$A_{\max}$ to be known. Note that the maximum entropy $E_{\max}$ and minimum classification accuracy $A_{\min}$ can be obtained from the reference case in which no activity types are grouped into a new class. In [6], the optimization parameter U was calculated only after the entropy and classification accuracy for all approx. 117,000 ATC combinations were calculated. Since calculating the entropy and classification accuracy for all possible combinations of ATCs is impossible given a large number of distinct activity types in the US NHTS (see Introduction), $E_{\min}$ and $A_{\max}$ need to be substituted.

The answer consists of allowing the trivial solution, that is, the case when all activity types are grouped into a single large class. In this trivial case, the entropy is zero (all activity type information is lost) and the classification accuracy is 100% (as only one class remains to predict). The results in [6] reveal that in practice $E_{\min}$ and $A_{\max}$ are in fact very close to, respectively, zero and one, thus supporting the proposed measure. Doing so, Equation 2 may be simplified to the following form in which all parameters (except A_i and E_i) can be calculated from the start:

$$U = \frac{A_i - A_0}{A_{\max} - A_{\min}} - a\frac{E_0 - E_i}{E_{\max} - E_{\min}} = \frac{A_i - A_0}{1 - A_0} - a\frac{E_0 - E_i}{E_0} \tag{4}$$

As a result, U may be calculated without the need to determine the classification accuracy and entropy for all possible ATC combinations beforehand. An iterative optimization approach is now possible.

4.2.2. The optimization algorithm

The parameter U (Equation 4) is used in the proposed optimization algorithm. The algorithm is described in pseudo code in Algorithm 1.

It starts with a seed (or 'reference') set of ATCs, for example this set of ten distinct activity types where no activity types were grouped into new classes (all activity types form their own group): [[1], [2], [3], [4], [5], [6], [7], [8], [9], [10]]. Different 'seed' sets may be used; each case could potentially converge to a different (and therefore local) optimum. If however consistently the same solution is found, this may be considered evidence for a global optimum. This

Algorithm 1 Pseudo code for the activity type class (ATC) optimization

Require:

Input data: activity-travel data ▷ e.g. from household travel survey

$E_0 \leftarrow$ entropy based on original input data (Equation 3)

$A_0 \leftarrow$ classification accuracy after training classifier on original data

$ATCs_0 \leftarrow$ chosen seed set of ATCs ▷ e.g. fully (dis-)aggregated

```
function U(ATCs)
    E ← entropy based on ATCs & input data (Equation 3)
    A ← classification accuracy after training a classifier for the ATCs
    U ← (A−A_0)/(1−A_0) − a (E_0−E)/E_0 (Equation 4)    ▷ Assume weighting factor a = 1
    return U
end function

ATCs_best ← ATCs_0
U_best ← U(ATCs_0)
while convergence criterion not satisfied do
    ATCs ← apply random change to ATCs_best
    if U(ATCs) > U_best then
        U_best ← U(ATCs)
        ATCs_best ← ATCs
    end if
end while
return ATCs_best
```

research uses both the fully *disaggregated* and the fully *aggregated* set of ATCs as reference set (seed). This is an approximate validation of the convergence to a single optimal solution.

Next, some random changes are applied to the set of ATCs, e.g. activity type '2' and '10' could be merged into a new class: [[1], **[2, 10]**, [3], [4], [5], [6], [7], [8], [9]]. A random change is defined as the exchange of one activity type from one group to another (this can be an empty group). The number of random changes that are applied are according to an exponential distribution: the probability of a single change is 64.4%, that of two changes 23.7%, that of three changes 8.7% etc. and this up to a maximum of ten random changes. Using this approach decreases the probability that the algorithm gets stuck in a local optimum and increases the probability that it will reach the global optimum. Note that this step is not completely random, as previously generated random grouping schemes are never used again (for obvious performance reasons). The random change generator is insensitive to the size of an existing group. This prevents a bias of large groups getting only larger, or vice versa. Multiple blocks of grouped ATCs can arise without biasness.

For this new set of ATCs, a DT is trained on the train set. The C4.5 (J48 in Weka [41]) DT classification algorithm yields an excellent classification accuracy and requires only a short time to train [6]. Therefore, this was the classifier of choice. Next, the activity classification accuracy is calculated based on the test set. The entropy retention in the data is determined as well. The optimization parameter U (Equation 4) is computed (considering $a = 1$ for this study). If the newly calculated U is larger than U_{best} of the best grouping scheme, the newly found grouping scheme will replace the previous best grouping scheme.

The previous steps of randomly changing the set of ATCs, training a DT classifier and calculating U is repeated until a stopping criterion is satisfied. The stopping criterion consists of a certain number of iterations without change in best U, which indicates that the algorithm converged to a (local) optimum (which is possibly equal to the global optimum). For the Seoul HTS data set, iterations were stopped after 100 cycles without a change in best U, whilst for the US NHTS data set this threshold was set to 4000 cycles. This predefined number was chosen after initial experimentation and may not be optimal. It is however critical that this number is chosen sufficiently large for cases of copious distinct activity types (more about this in subsection 5.1). This is important since more combinations of classes are possible and thus the optimum becomes more difficult to find. The algorithm needs sufficient time to try random different combinations before one can conclude convergence.

Note that, given a minimum amount of data, the sample size (see Table 2) has little influence. This is i.a. a consequence of the pruning stage in the DT learning phase. However, if not only temporal variables are used, then the sample size should perhaps also increase proportionally to ensure as much behavior as possible is captured in the DT.

In the experiments described in this paper, the algorithm was run for 10 (Seoul HTS 2010) or 32 (US NHTS 2009) times. Of these 32 runs on the US NHTS, 17 used the fully *disaggregated* set of ATCs as seed, the other 15 considered the fully *aggregated* set of ATCs as reference set. Due to the random changes applied to the ATCs during iterations, each run had a different path of convergence. Yet, as will be discussed in the results section, consistently the same optimum was found giving evidence for a global optimum.

5. Results

5.1. Convergence of the local search algorithm

First the proposed algorithm was run for the Seoul HTS data set, similar as in [6]. The intention of this experiment is to confirm that the proposed algorithm works, and that it yields major improvements in performance. Compared to [6], slightly different values for U are expected since an adapted formula is used in this study. The algorithm ran for 10 times (independently) and converged each time to the same optimum, which was reassuringly also the same as was found in [6]. Figure 1 illustrates the convergence of these runs. Although each run started at a different U value due to the random change at the start of

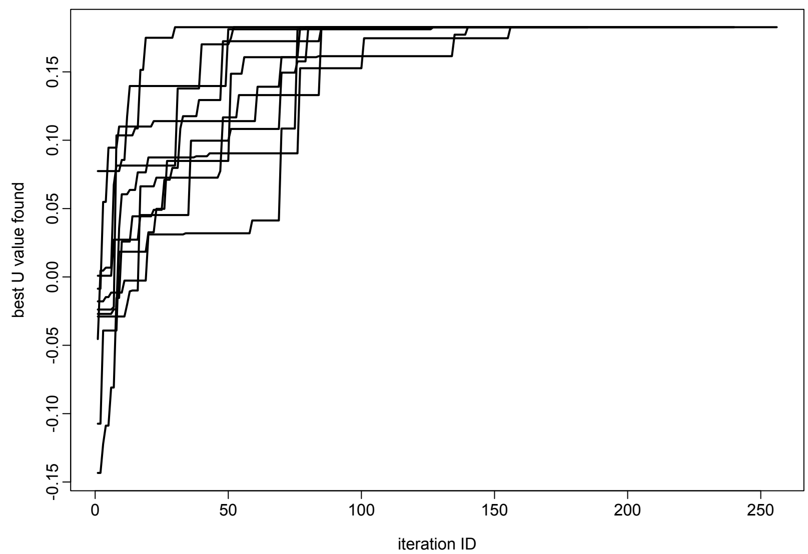

Figure 1: Convergence plot for 10 independent runs on the Seoul HTS 2010 data set. The slowest run finished in just under 13 minutes (Intel Core i5-4210M CPU @ 2.60GHz) and needed 255 iterations (100 part of the stopping criterion). All 15 found the same optimum.

the algorithm, they all converged to the same optimal U value. Figure 2 shows how each of the runs also followed a distinct conversion path, here shown by the evolution of the number of ATC during the conversion process. Notably, this exact same result could be found in just a couple of minutes, whilst in the approach of [6] approximately 30 hours on 20 threads of a high-end server were needed. As also concluded in that study, this optimum is considerably better than many ‘expert opinion’ activity type classification systems being used.

After having confirmed the excellent performance of the method on the Seoul HTS, the experiment was repeated on the US NHTS data set. In 32 independent runs, each time the same set of ATCs was found. This is a strong evidence that the revealed set of ATCs is indeed a globally optimal set of ATCs for the US NHTS. Figure 3 illustrates the convergence graphically. Of those 32 runs, 17 employed the original set of ATCs (fully disaggregated) as seed set (‘reference’ set), similarly as in the experiment on the Seoul HTS. The other 15 runs used the fully aggregated seed. The two types of runs used the extremes of possible

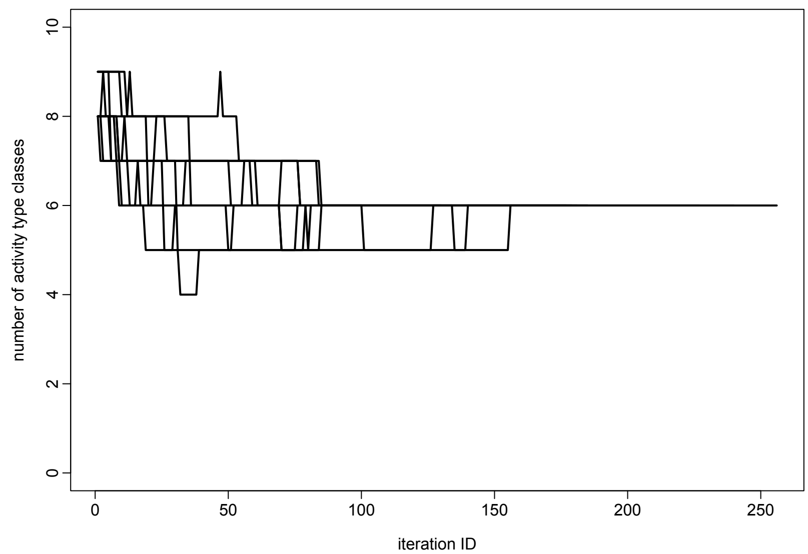

Figure 2: Evolution of the number of activity type classes during convergence of 10 independent runs on the Seoul HTS 2010 data set.

seed sets of ATCs: the first having as many as possible initial ATCs, the second as few as possible. Finding the same optimum starting from these two extremes using an iterative algorithm with such a large search space is a very positive result. Figure 4 shows the evolution of the number of ATCs for both types of runs. One clearly sees how both experiments start from 2 extremes (the maximum number of ATCs vs the minimum number of ATCs) and that they converge to the same optimum.

Within 32 runs, approximately 309,290 combinations of ATCs were trialed. As described in Algorithm 1, an equal number of DTs was trained. Combining the results of the 32 runs there were 261,537 unique combinations. Table 4 lists a selection of all those combinations, including also the most optimal set of ATCs as the first entry in the table (see also next section for more discussions).

Of the runs with as seed the fully *disaggregated* set of ATCs, the fastest run found the optimum in just under 79 hours (after 7,468 iterations), the slowest one in a bit over 140 hours (16,256 iterations). Of the runs based on the fully *aggregated* set of ATCs as seed set, the fastest run found the optimum in just under 47 hours (5,636 iterations), the slowest one in a bit over 76 hours (9,687 iterations). Mind that in all runs the final 4,000 iterations were part of the stopping criterion. Note that this is a mere fraction of the $3.82 \cdot 10^{30}$ sets of ATC combinations that would have to be analyzed with the method of [6]. Compared to the Seoul HTS, processing time for US NHTS took considerably longer. This is a consequence of the increased time to transform the ATCs in the data and subsequently train the DTs on this large data set.

The runs which started from the fully aggregated ATC converged faster than the ones starting from the set of disaggregated ATCs (see also Figure 3). This was somewhat unexpected, since the optimal set of ATCs (see subsection 5.2 and Table 4) appears to be more similar (i.e. 'closer') to the fully disaggregated seed than to the fully aggregated seed. This effect can however be explained by the fact that the search space is initially smaller (fewer possible alterations are possible starting from the fully aggregated ATC) and the random changes are more likely to be in the right direction.

The conversion results presented above are sensitive to the convergence criterion used in the algorithm. In earlier experiments, some runs failed to converge to the same optimal set of ATCs found by other runs. The U value at which these runs reached the stopping criterion was inferior to those of the other runs. Two likely causes were identified: (i) too few iterations occurred before the stopping criterion was fulfilled (set to 2,000 in these initial experiments) or (ii) these runs were stuck in a local optimum. The latter is unlikely, since by allowing up to 10 random changes on the previous best scheme (see section 4.2.2) it would be very likely that any local optimum could be avoided, on the condition that the algorithm is given enough time.

The first hypothesis was confirmed by the finding that the final set of ATCs of those unexpected runs also could be found within the set of iterations of the other runs (which *did* converge to the optimum). This means that there existed a direct path that could lead to the same optimum. By chance (i.e. too few iterations before stopping criterion was fulfilled) such a path was not

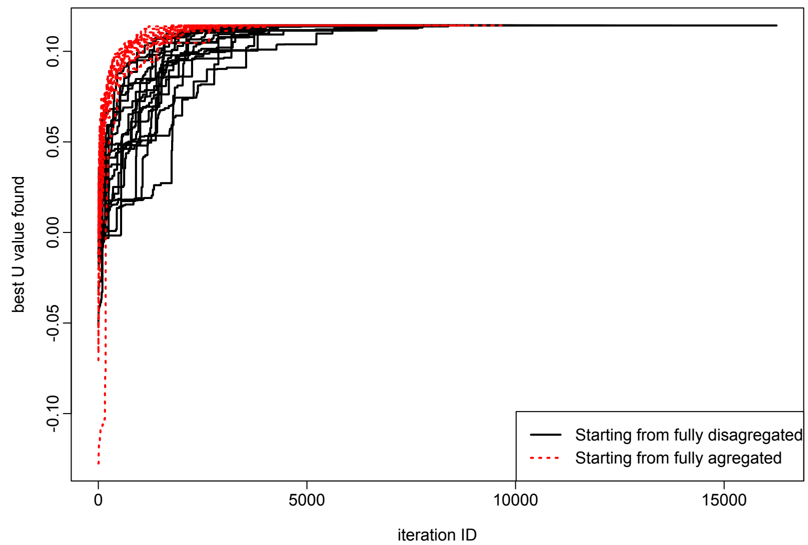

Figure 3: Combined convergence plot for 32 independent runs on the NHTS 2009 data set, for the two experiments with different seed sets of activity type classes (fully disaggregated (17 runs) and fully aggregated (15 runs)).

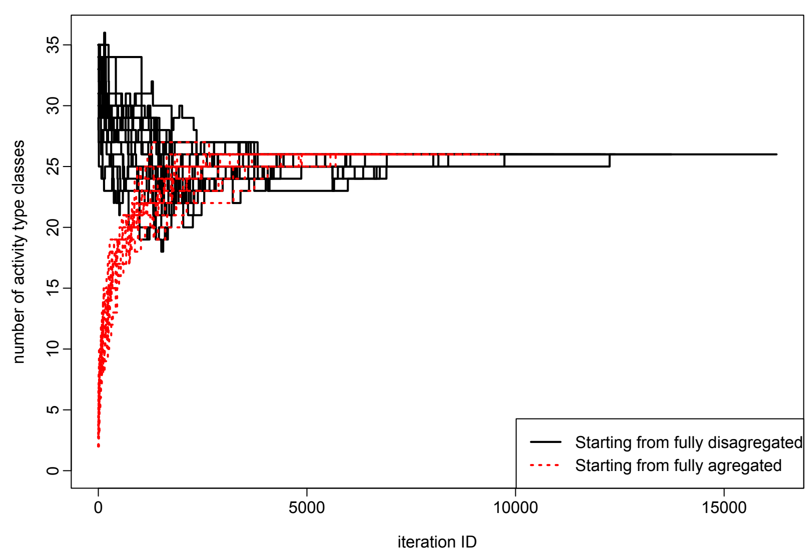

Figure 4: Combined evolution plot of the number of activity type classes during convergence of 32 independent runs on the NHTS 2009 data set, for the two experiments with different seed sets of activity type classes (fully disaggregated (17 runs) and fully aggregated (15 runs)).

found in these runs. Sufficient iterations should be allowed before convergence is concluded.

5.2. Optimal activity type classes for annotation

Table 4 lists some interesting results from the experiments on the NHTS 2009 data. The first entry is the most optimal ATC combination: 10, 11, 12, 13, 14, 20, 21, 23, 24, 50, 52, 54, 55, 60, 61, 62, 63, 64, 65, 70, 72, 80, 81, 83, 97, [22, 30, 40, 41, 42, 43, 51, 53, 71, 73, 82]. Compared to the reference case, its test set classification accuracy has more than doubled from 34.0% to 73.4%. This comes at a cost of losing 2.06 bits of information. It suggests these 26 distinct classes (excl. 'Home') are more optimal compared to the original 36. It merges the following activity types into a new class:

- 22: Go to religious activity
- 30: Medical/dental services
- 40: Shopping/errands
- 41: Buy goods: groceries/clothing/hardware store
- 42: Buy services: video rentals/dry cleaner/post office/car service/bank
- 43: Buy gas
- 51: Go to gym/exercise/play sports
- 53: Visit friends/relatives
- 71: Pick up someone
- 73: Drop someone off
- 82: Get/eat meal

This is a class which is hard to define. Some are flexible in nature (buying goods,) yet others have obligations to third parties and are not flexible (e.g. picking up or dropping off someone). However, none of them usually have a very

Table 4: Most optimal sets of the combined results of the 32 independent runs on the US NHTS 2009 data set and some interesting sets to compare with. Table 5 lists the encoding of the activity types. The best set of activity classes (1^{st} row in table) was the optimal set in all 32 runs.

Sets of activity classes (only grouped activity types are shown)	**Test Set Accuracy**	**Entropy**	$U(\downarrow)$
[22, 30, 40, 41, 42, 43, 51, 53, 71, 73, 82]	0.734	2.216	0.114273
*[10, 23], [22, 30, 40, 41, 42, 43, 51, 53, 71, 73, 82] **	0.734	2.216	0.114268
[23, 70], [22, 30, 40, 41, 42, 43, 51, 53, 71, 73, 82]	0.734	2.214	0.113756
[10, 23, 70], [22, 30, 40, 41, 42, 43, 51, 53, 71, 73, 82]	0.734	2.214	0.113751
*[10, 62], [23, 70], [22, 30, 40, 41, 42, 43, 51, 53, 71, 73, 82] **	0.734	2.214	0.113751
[10, 11, 12, 13, 14], [20, 21, 22, 23, 24], [30, 40, 41, 42, 43, 50, 51, 52, 53, 54, 55, 60, 61, 62, 63, 64, 65, 70, 71, 72, 73, 80, 81, 82, 83, 97] (ref.: [42])	0.851	0.977	0.001754
Reference case (original 36 activity types)	0.340	4.276	0
[10, 11, 12, 13, 14], [20, 21, 22, 23, 24, 30, 40, 41, 42, 43, 50, 51, 52, 53, 54, 55, 60, 61, 62, 63, 64, 65, 70, 71, 72, 73, 80, 81, 82, 83, 97] (ref.: [4])	0.895	0.618	-0.014185
[10, 11, 12, 13, 14], [23, 24, 30, 40, 41, 42, 43, 50, 51, 52, 53, 54, 55, 63, 64, 80, 81, 82, 83, 97], [20, 21, 22, 60, 61, 62, 65, 70, 71, 72, 73] (ref.: e.g. [43])	0.733	1.271	-0.107685
[10, 11, 12, 13, 14], [20, 21, 22, 23, 24], [40, 41, 42, 43], [50, 51, 52, 53, 54, 55], [60, 61, 62, 63, 64, 65], [70, 71, 72, 73], [80, 81, 82, 83] (ref.: [8] (first digit NHTS codes))	0.476	2.754	-0.150825
[24, 30, 40, 41, 42, 43, 61, 64, 65, 82], [10, 11, 12, 13, 14, 20, 21, 70, 71, 72, 73], [22, 23, 50, 51, 52, 53, 54, 55, 60, 62, 63, 80, 81, 83, 97] (ref. e.g. [44])	0.632	1.539	-0.197553
[10, 11, 12, 13, 14], [20, 21, 22, 23, 24], [30, 40, 41, 42, 43], [50, 51, 52, 53, 54, 55, 60, 61, 62, 63, 64, 65, 70, 71, 72, 73, 80, 81, 82, 83, 97] (ref.: [30])	0.599	1.741	-0.200993
[40, 41, 42, 43], [70, 71, 72, 73], [10, 11, 12, 13, 14], [20, 21, 22, 23, 24], [50, 51, 52, 53, 54, 55], [30, 60, 61, 62, 63, 64, 65, 80, 81, 82, 83, 97] (ref.: [45])	0.485	2.429	-0.213240

Note: *italic** sets of activity classes represent multiple variations with 10: ‘Work’ (see text) long duration compared to others such as e.g. ‘Work’, and they could in theory occur at almost any time within a day. All activity types occur at a relatively high frequency (see Table 5). Many of these activities are likely to be chained together: picking up or dropping off people whilst visiting friends/relatives or going to play sports, or getting something to eat before (or after) doing some shopping etc. This makes it hard to distinguish between these activity types based on only temporal profiles, and hence it makes sense to merge them into a single class.

The next ATC schemes in the list combine 10: ‘Work’ with all other activity types which are not in the large group of the most optimal scheme (24 distinct combinations, e.g. [10, 23]; [10, 70]; [10, 62]; etc.) and finally it also joins the large group. Because of space constraints, only the best performing of all those variations is listed in italics in Table 4. From Table 5 one observes that activity type 10 is slightly peculiar, as its weighted frequency is many orders of magnitude smaller than other activity types. It is clearly different from ‘Go to work’ as the latter has a frequency which is approximately 10^5 times larger. The exact definition of the ‘Work’ activity could not be found. Because of the very low frequency, the impact of this activity type on the classification accuracy and entropy retention is very small. This experiment concludes that in practice these variations with activity type 10 may not be different from the most optimal scheme and one could most likely ignore them.

Subsequently in Table 4 one finds the scheme where, in addition to the large group from before, also 23: ‘Go to library: school related’ and 70: ‘Transport someone’ are merged into a single class. This could make sense as this experiment used only time-related variables to train the DTs, and one could intuitively think the temporal distributions of both activity types may be similar. Again different combinations with activity type 10 are listed afterwards. The schemes discussed so far perform similar as the most optimal scheme. One has to be cautious when interpreting the rank in Table 4 as the algorithm does not guarantee to find all ATC combinations.

Next in Table 4 are seven interesting activity class combination schemes from

Table 5: Trip motive codes in US NHTS 2009 which were used in this studys optimization of activity type classes. There are 37 distinct codes (including Home)

NHTS 2009 codes	**Description of trip motive**	**Weighted frequency**
1	Home	1.35E+11
10	Work	2.16E+05
11	Go to work	3.11E+10
12	Return to work	5.73E+09
13	Attend business meeting/trip	1.07E+09
14	Other work related	7.90E+09
20	School/religious activity	1.13E+09
21	Go to school as student	1.18E+10
22	Go to religious activity	6.98E+09
23	Go to library: school related	4.54E+08
24	OS - Day care	8.29E+08
30	Medical/dental services	6.30E+09
40	Shopping/errands	7.10E+09
41	Buy goods: groceries/clothing/hardware store	4.40E+10
42	Buy services: video rentals/dry cleaner/post office/car service/bank	1.12E+10
43	Buy gas	6.60E+09
50	Social/recreational	3.78E+09
51	Go to gym/exercise/play sports	1.34E+10
52	Rest or relaxation/vacation	3.28E+09
53	Visit friends/relatives	1.76E+10
54	Go out/hang out: entertainment/theater/sports event/go to bar	6.84E+09
55	Visit public place: historical site/museum/park/library	1.85E+09
60	Family personal business/obligations	4.48E+09
61	Use professional services: attorney/accountant	1.11E+09
62	Attend funeral/wedding	6.68E+08
63	Use personal services: grooming/haircut/nails	1.47E+09
64	Pet care: walk the dog/vet visits	2.94E+09
65	Attend meeting: PTA/home owners association/local government	1.61E+09
70	Transport someone	3.09E+08
71	Pick up someone	1.10E+10
72	Take and wait	1.19E+09
73	Drop someone off	1.20E+10
80	Meals	7.92E+08
81	Social event	2.49E+09
82	Get/eat meal	2.04E+10
83	Coffee/ice cream/snacks	2.98E+09
97	Other reason	2.59E+09

literature to compare with the optimal scheme. An attempt was made to merge ATCs in a similar fashion as in these studies. The most obvious comparison may be made with an ATC scheme based on the first digit of the NHTS codes [8]. Even though there are much fewer activity classes to predict compared to the most optimal scheme, its classification accuracy is much lower at 47.6% compared to 73.4%. This deficiency outweighs the fact that this scheme retains slightly more information than the optimal scheme. The scheme based on [45] performs similarly. The ones inspired by [44] and [30] perform worse than the optimal scheme on both the classification accuracy and information retention. The schemes inspired by [4],[42] and [43] have similar of better classification accuracies compared to the optimal scheme, however these lost a major portion of their information content as a consequence.

Depending on the research, there might exist a reason for employing one of the suboptimal schemes (e.g. some activity types need to be predicted and may not be merged, or a predefined number of ATCs is required). Yet, without such justification, this work suggests one should strongly consider using the revealed most optimal set of ATCs in order to simultaneously maximize the prediction accuracy and the information in that prediction.

6. Conclusion

As demonstrated in previous research [6], there is a strong need for activity categorization standards in the domain of trip purpose annotation (i.e. activity type classification). Most existing researches use a suboptimal set of ATCs in their methodology (without providing a justification), leading to high classification accuracies, but low information in the prediction. An optimization strategy that was proposed in previous research [6], has shown a limitation: the issue of copious distinct ATC combinations and its associated long computation time. This issue makes it practically *impossible* to apply the optimization strategy to data sets having copious activity type classes (ATCs).

The aim of this paper is to optimize which original activity types should be merged into a new class, and this for data sets for which it is impractical or impossible to simply calculate all ATC combinations due to an extremely large amount of combinations. The paper suggests a revision of the optimization method in [6]. The proposed method defines an optimization parameter U, based on classification accuracy and information retention, which is maximized in an iterative search algorithm.

The local search algorithm starts from a predefined set of ATCs and iteratively tries to optimize it by applying random changes. In each iteration, the search algorithm randomly combines or disjoins (with constraints) some of the activity types into a new class, and subsequently calculates the classification accuracy based on temporal variables, as well as the retained information (entropy). The optimization parameter U which is based on both indices is maximized.

Tests on the Seoul household travel survey (HTS) concluded that the proposed algorithm is (i) able to find the optimal set of ATCs and (ii) that it is much more efficient at this task compared to previous research [6].

Experiments with the very large national household travel survey of the U.S. (US NHTS 2009), which is to the authors' knowledge the HTS with the most copious trip purpose (activity type) variable, concluded that (i) the algorithm is capable of running this 'worst-case-scenario' in a reasonable amount of time (ii) a *global* optimum was found thanks to using seeds at two 'extremes'.

Depending on the research, there might exist a reason for employing one of the suboptimal schemes (e.g. some activity types need to be predicted and may not be merged, or a predefined number of ATCs is required). Yet, without such justification, this work suggests one should strongly consider using the proposed algorithm to objectively determine the most optimal set of ATCs in order to simultaneously maximize the prediction accuracy and the information in that prediction.

Future research will also include spatial and regional variables to apply the methodology to a big transport data activity type annotation problem. Furthermore, the application of data fusion based on annotated optimized ATCs will be investigated. Models based on traditional ATCs and optimized ones can be compared. The concept of an optimized set of classes could also be transferred to other classification problems such as trip mode inference from GPS data.

References

[1] R. Kitchin, Big data and human geography: Opportunities, challenges and risks, Dialogues in Human Geography 3 (3) (2013) 262–267. doi:10.1177/2043820613513388.

[2] W. D. Lee, K. Choi, T. Bellemans, J. H. Hwang, S. Cho, Data Mining Method for Smart Card Data Using Household Travel Survey: A Pilot Study of Public Transportation in Suwon, South Korea, in: International Association of Traveler Behavior Conference, 2015.

[3] T. Feng, H. J. P. Timmermans, Detecting Spatial and Temporal Route Information of GPS Traces, in: I. Ivan, I. Benenson, B. Jiang, J. Horák, J. Haworth, T. Inspektor (Eds.), Geoinformatics for Intelligent Transportation, Springer International Publishing, Cham, 2015, pp. 61–75.

[4] Y. Lu, L. Zhang, Imputing trip purposes for long-distance travel, Transportation 42 (4) (2015) 581–595. doi:10.1007/s11116-015-9595-0.
URL http://dx.doi.org/10.1007/s11116-015-9595-0

[5] L. Montini, N. Rieser-Schüssler, A. Horni, K. W. Axhausen, Trip Purpose Identification from GPS Tracks, Transportation Research Record: Journal of the Transportation Research Board 2405 (2014) 16–23. doi:10.3141/2405-03.
URL http://trb.metapress.com/openurl.asp?genre=article{&}id=doi:10.3141/2405-03

[6] W. Ectors, S. Reumers, W. D. Lee, K. Choi, B. Kochan, D. Janssens, T. Bellemans, G. Wets, Developing an optimised activity type annotation

method based on classification accuracy and entropy indices, Transportmetrica A: Transport Science 13 (8) (2017) 742–766. doi:10.1080/23249935.2017.1331275.

[7] A. Ahern, G. Weyman, M. Redelbach, A. Schulz, L. Akkermans, L. Vannacci, E. Anoyrkati, A. Von grinsven, Analysis of National Travel Statistics in Europe, Tech. rep., European Commission - Joint Research Centre (2013). doi:10.2788/59474.
URL https://publications.europa.eu/en/publication-detail/-/publication/90d01fd1-0c91-437b-8af6-edc3b4ff7f9a

[8] U.S. Department of Transportation, Federal Highway Administration, 2009 National Household Travel Survey (2009).
URL http://nhts.ornl.gov

[9] NASA/WMAP Science Team, WMAP- Age of the Universe (2012).
URL https://map.gsfc.nasa.gov/universe/uni{_}age.html

[10] Department for Transport, National Travel Survey, 2002-2014 [computer file]. 9th Edition. (2015). doi:10.5255/UKDA-SN-5340-5.

[11] Department of Economic Development; Jobs; Transport and Resources (DEDJTR), Victorian Integrated survey of Travel and Activity 2007 (2007).
URL www.economicdevelopment.vic.gov.au/vista

[12] Department of Economic Development; Jobs; Transport and Resources (DEDJTR), Victorian Integrated survey of Travel and Activity 2009 (2009).
URL www.economicdevelopment.vic.gov.au/vista

[13] E. Cornelis, M. Hubert, P. Hyunen, K. Lebrun, G. Patriarche, A. De Witte, L. Creemers, K. Declercq, D. Janssens, M. Castaigne, L. Hollaert, F. Walle, La mobilité en Belgique en 2010 : résultats de l'enquête BELDAM (2012).

[14] D. Janssens, K. Declercq, G. Wets, Onderzoek Verplaatsingsgedrag Vlaanderen 4.5 (2012-2013), Tech. rep., Hasselt University, Transportation Research Institute (IMOB) (2014).

URL http://www.mobielvlaanderen.be/pdf/ovg45/ovg45-analyse-globaal.pdf

[15] M. Loechl, Stability of Travel Behaviour: Thurgau 2003 (2005).
URL http://archiv.ivt.ethz.ch/vpl/publications/tsms/tsms16.pdf

[16] V. Chalasani, K. W. Axhausen, Mobidrive: A six week travel diary (2004).
URL https://www.ethz.ch/content/dam/ethz/special-interest/baug/ivt/ivt-dam/vpl/tsms/tsms2.pdf

[17] Liikennevirasto - Finnish Transport Agency, National Travel Survey 20102011, Tech. rep., Liikennevirasto - Finnish Transport Agency.
URL http://www.liikennevirasto.fi/documents/21386/152020/HLT{_}2010{_}2011{_}esite{_}ENG.pdf/025982d6-0b3e-4919-b49d-7c0ef61eaf0d

[18] J. Armoogum, J.-P. Hubert, D. Francois, B. Roumier, M. Robin, S. Roux, Enquête nationale transports et déplacements 2007-2008 (ENTD 2007-2008) (Rapport technique), Tech. rep., IFSTTAR; INSEE; SOeS (2011).
URL http://www.statistiques.developpement-durable.gouv.fr/fileadmin/documents/Themes/Transports/Transport{_}de{_}voyageurs/Deplacements/Fichiers{_}details{_}2011/GuideMethodologique.pdf

[19] Central Statistics Office, National Travel Survey 2009, Tech. rep., An Phríomh-Oifig Staidrimh - Central Statistics Office, Dublin, Ireland (2011).
URL https://www.ucd.ie/t4cms/NTSReport2009.pdf

[20] Metropolitan Transport Authority, The Report of Household travel survey in Seoul Metropolitan Area [In Korean], Tech. rep., Seoul (2012).

[21] Korea Transportation Institute, National Transportation Demand Survey and Database Establishment in 2010: Passenger O/D Survey on the National Area, Tech. rep. (2011).

[22] Centraal Bureau voor de Statistiek (CBS), Rijkswaterstaat (RWS), Onderzoek Verplaatsingen in Nederland 2013 - OViN 2013 (2014).
URL http://dx.doi.org/10.17026/dans-x9h-dsdg

[23] M. Klemenčič, M. Lep, B. Mesarec, B. Žnuderl, Potovalne navade prebivalcev v Mestni občini Ljubljana in Ljubljanski urbani regiji (2014).

[24] Trafik Analys, RVU Sverige 20112014 - Den nationella resvaneundersökningen (RVU Sweden 2011-2014 - national travel survey), Tech. rep., Stockholm (2015).
URL http://www.trafa.se/globalassets/statistik/resvanor/rvu-sverige-2011-2014.pdf

[25] S. Hasan, S. V. Ukkusuri, Urban activity pattern classification using topic models from online geo-location data, Transportation Research Part C: Emerging Technologies 44 (2014) 363–381. doi:10.1016/j.trc.2014.04.003.
URL http://dx.doi.org/10.1016/j.trc.2014.04.003

[26] J. Wolf, R. Guensler, W. Bachman, Elimination of the travel diary: Experiment to derive trip purpose from global positioning system travel data, Transportation Research Record 1768 (1) (2001) 125–134. doi:10.3141/1768-15.

[27] J. Wolf, S. Schönfelder, U. Samaga, M. Oliveira, K. Axhausen, Eighty Weeks of Global Positioning System Traces: Approaches to Enriching Trip Information, Transportation Research Record: Journal of the Transportation Research Board 1870 (August) (2004) 46–54. doi:10.3141/1870-06.

[28] P. R. Stopher, C. FitzGerald, J. Zhang, Search for a global positioning system device to measure person travel, Transportation Research Part C: Emerging Technologies 16 (3) (2008) 350–369. doi:10.1016/j.trc.2007.10.002.

[29] W. Bohte, K. Maat, Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands, Transportation Research Part C: Emerging Technologies 17 (3) (2009) 285–297. doi:10.1016/j.trc.2008.11.004.

[30] L. Shen, P. R. Stopher, A process for trip purpose imputation from Global Positioning System data, Transportation Research Part C: Emerging Technologies 36 (November) (2013) 261–267. doi:10.1016/j.trc.2013.09.004.

[31] T. Feng, H. J. Timmermans, Detecting activity type from gps traces using spatial and temporal information, European Journal of Transport and Infrastructure Research 15 (4) (2015) 662–674.

[32] P. McGowen, M. McNally, Evaluating the Potential To Predict Activity Types from GPS and GIS Data, in: Transportation Research Board 86th Annual Meeting, 2007.

[33] M. Allahviranloo, W. Recker, Mining activity pattern trajectories and allocating activities in the network, Transportation 42 (4) (2015) 561–579. doi:10.1007/s11116-015-9602-5.
URL http://dx.doi.org/10.1007/s11116-015-9602-5

[34] S. T. Doherty, A. Mohammadian, The validity of using activity type to structure tour-based scheduling models, Transportation 38 (1) (2011) 45–63. doi:10.1007/s11116-010-9285-x.

[35] M. D. Meyer, Institute of Transportation Engineers., Transportation Planning Handbook, 4th Edition, fourth edi Edition, John Wiley & Sons, 2016.
URL https://books.google.rs/books?id=qFiBDAAAQBAJ{&}pg=PA162{&}lpg=PA162{&}dq=Institute+of+Transportation+Engineers+2009+Transportation+Planning+Handbook{&}source=bl{&}ots=H5KfBpA144{&}sig=LOGHSPJk793WyH6B315mBG2JP8M{&}hl=sr{&}sa=X{&}ved=0ahUKEwiM4vG2nPzWAhWrBsAKHTrdAB8Q6AEIXj

[36] J. Choi, W. D. Lee, W. H. Park, C. Kim, K. Choi, C. H. Joh, Analyzing changes in travel behavior in time and space using household travel surveys

in Seoul Metropolitan Area over eight years, Travel Behaviour and Society 1 (1) (2014) 3–14. doi:10.1016/j.tbs.2013.10.003.

[37] T. Kusakabe, Y. Asakura, Behavioural data mining of transit smart card data: A data fusion approach, Transportation Research Part C: Emerging Technologies 46 (2014) 179–191. doi:10.1016/j.trc.2014.05.012.

[38] C. Zhong, X. Huang, S. Müller Arisona, G. Schmitt, M. Batty, Inferring building functions from a probabilistic model using public transportation data, Computers, Environment and Urban Systems 48 (2014) 124–137. doi:10.1016/j.compenvurbsys.2014.07.004.
URL http://dx.doi.org/10.1016/j.compenvurbsys.2014.07.004

[39] S. Reumers, F. Liu, D. Janssens, M. Cools, G. Wets, Semantic Annotation of Global Positioning System Traces, Transportation Research Record: Journal of the Transportation Research Board 2383 (2013) 35–43. doi:10.3141/2383-05.
URL http://trrjournalonline.trb.org/doi/10.3141/2383-05

[40] S. Reumers, F. Liu, D. Janssens, G. Wets, The Annotation of Global Positioning System (GPS) Data with Activity Purposes Using Multiple Machine Learning Algorithms, in: S. Rasouli, H. Timmermans (Eds.), Mobile Technologies for Activity-Travel Data Collection and Analysis, Vol. i, IGI Global, 2014, pp. 119–133. doi:10.4018/978-1-4666-6170-7.ch008.

[41] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. H. Witten, The WEKA data mining software, ACM SIGKDD Explorations 11 (1) (2009) 10–18. doi:10.1145/1656274.1656278.

[42] S. Lee, M. Hickman, Trip purpose inference using automated fare collection data, Public Transport 6 (1-2) (2014) 1–20. doi:10.1007/s12469-013-0077-5.
URL http://dx.doi.org/10.1007/s12469-013-0077-5

[43] B. Kochan, Implementation, validation and application of an activity-based transportation model for Flanders, Ph.D. thesis (2012).
URL https://uhdspace.uhasselt.be/dspace/bitstream/1942/15067/1/PhDthesisBrunoKochan.pdf

[44] M. Bradley, P. Vovsha, A model for joint choice of daily activity pattern types of household members, Transportation 32 (5) (2005) 545–571. doi:10.1007/s11116-005-5761-0.

[45] Y. Lu, S. Zhu, L. Zhang, Imputing Trip Purpose based on GPS Travel Survey Data and Machine Learning Methods, Transportation Research Board 1250.