Travel-Mode Classification for Optimizing Vehicular Travel Route Planning

Lijuan Zhang, Sagi Dalyot and Monika Sester

Abstract Navigating and travelling between destinations with the help of Geographic Information Systems route planning is a very common task carried out by millions of commuters daily. The route is mostly based on geocoding of the addresses given by the traveller based on static road network into digital-map positions, and thus the creation of path and directions needed to be taken. Today's navigation data sets rarely contain information about parking lots, related to building entrances, and walking paths. This is especially relevant for large building complexes (hospitals, industrial buildings, city halls, universities). A fine-tuned route tailored for the driver requirement, e.g., park the car close-by to destination, is required in such cases to save time and frustration. The idea of this chapter is to extract this information from the navigational behaviour of users, which is accessible via an analysis of GPS traces; analysis of car commuters in relation to their point of departure and destination by analysing the walking path they took from—and to—their parked car in relation to a specific address. A classification scheme of GPS-traces is suggested, which enables to classify robustly different travel modes that compose a single GPS trace. By ascribing the classified vehicular car trace, which is accompanied by a walking path to/from the car, to a specific address, it is made feasible to extract the required ascribed data: parking places corresponding to that address. This additional data can later be added to the road network navigation maps used by the route planning scheme to enable the construction of a more fine-tuned optimal and reliable route that will prevent subsequent detours.

Keywords Data mining · Classification · GPS · Route planner · Optimization

L. Zhang (✉) · S. Dalyot · M. Sester
Institut für Kartographie und Geoinformatik (IKG), Leibniz Universität Hannover,
Appelstraße 9a, 30167 Hannover, Germany
e-mail: Lijuan.Zhang@ikg.uni-hannover.de

J. M. Krisp (ed.), *Progress in Location-Based Services*,
Lecture Notes in Geoinformation and Cartography, DOI: 10.1007/978-3-642-34203-5_16,

1 Introduction

Nowadays, vehicle drivers use commercial route planners via electronic maps to guide them from their point of departure to destination: from finding an office to planning cross-country excursions. The routes given by mass-use route planning systems, such as Yahoo! Maps, Google Maps, to name a few, are commonly constructed and based usually on static data derived from the use of road networks existing in geospatial libraries and databases. Most algorithms that rely on such road networks are usually based on designed cost-function schema—or a mix of such functions—associated with the network 'edges'; most commonly used ones are travel time, distance, scenic value, etc., which together construct the 'most optimal' route. Relying solely on the road network yields that not all available data regarding neighbouring attributes, for example parking lots or means of access, are taken into consideration when the route is being constructed. Moreover, the destination point, which is normally an address, is being geocoded into a single position, e.g., coordinate, on the digital-map. A driver who is unfamiliar with the area and uses a vehicle to get to that destination address will usually require a car parking, which adds some ambiguity since this might not exist or be accessed directly from the destination address; thus, a detour to the assigned route is unavoidable. Consequently, an alternate route that is the optimal one for the driver's needs might have been constructed in case the driver's destination was the parking place affiliated with his specific destination address that he wishes to get to. Figure 1 depicts such an example, in which a fine-tuned route (right image) that directs the driver from point of departure (denoted as *A* in green bubble) to the affiliated parking place (brown circle) of his desired address (denoted as green arrow) is shorter by 400 m than the optimal preliminary route constructed (left image). A detour to find the affiliated parking (denoted as red arrow) will cost the driver in this case to drive an additional distance of 1,100 m. These translate to a fine-tuned solution that will evidently save the driver, who is unfamiliar with that area, time and unnecessary frustration.

GPS-data today are often collected through mobile handheld devices. As a result, roads, paths and routable traces derived by GPS measurements are collected straightforwardly by pedestrians, vehicular commuters, bicycle riders, and more. The assumption is that every GPS-trace stores some unique and relevant characteristics that are dependent on a specific travel-mode resultant by the road-type it was acquired on. A single such trace is usually composed from several sub-traces; each corresponding to a different travel-mode. Based on this assumption, a GPS-trace composed of vehicular driving data accompanied by walking paths before and/or after points to a route where parking took place; parking that apparently is essential to get to a specific address or location. By being able to classify correctly such trace should deliver the data that is required to construct the optimal route planning, and hence update the existing navigation map with this significant and valuable data.

This research proposes a classification scheme of GPS-traces, which enables to classify robustly different travel modes that compose a single GPS trace.

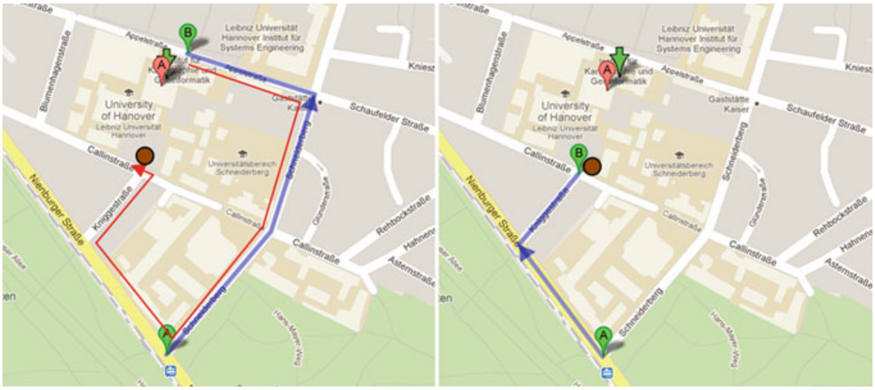


Fig. 1 An example showing different planned routes using Google maps (*left*): directly to the address (*denoted as green arrow*), and to the address's affiliated parking place (*brown circle*), yielding a route that is 400 m shorter (*right*). *Red arrow* (*left*) depicts the additional 1,100 m required in the first route to get to the parking place (*source* http://maps.google.com/)

By ascribing the classified vehicular car trace, which is accompanied by a walking path to/from the car, to a specific address (building), it is made feasible to extract the required ascribed data: parking places corresponding to a specific address. This additional data is added to the route planning scheme enabling to extract a more fine-tuned optimal route.

2 Related Work

Since data in road networks is usually static, path construction is mostly based to some extent of pre-computations, which are stored and re-used for acceleration reasons. Classical static algorithms might suggest the Dijkstra's Algorithm (Dijkstra 1959), which maintains an array of tentative distances for all nodes exist in the road network. As shown by Delling et al. (2009), though most algorithms used today will aim to guarantee (to some extent) that indeed the shortest route (path) is found, the commercial route planning systems usually settle for an approximate result. This usually happens due to the fact that these systems neglect certain data as being unimportant for certain preference issues, thus moving away from the optimal solution.

Due to increase in mobile navigation systems, as car navigation and the use of smartphones equipped with GPS, dynamization is integrated into the route planner: traffic jams and such can also be incorporated (for example: Biagioni et al. 2011). This also yields the use of Multi-Criteria Routing (for example: Li et al. 2011; Nadi and Delavar 2011), which states that the fastest or shortest route in the analysed network is often not the 'best' one—since other criteria are to be taken into account. For example: price—or in this chapter's case—time spent for finding a parking spot. This also yields Multimodal Routing constraints (walking to and from) that might also have a significant influence and criteria on the optimal route chosen. Jariyasunant et al. (2010) had suggested a real-time route planner system to predict shortest path between any points—while relying on static maps only—together with incorporating third-party information. It also integrated to the analysis user-defined data that considered the travel mode in different regions of the path to achieve close to optimal routing. Axhausen et al. (2003) attempted to construct route planning scheme that is based solely on a vast number of GPS traces collected; still, no attempt was carried to try and classify the different traces—assuming all where only vehicular ones—in order to produce a fine-tuned optimal route tailored for a specific need, for example: means of transportation. Adapting such scenarios is perhaps one of the main challenges in near future.

Since GPS observations alone supply only with geometric and temporal data, specific data-mining methods have to be applied in order to extract the required information of travel-mode type classification. Most approaches include two steps in such process: a segmentation of the trajectory into a series of single travel-mode; and, assigning a specific travel-mode to all segments exist in the series. A basic assumption is usually made (Chung and Shalaby 2007; Zheng et al. 2008) that walking is necessary when a mode-change occurs, e.g., change point-based segmentation method. Segmentation of such usually relies on small or no position change accompanied by low speed value and time-length of segments. Though usually found to be accurate, the research proposed here suggests using additional characterization values and parameters, such as heading and single travel-mode pattern-classifiers, thus introducing more robust and non-ambiguous segmentation to a given GPS-trajectory. This is a key-element here, since we look here for vehicular travel modes accompanied by walking (specifically: from/to car).

As for classification, most of the existing methods compare some known preliminary travel-mode related measures, e.g., rule-based values, to empirically determined values. Most commonly used values are derived from the speed and acceleration of a segment (single travel-mode), such as maximum and mean speed (Bohte and Maat 2009; Oliveira et al. 2006). Still, it was shown that these approaches might present ambiguous-classification, thus yield errors and lack the flexibility to examine properly change in pattern and uncertainty of the travel-mode. Also, the thresholds depend on a specific study-area and supplementary data, making them not generic to be implemented for all environments and test-data. To overcome the uncertainty and ambiguity existing in the data, the use of fuzzy logics as a replacement for the empirically determined values is also suggested for classification. The speed and acceleration measures are related as fuzzy sets, while fuzzy membership patterns are structured to enable travel-mode classifiers via linguistic rules (Tsui and Shalaby 2006; Schüssler and Axhausen 2008). Although these researches show an improvement in robustness of classification, the determination of bounds for each linguistic rules associated with each measure, such as fast speed and long travel, was found to be depended on subjective experience exist in the travel-logs. A Decision Tree is also used (for example: Reddy et al. 2008), where the authors present its superiority to other approaches commonly used. Still, some perquisites in the classification process were considered (grouping several different travel mode classes to a single one) together with training data that was relatively small. In past research we had presented a first version of segmentation and multi-stage classification approach dividing it to two supplementary stages. The first stage used a fuzzy-logic classification for identifying walking, cycling and vehicular travel-modes. The second stage entailed the classification of the vehicular travel-mode class using Support Vector Machines (SVMs) schema into car, train, tram, and bus travel-modes. Adopting that scheme showed promising and qualitative classification results of all six different travel-modes. In this chapter, this is explained in more detail with additional examples, while data extracted (classified) is used for the determination of important locations in the network, namely parking lots and access paths to buildings.

3 Segmentation and Classification Methodology

Classifying correctly car and walk travel modes is the main emphasize given here; by doing so, the added data required to fine-tune the GPS routes and make them optimal is made feasible. Since the research carried here try to simulate the natural way of commuters in their everyday all available data is collected and classified. Work schema of the proposed methodology is presented in the diagram, depicted in Fig. 2.

3.1 Data Collection

GPS traces involved with vehicular car routes in the urban region of Hanover City were collected using android-based handheld mobile devices (smartphones). These devices, which are equipped with GPS for collection of positions, used a designated application for data collection that was developed specifically for this research. To evaluate the reliability of the approach presented here, travel-modes are recorded also by the application as reference—together with the collection of the traces' positions. Emphasis was given to simulate the natural way of commuters in their everyday life during data collection without applying any special concerns or restrictions.

Since the reliability and consistency of the travel-mode classification process is a primer key in this approach, training data with supplementary information is required. The Android-based application, which was programmed in Java, collects GPS data and reference travel-mode tagging specified by the data-collection user. The Graphical User Interface (GUI) of the TRACER, depicted in Fig. 3, presents specific and easy-to-use functions. These functions include: a toggle button for starting and stopping data acquisition (left); and, a button enabling the user to

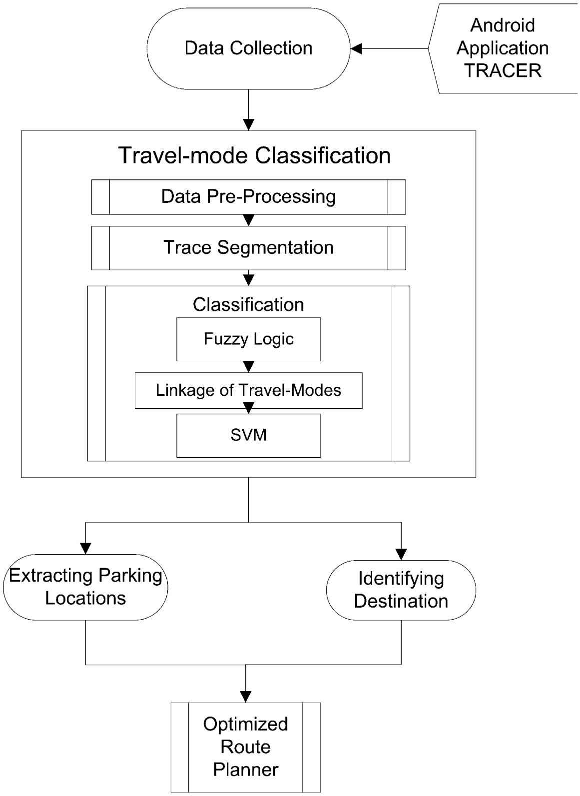


Fig. 2 Proposed methodology workflow diagram

select (and modify) the current travel-mode (right). The user can choose from six different travel-modes, which were chosen to simulate all travel-modes commonly exercised: Walk, Bike, Car, Bus, Train, and Tram. Though this chapter addresses two specific travel-modes: Car and Walk; still, emphasis was given to verify all travel-modes are enabled and collected, and 'used' during the classification process.

Since the data acquisition is supposed to be a passive procedure, the TRACER application provides with a notification system that requires the user attention on specific predefined events. The notification system utilizes all modes of user notifications provided by modern smartphones, e.g., visual, sound and haptical. If required by the user, a checkbox labelled "silent" allows the user to choose not to be notified by certain predefined events. The idea is that the application will not irritate the user as much as possible, so the data-collection will be a smooth and almost 'transparent' process to the user. Still, certain notifications are used here, while their common goal is to obtain the user's current travel-mode, which is vital for the handling of the training data.

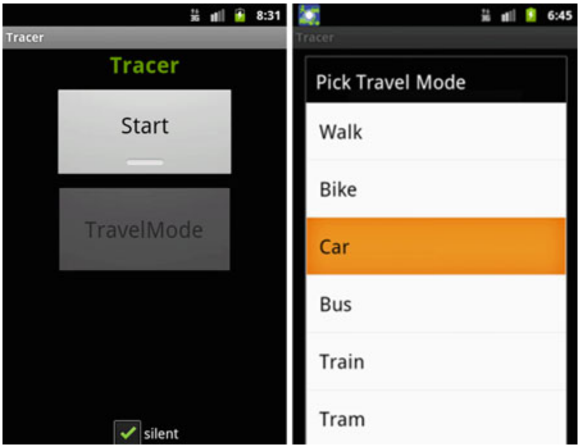


Fig. 3 TRACER GUI: main view (*left*); and, travel-mode selection (*right*)

The TRACER application implements the following events:

- Constant travel-mode update: this event forces the user to update the current travel-mode every 10 min. This is implemented to prevent the user from forgetting to do so, i.e., did not change travel-mode whilst travel-mode had changed.
- GPS-signal loss: this event is triggered only after gaining back of GPS-signal, which was lost for more than 20 s (tunnel, non-coverage area, etc.). This event also prevents cases where travel-mode had changed while no GPS-signal was available.
- Speed inconsistency: this event is triggered when there seems to be speed anomalies in respect to the travel-mode assigned by the user. Thus, speed limits, which are based on some coarse preliminary knowledge, for walking and cycling travel-modes for more than 10 consecutive seconds are implemented. Since thresholds used are coarse, as such they are only a type of warning enforcing and travel-mode change during the data collection process.

3.2 Data Pre-Processing

The positional accuracy of mobile handheld GPS signal in normal conditions can reach several meters (Wolf 2006). Still, it is quite common that the positional accuracy is even worse, in cases where there is a lack of sufficient satellites coverage, equipment that is not being ideally positioned (this is common with the use of GPS in smartphones), signal that is being reflected by tall buildings, bad weather—to name a few. The errors are reflected directly on the position of the acquired GPS data. Moreover, since travelling attributes are the key-features for enabling reliable travel-mode classification, positional errors are projected directly on such attributes, such as speed and acceleration. Thus, preliminary reduction of error affects before the calculation of attributes and parameters is implemented. The use of smoothing method to reduce speed errors by averaging its neighbourhood is introduced. The range of the smoothing is five travel-epochs, or seconds under common conditions. Heading smoothing was not implemented here, because heading is not a continuous phenomenon by nature, thus smoothing might remove its characteristic and degrade its reliability as a travel-mode parameter that is required on latter stages of classification.

3.3 Traces Segmentation

Since the relations between both car and walk travel-modes are mandatory to achieve the approach addressed here, it is only logical that a GPS trace is not derived from a single travel-mode; instead, it is composed of several different travel-modes: walking to the car, driving to work and parking the car, and walking from the car to the office (for example). Thus, before any classification can be implemented, a division of the single GPS trace into segments of (still unknown) individual travel-modes has to be implemented. These different segments are categorized and characterized as sub-traces. A definition is made, which states that a sub-trace is composed of a single travel-movement segment that is separated (divided) by two stops. A stop can be a temporal pause in movement (no change in position over time), but also a change in travel-mode (e.g., change point-based segmentation method). Figure 4 (top) depicts an example of stops, which consist of a sequence of observations—black segments—that have very low speed and very small distance changes that on the same time are not classified and defined as a walk. Thus, identifying stops and consequently filtering-out the stops-data before actual classification takes place is important (calculating travel-modes parameters should avoid using stop-data, otherwise modifying these parameters and weakening the reliability of the classification process).

Most researches trying to identify change point-based segmentation commonly use values associated with stops—namely small distance changes per-time and low speed value. This research proposes the use of magnitude in heading change; this parameter was found to be vital for a robust identification of stops. Figure 4 (bottom) depict the idea behind this, where stops are always accompanied with large magnitude values in heading changes (in black), which cannot be explained by realistic movement changes (as the ones in blue or red, for example, which correspond to specific travel-modes). The reason of the existence of these accompanied large values is the result when no change in position occurs, i.e., stops or low speed values, thus large and random magnitude values in heading changes exist due to relatively small change in position. This can be interpreted as heading change noise, which expresses a very limited change in position over time.

The following thresholds are used in this research to form the sub-traces segmentation existing in a single GPS-trace: (All values given below are in respect to

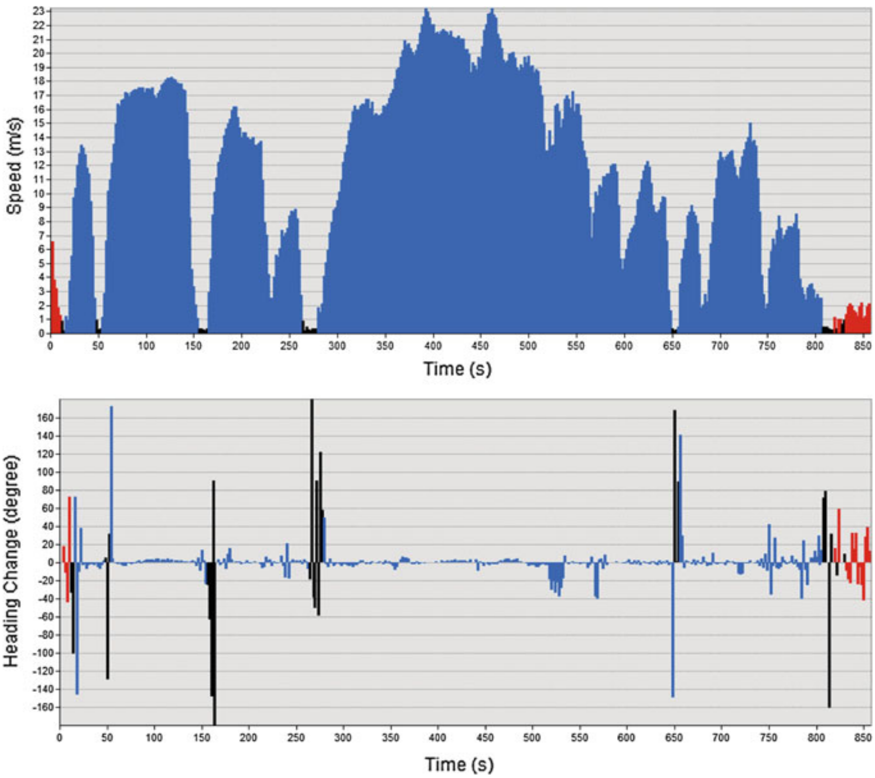


Fig. 4 Speed derived from a GPS-trace representing approximately 15 min of travelling divided into different individual segments in blue (*walk is in red*), and identified stops in *black* (*top*); corresponding heading changes magnitude (*bottom*), showing high correspondence exist between stops (*in black*) having high heading change values

the 1 s time-stamp of the GPS-traces locations collected in this research; modifying these values should be considered when other time-stamp values are used).

- Variation in position: cases where distance change for 5 consecutive seconds is less than 5 m are identified as stop.
- Speed values: cases where for 5 consecutive seconds speed value is less than 0.5 m/s are identified as stop.
- Magnitude in heading change: cases where heading change value for 5 consecutive seconds is larger than 100 decimal degrees are identified as stop.

The algorithm workflow is as follows:

- From the first observation point on, in case accumulative distance from that observation point to the fifth consecutive observation point neighbour is less than 5 m—break the trace from that point; go to step 2.

- Check all points existing in this 5-points segment: if speed is smaller than 0.5 m/s and/or change of heading magnitude is larger than 100°—check next point. Else—break the trace from that point and go to step 1. If no break occurs—go to step 3.
- Assign the sixth point (in regard to the first observation point examined) as the beginning point—go to step 1.
- Stop when reaching the end of the trace.

3.4 Segments Travel-Mode Identification

Classification is applied to the separate segments identified earlier, which together compose a single GPS-trace. This is finalized by linking of neighbouring segments that have been classified with the same travel-mode to form a sub-trace. It was found that the characteristics of walk and bicycle travel-modes are prominently different from all other travel-modes, categorized here as vehicular class. Additionally, a classification all four vehicular, namely car, bus, tram and train, which is based solely on the segments, might result in an ambiguous results. This is due to the fact that the divided segments might present similar characteristics (parameters values), while the characteristics of the whole sub-trace of a specific travel-mode are not utilized. For example, buses and cars are specifically different on the fact that buses have regular stops and cars do not. Still, by examining a single segment only might show otherwise.

As a result, adopting a multi-stage method is required: on the first stage, pedestrian and bicycle travel-modes are differentiated from motorized vehicles based on specific characterizations and specification of their segments. On the second stage, segments are linked up to form sub-traces, and consecutively car, bus, tram and train travel-modes are classified based on the specific characterizations and specification of the sub-traces. Still, the proposed idea behind this research is concerned mainly with car travel-modes associated with walking travel modes to/from the car, thus the first stage should satisfy this requirement. Nonetheless, the second stage is presented in short here, which is required where validating that the travel-mode classified is indeed of a car is required.

First stage classification, which is derived from two main parameters associated with travel modes, is employed based on fuzzy logic classification scheme:

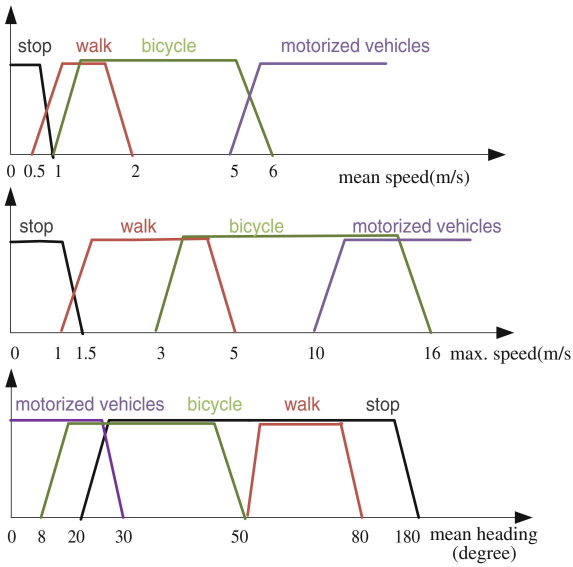
- Speed related characteristics are very important for identifying travel-modes, mainly in this stage. Four principal parameters are used here, namely: mean speed, maximum speed, mean acceleration, and maximum acceleration. These are calculated for each individual segment of the trace. To achieve more reliable parameters and reduce classification errors and bias that might exist when using one observation alone for each segment, the parameters of maximum speed and maximum acceleration are calculated based on the average values of the largest five (5) values exist in the segment.

- Heading related parameters, namely mean and maximum heading magnitude changes, are also employed. The heading change is calculated in a way that it is in the range of (−180°, 180°) decimal degrees. When calculating mean heading changes all values are transferred to positive because the magnitude alone—and not the sign—is of importance. As shown earlier in Fig. 4, while maximum heading change corresponds to stops, walking always show a large magnitude value for the average heading changes, thus making this parameter very useful for this stage classification.

Figure 5 depicts three of the classifiers, namely mean speed, maximum speed and mean heading changes, used in the fuzzy logic classification stage. Wide enough ranges for the three classifiers are used to include all possible segments into consideration whilst avoiding making wrong classification. As stated earlier, additional classifiers are used to validate correct travel-mode separation and classification. From Fig. 5 it is clear that there are some overlay areas for the parameters used, between: stop and walk, walk and bicycle, and bicycle and motorized vehicles. Taking all three under consideration, a minority of segments will fall into these overlay areas simultaneously. Solving these ambiguities yielded extra parameters to be incorporated into the fuzzy logic classification process:

- Stop and walk overlay area: maximum heading change is introduced. Since stops are always accompanied with high magnitude values of heading change, it was found that in case this value for a specific segment is larger than 80 decimal degrees, the entire segment can be classified as a stop (and not walk).
- Walk and bicycle overlay area: the use of second order polynomial is performed, where the polynomial is fitted to the segments speed pattern. The coefficients (a,

Fig. 5 Fuzzy logic classification: value range used for different travel-modes—mean speed (*top*), maximum speed (*middle*), and, mean heading changes (*bottom*)



b, and c) of the best fitting polynomial are used, where it was found that walking usually show a constant value of the second coefficient (b)—as opposed to bicycle, which is noisier.

- Bicycle and motorized vehicles overlay area: the use of maximum acceleration is introduced. It was found that when bicycle travels with a relatively high speed (mean speed >5 m/s) it is accompanied with high value of acceleration (maximum acceleration >4 m/s^2)—as opposed to motorized vehicles, which will usually show a smaller maximum acceleration value when travelling at these speeds.

3.5 Sub-Traces Construction: Linkage of Travel-Modes

Neighbouring segments with the same travel-mode are linked up to form sub-traces, which are assigned a travel-mode of walk, bicycle or motorized vehicle. The segments are checked with specific predefined rules to ensure they are not incorrectly classified before joined to a sub-trace. For example, a vehicular segment, which has a relatively low speed (parking process, for example), may be wrongly identified as bicycle. However, if this segment's neighbouring segments both are classified as vehicular segment travel-mode, and since it is not possible to transfer directly from bicycle to car, for instance, without stop or walk, the travel-mode is corrected and re-classified accordingly.

In order to correct the possibly wrongly classified segments, rules are applied during the linking procedure according to the basic travel knowledge:

- A travel-mode should exceed the period of 120 s—the use 120 s is designed to eliminate sub-traces that are too short and thus have no significance on the sub-trace, or are wrongly classified.
- Stop duration between two neighbouring segments of one sub-trace should be less than 120 s—if the stop duration is longer than 120 s then the trace should be treated as two individual sub-traces.
- No direct transformation from bicycle to any of the motorized vehicle class is possible—unless at least 120 s of walking or stop took place. The time duration threshold of 120 s is used to avoid linking two different modes together. This linkage rule is of major importance here, since it helps in identifying walk travel modes accompanied with car travel mode, and perhaps more importantly, avoids the classification of bicycle travel mode that come before/after a car travel mode.

3.6 SVM Classification

Since this research is aimed at classifying car and walk travel-modes only, the first stage facilitates this. To assure the classification of car travel-mode (and not public bus or train, which is normally accompanied with walking, for example), the use of the supervised learning method SVM is employed, aimed at classifying the different motorized vehicles (that includes car, bus, tram and train). SVMs are a popular machine learning method used in recent years for classification and other learning tasks. This method projects the parameters to a high–or infinite–dimensional space and constructs a hyperplane, which can be used for classification (Smola and Schölkopf 1998). The SVM produces a model based on a set of training data (attributes together with target values), and then uses this model to predict the target value of the test data with attributes only to find the solution for the optimization problem. A kernel function is used; in this case a Gaussian Radial Basis Function (RBF) that is suitable for cases where the relation between class and attributes is simultaneously nonlinear and linear (Hsu et al. 2003).

SVM classification is based solely on the constructed sub-traces of the already-classified motorized vehicles. The entire sub-trace is treated as a single object, and the attributes of each sub-trace are presumed to describe the characteristics of a unique travel-mode. An assumption is made (for a more reliable classification of cars) stating that bus, tram and train travel-modes should present regular stops, which are longer than those of car; this together with similar travel duration between two consecutive stops. Additional parameterization of travelling characterization—11 parameters in total—is used as attributes in the SVM implementation:

- Mean and standard deviation of maximum speed
- Mean and standard deviation of average speed
- Mean and standard deviation of acceleration
- Mean and standard deviation of average acceleration
- Mean and standard deviation of travelling duration
- Ratio of stop duration in respect to travelling duration.

Each segment within an individual sub-trace is used for the calculation of the aforementioned attributes. The attributes are scaled before applying SVMs to range (0, 1). Both the corresponding attributes of training and testing data are scaled in the same way. The main advantage of doing so is avoiding the attributes in greater value ranges dominating those in smaller numeric ranges, together with benefit of reducing calculation complexities.

4 Travel Route Plan Optimization

4.1 Extracting Parking Locations

After classification of all GPS-traces is finalized, all traces composed of car accompanied by walking travel-modes are analysed for the proposed optimization route planning process. First, all change points are identified where travel-modes are altered from car to walking—or vice versa. These locations refers and points to the desired information about parking lots related to building entrances. Also, these locations refer to starting/ending point of walking paths taken by the driver that are located nearby the point of departure—or destination. Consequently, these locations will have to be associated to specific addresses in their vicinity to be associated as their parking lots.

4.2 Identifying Destination

Since the positional accuracy of the GPS is at least several meters (and worse in build-up areas), and GPS multipath signal errors and signal-loss is common in the vicinity and inside of buildings, identifying correctly the building in which the walking started/ended in is important. This stage is achieved by performing a buffering process around building features in the close vicinity of the start and end points locations identified earlier. A building feature buffer that contains in its spatial extent (intersect) the first/last walking signal acquired in the trace, or some walking trace-segments, is assumed to be associated to that trace; hence, destination address can be identified. An example of this process is depicted in Fig. 6. Since parking lots are mostly situated in the vicinity of large building complexes (hospitals, industrial buildings, and such), the assumption is that if several building buffers contain the same walking positions, all can be attributed to the same parking lot position since it probably serves all these buildings. Buffer size used is derived from the positional accuracy of the building features (existing in the given database), but also from the positional certainty of the GPS signals acquired during the walking trace. Since build-up areas is of poor positional certainty, this value should be considered based on some knowledge of the errors at hand. In the examples analysed in the next chapter, buffer size used had the magnitude of 10 m, which was found to be sufficient.

5 Experimental Results

5.1 Classification

Evaluating the reliability and certainty of the classification methodology is carried out. This is an important assessment to verify that indeed when all the collected observations are processed, only those GPS-traces that are relevant to the fine-tuning process, namely car traces accompanied by walk, are identified correctly. 149 GPS-traces were collected in the study-area of Hannover City. Almost all traces are composed of two or more travel-modes. Table 1 depicts the fuzzy-logic classification results, showing high statistical classification certainty of close to 100 % for all classified travel-modes. When compared to the available reference

Fig. 6 Building buffer containing walking travel modes to/from car parking place: car traces are represented in *blue*, walk traces are represented in *red*

Table 1 Fuzzy-logic classification results

Travel-mode	Total	Correct	Statistical classification certainty (%)
Walk	47	44	94
Bicycle	19	18	95
Motorized vehicle	170	165	97

data (TRACER tags) the most common erroneously classification is walk-stop; this usually occurs when walking presents very low speed together with rapid stops.

Table 2 depicts the SVM classification results after comparing all observations to the available reference data (TRACER tags). Perhaps the most significant figure in this table is depicted in the upper row, showing 98 % certainty for car travel-modes that are correctly classified; this is very important to the scope of this research since these traces are later used for the optimization process. Analyzing the error matrix received for the SVMs classification showed that no other type of vehicular travel mode was classified as car, which also strengthens the assurance of this classification process.

Table 2 SVMs classification results

Travel-mode	Training data	Testing data	Correct	Statistical classification certainty (%)
Car	49	50	49	98
Bus	11	10	8	91
Tram	19	9	7	78
Train	4	2	2	100
Total	83	71	66	93

5.2 Identifying Parking Lots and Entrances

Since all GPS-traces are classified automatically and with high statistical certainty, the final stage in which the route plan is being optimized can take place. All traces composed of car accompanied by walking travel-modes are identified, so the extraction of parking locations and the identification of buildings/addresses can take place. An example is depicted in Fig. 6, which shows a building buffer polygon containing traces of walking travel-modes (in red) emerging from car travel modes (in blue). It is visible that the 10 m building buffer covers in its extent all walking traces leading to/from it. Since this building has two entrances, some walking traces lead to its north entrance associated to the east parking lot, while the other lead to its south entrance associated to the south parking lot. Consequently, both parking lot positions are associated to this building. Also, it is visible that other building in the vicinity will also be associated to one (or both) parking lots extracted here.

Figure 7 depicts a scenario of the proposed route optimization: blue traces represent car travel-mode and red traces represent walking travel-mode (right). The left image depicts the default route plan produced via Google Maps, while the right image depicts the route which would be taken if the parking lot was available in the data set, which is devised by knowing the digital-map position (coordinates) the car should get to in order to park. The position of the parking place is the one that was associated with the walking route—represented by the red traces. This scenario shows that the optimized route is much shorter—approximately 800 m. Not only that the route is shorter, thus saving driving time, the driver will most probably save some time and frustration in finding this specific parking place.

Figure 8 depicts two additional scenarios, in which by knowing the parking place associated with the desired address the driver is directed to a different but more appropriate and optimized location. By doing so, the driver is not directed straight to the address, represented by a pink buffer in both images. In this way, the system is able to automatically detect the most appropriate locations for parking spaces related to certain addresses. This information is revealed from the behaviour of users.

These examples amplify the argument presented in this research: by including the knowledge regarding parking places in the vicinity of building and facilities addresses into the navigation maps and databases, it is possible to construct a more

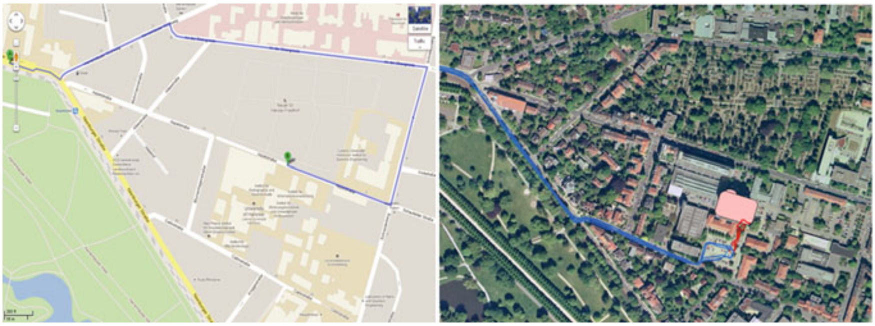


Fig. 7 Route planning scenario: 'off-the-shelf' Google maps route (*left*) (*source* http://maps.google.com/), and optimized route proposed in this research, showing precise classification and extraction of parking place associated with desired address

Fig. 8 Optimized route planning scenarios showing precise classification and extraction of parking place associated with desired address avoiding detour and loss of valuable time: car trace in *blue*, walking trace in *red*, destination in *pink polygon*

fine-tuned route that answers the user requirements; the closest location where he or she can park the car in order to get to the desired address. This is done automatically and with high certainty and positional reliability.

6 Conclusions and Discussion

Routing techniques using static road networks have made tremendous progress in the last few years, mainly in the speed-up domain where fast query response time is critical. Still, answering also requirements derived from multi-criteria optimization for specific users is under constant development. This chapter presents a working frame that facilitates a more fine-tuned, optimised and fully-automatic route planner that answers specific user demands. In this case, giving driving instruction to a parking place associated with the desired address (destination). This modified route planner makes use of an automatic detection and classification travel-modes working schema of GPS-traces, which are acquired today straightforwardly by travellers. Segments of GPS-traces and sub-traces comprised of an individual travel-mode are found with very high certainty, reliability, and efficiency. Preliminary tests and examples presented in the chapter are promising, showing good results in producing routes that are usually shorter, but perhaps more important—tailored to the problem at hand. Thus proving technical feasibility as well as having positive effects on the drivers travelling these routes.

Future work will involve the analysis and verification of the presented process on larger datasets, while integrating the extracted knowledge with navigation maps of road networks that are used for real-time route planning. It will also involve the adaptation of the proposed classification methodology to other multi-criteria scenarios, such as adding bicycle path to the route planner. As the appropriate parking locations associated with an address are automatically derived from the users' behaviour, it is also possible to reveal temporal properties, e.g., that the best parking place varies over the day: in the morning there might be enough parking lots in front of a building, whereas in the afternoon a nearby alternative parking lot may be the most appropriate one to use. Also, our first approach for identifying the entrances is rather straightforward using buffer threshold; future work will entail the use of clustering approaches. These are issues of future research we plan to deal with.

References

Axhausen KW, Schönfelder S, Wolf J, Oliveria M, Samaga U (2003) 80 Weeks of GPS-traces: approaches to enritching the trip information. Arbeitsbericht Verkehrs- und Raumplanung, 178, Instut für Verkehrsplanung und Transportsysteme, ETH Zürich, Zürich

Biagioni J, Gerlich T, Merrifield T, Eriksson J (2011) Easytracker: automatic transit tracking, mapping, and arrival time prediction using smartphones. In SenSys, pp 68–81. ACM

Bohte W, Maat K (2009) 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. Emerg. technol. 17(3):285–297

Chung E, Shalaby A (2007) A trip reconstruction tool for GPS-based personal travel surveys. Transp Planning Technol 28(5):381–401

Delling D, Sanders P, Schultes D, Wagner D (2009) Engineering route planning algorithms. In: Algorithmics of large and complex networks. Lecture Notes in Computer Science, vol 5515. Springer, Berlin, pp 117–139.

Dijkstra EW (1959) A note on two problems in connexion with graphs. Numer Math 1:269–271

Hsu C, Chang C, Lin C (2003) A practical guide to support vector classification. Technical report, Department of computer science and information engineering, National Taiwan University

Jariyasunant J, Work DB, Kerkez B, Sengupta R, Bayen AM, Glaser S (2010) Mobile transit trip planning with real-time data. Transportation Research Board Annual Meeting, p 17

Li YT, Huang B, Lee DH (2011) Multimodal, multicriteria dynamic route choice: a GIS-microscopic traffic simulation approach. Annals of GIS 17(3):173–187

Nadi S, Delavar MR (2011) Multi-criteria, personalized route planning using quantifier-guided ordered weighted averaging operators. Int J Appl Earth Obs Geoinf 13(3):322–335

Oliveira M, Troped P, Wolf J, Mattheww C, Cromley E (2006) Mode and activity identification using GPS and accelerometer data. In: Proceedings of transportation research board 85th annual meeting, p 12

Reddy S, Burke J, Estrin D, Hansen M, Srivastava M (2008) Determining transportation mode on mobile phones. In: 12th IEEE international symposium on wearable computers. ISWC, pp 25–28

Schüssler N, Axhausen KW (2008) Processing GPS raw data without additional information. Working paper, p 15

Smola AJ, Schölkopf B (1998) On a kernel–based method for pattern recognition, regression, approximation and operator inversion. Algorithmica 22:211–231

Tsui SA, Shalaby AS (2006) Enhanced system for link and mode identification for personal travel surveys based on global positioning systems. J Transp Res Board 1972:38–45

Wolf J (2006) Applications of new technologies in travel surveys: travel survey methods, Quality and future directions. Elsevier, Oxford, pp 531–544

Zheng Y, Liu L, Wang L, Xie X (2008) Learning transportation mode from raw GPS data for geographic applications on the web. In: Proceedings of the 17th international conference on world wide web, WWW 08, pp 247–256