An Approach to Semantic Processing of GPS Traces

K. Rehrl[1], S. Leitinger[2], S. Krampe[2], R. Stumptner[3]

[1]Salzburg Research, Jakob Haringer-Straße 5/III, 5020 Salzburg, Austria
Email: k.rehrl@salzburgresearch.at

[2] TraffiCon - Traffic Consultants GmbH, Schillerstraße 25, 5020 Salzburg, Austria
Email: krampe@trafficon.eu

[3]Institute for Application Oriented Knowledge Processing, University of Linz, Altenberger Straße 69, 4040 Linz, Austria
Email: rstumptner@faw.jku.at

1. Introduction

Recording the movement of objects or persons with GPS-technology has been widely adopted. So called GPS traces are used for enriching travel diaries (Wolf et al. 2001), for learning significant places (Ashbrook and Starner 2003), for monitoring animal movement (Steiner et al. 2000) or for building road maps (Cao and Krumm 2009). While all of the mentioned applications rely on GPS traces, the methods for collecting and processing the data differ. One of the basic goals in processing GPS raw data is to extract re-occurring motion patterns (Laube et al. 2005, Zheng et al. 2008). In a current research project called HOTSPOT, we try to answer the question, which knowledge about an object or person's movement can be extracted from a single GPS trace. One of the unique aspects of the project is that our approach to knowledge extraction only relies on GPS raw data and explicitly avoids any kind of map matching or usage of additional data sources.

In this position paper we tackle aspects of semantic data processing. With semantic data processing we refer to methods attaching meaning to sub-sequences of GPS traces. We propose a method for semantic processing of GPS traces, resulting in basic motion and course change activities. The method can be applied to any kind of raw GPS data and prepares the data for further analysis of motion patterns.

2. GPS traces

For testing the algorithms for semantic data processing the project team collected a reference data set of about 400 GPS traces. Traces were collected in a representative manner, considering different kinds of GPS receivers, different modes of transport (on foot, running, hiking, bicycle, bus, tram, train, car and combined modes), different geographic regions (intensively built areas, rural landscapes, tracks with tunnels, forests), different road types, different days with different satellite constellation, different daytimes as well as different weather conditions. It was necessary to collect our own GPS traces, since the automatic recognition of motion patterns can only be validated by comparing the patterns to real motion behaviour. However, a carefully annotated reference set of GPS tracks would be a great help in validating the algorithms.

From the collected raw GPS data only lat/lon coordinates, elevation and timestamps were used. Other parameters like velocity, acceleration and course changes were calculated from subsequent positions and timestamps. We found it a good practice to post-calculate all relevant parameters from the basic positions, since velocities, accelerations and courses vary between receivers.

For recording GPS traces we used a sampling rate of 1 Hz, although a sampling rate up to 20Hz with high-end receivers and a sampling rate of 5Hz with semi-professional receivers (e.g. QStarz BT-Q1000EX) would be feasible. Some of the receivers (e.g. the Garmin Forerunner series) optimize the amount of logging data by providing adaptive logging (e.g. logging a position every 1-5 seconds depending on the travelled distance or absolute course change).
In a pre-processing step we detected and removed severe GPS errors (sudden drifts, unrealistic values in velocity and course changes).

3. Semantic processing

Basically, a GPS trace can be interpreted as a discrete capture (according to the chosen sampling rate) of the motion of objects over time. One of the goals of semantic processing is to abstract point data to higher level motion patterns. Before starting with the processing it is worth to think about the basic motion patterns an object or person can accomplish while moving. Certainly, each of the patterns is depending on physics of the moving object (e.g. vehicle) or the physiology of a person and the underlying surface (e.g. uphill or downhill, road infrastructure used for movement). The basic parameters to express motion in space and time are *velocity* and *course* (Zheng et. al., 2008). Both parameters can be computed from two sequent GPS measurements. By describing changes of these basic parameters over time, a set of six basic motion patterns can be defined (Table 1).

Table 1: Basic motion patterns

Parameter	Basic motion pattern	Description	Units
Velocity	Stand still	No motion	
	Steady motion	Motion with steady velocity	m/s
	Positive acceleration	Increase of velocity	+m/s^2
	Negative acceleration	Decrease of velocity	-m/s^2
Course	Positive course change	Increase of degrees	+°/s
	Negative course change	Decrease of degrees	-°/s

The first step in the semantic processing of GPS traces is the extraction of these basic motion patterns. Therefore, a GPS trace is analysed by searching sub-sequences of the patterns. Pattern matching is done with parameter thresholds, e.g. which change in velocity should be defined as positively accelerated motion or which change in heading should be identified as positively change of heading. Table 2 shows empirically founded rules and parameter thresholds.

Table 2: Rules for detecting motion patterns

Parameter	Basic motion pattern	Rules
Velocity	Standstill	velocity < 1 m/s
	Steady motion	velocity > 1 m/s & acceleration < +0.3 m/s^2 and acceleration > -0.3m/s^2
	Positive acceleration	velocity > 1 m/s & acceleration > +0.3 m/s^2
	Negative acceleration	velocity > 1 m/s & acceleration < -0.3m/s^2

Course	Positive course change Negative course change	course change > +0.4°/s course change < -0.4°/s

3.1 Semantic classification of motion

Figure 2 shows an example of detected motion patterns in a GPS trace recorded during a train ride. Slowdown (red), stop (yellow) and speedup (green) at a train station were automatically detected. Also a short period of steady motion (blue) could be successfully detected.

Figure 1: Part of a GPS trace recorded during a train ride

Figure 2: Automatically detected motion patterns in the GPS trace. (1) Standstill (yellow), (2) positive acceleration (green), (3) negative acceleration (red), (4) steady motion (blue), (5) negative course change (orange)

For reliable pattern detection a set of fuzzy rules is used. E.g. steady motion will only seldom be represented as a sequence of exactly the same velocity in raw GPS data. Due to various reasons, minor variances in the calculated velocity occur. By applying fuzzy rules, a sequence of GPS points representing nearly the same velocity can be mapped to the semantic class of *steady motion*. As Table 2 reveals, a good empirically founded threshold between steady motion and acceleration is an acceleration rate of +-0.3m/s^2. Due to fuzzy rules, all values within this threshold (-0.3m/s^2 and +0.3m/s^2) are mapped to steady motion.

Despite of fuzzy matching some unrealistic classification remains. Figure 2 shows a longer sequence of steady motion, which is interrupted by negative acceleration for three times (red and blue patterns on the left). Since the interruptions only last for 1 second, we assume that the classification is not a realistic since 1 second slowdowns are not the expected behaviour of a train. Since the interruptions are located at the beginning of a longer phase of slowdown, values slightly below or above the threshold cause varying classifications. A solution to this problem could be to detect longer sub-sequences and subsume short sub-sequences (e.g. shorter than 3 seconds) in the longer ones.

3.2 Semantic classification of course changes

As reported by Zheng et al. (2008), the rate of course changes varies significantly with the mode of transportation. Moving with higher velocity allows a lower rate of course changes. Moreover, absolute course changes within short time periods may be significantly higher when moving with lower velocity, e.g. a pedestrian may accomplish 90 degree turns within a short time period (within 1 or 2 seconds). A train however is bound to the railway infrastructure and is limited in its course changes. For detecting course changes we use the mean course change rate in one time period (+- °/s).

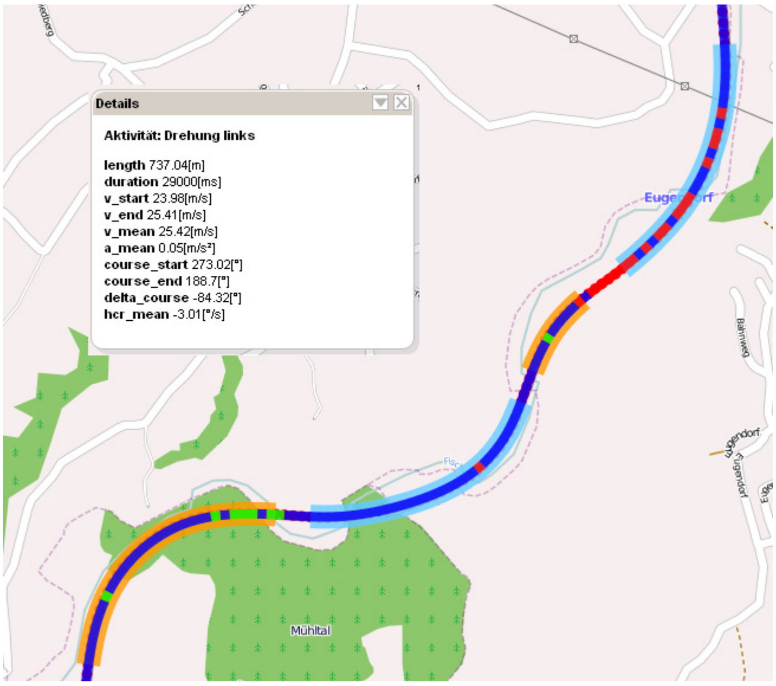

Figure 3: Detection of course changes. (1) Positive course change (light blue), (2) negative course change (orange)

Figure 3 shows course changes in a GPS trace collected during a train ride. For semantic classification of course changes, we again use a fuzzy rule set. When the course change rate within one time period (the change of degrees within one second) is above or below a threshold value, the GPS point is matched either to the semantic class *positive course change* or *negative course change*. In comparison to the motion detection, for the course change detection we use velocity-dependent thresholds. Velocity-dependent thresholds means, that for trace segments with low velocity, the threshold for course changes is set significantly higher compared to trace segments with higher velocity. In other words, as higher the velocity gets, as lower the threshold for course changes is set. The empirically estimated range of thresholds are +-5°/s (pedestrians moving at 1.5m/s), +-1°/s (car moving at 20m/s), +-0.4°/s (high speed train moving at 55m/s). The relatively high threshold for slow movement is necessary to deal with inaccuracies of GPS positioning. While moving slowly, minimal inaccuracy of positions has higher impact on the course change rate compared to moving at a higher speed (since the distance between two measurement is longer at higher speed and occurring positions errors have relatively less impact on calculated velocity and course). Increasing the threshold to +-5°/s allows detection of significant course changes only (Figure 4).

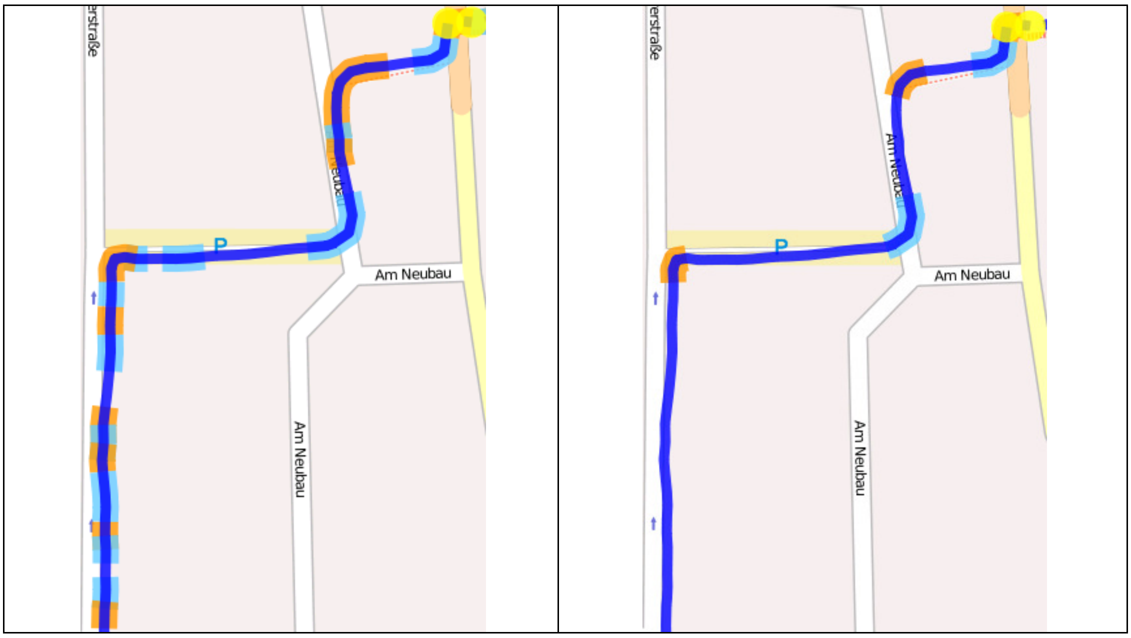

Figure 4: Course change classification (orange and light blue shading) in a GPS trace recorded during a walk with a threshold of +-1°/s (left) and +-5°/s (right). The higher threshold results in the detection of only realistic course changes.

4 Conclusions and open issues

In the paper we introduced an approach to semantic processing of GPS traces by classifying sequences of GPS points with motion and course change patterns. Although the reported approach works basically well, a number of open issues remain.

Firstly, the threshold-based classification of motion only reveals basic, coarse-grained motion patterns. A more fine-grained classification of steady motion as well as acceleration is expected. For further processing we find it useful not to rely on fixed thresholds, but to apply cluster analysis e.g. to reveal fine-grained clusters of steady motion. In first tests we found density-based clustering (e.g. ST-DBSCAN) (Birant and Kut 2007, Tietbohl et al. 2008) to be a well suited method for sub- classification. At time of writing we are not able to provide final results, but promising preliminary results. The same method can be applied to cluster acceleration (mean acceleration), standstills (using time spans) and course changes (using mean change rates).

A second open issue concerns the semantic classification of multimodal traces (traces including motion with more than one means of transport). The first open question is the automatic detection of transport modes and change points. Although some authors tackled this problem (e.g. Zheng et al. 2008) we could not find robust methods for automatically detecting any transportation mode. Authors either rely on map matching (e.g. by matching standstills to bus stops and thus deriving the transport mode bus) or only focus on some modes, but miss others. Currently we are working on robust algorithms, using case-based reasoning as well as cluster analysis.

Another open question is how to deal with severe errors of GPS positioning during walks as well as at interchange points. Since interchange points are typically buildings with underground passages, a high number of GPS errors occur. Although we apply methods for error correction, the semantic processing of such trace parts is a challenging question (Figure 5).

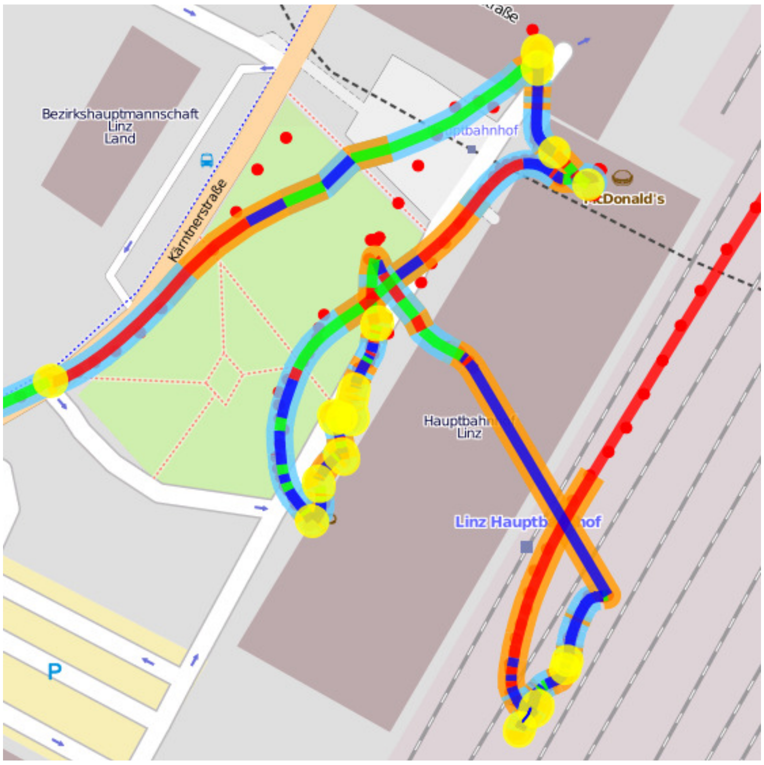

Figure 5: An example of semantic processing of a GPS trace in situations of modal change. The trace shows a change from train to car with a walk through the interchange building.

Acknowledgements

The project HOTSPOT is funded by the Austrian Federal Ministry for Transport, Innovation and Technology (BMVIT) within the programme "IV2Splus – intelligent transport systems and services".

References

Ashbrook D and Starner T, 2003, Using GPS to learn significant locations and predict movement across multiple users. *Personal & Ubiquitous Computing*, 7: 275-286.

Birant D and Kut A, 2007, ST-DBSCAN: An algorithm for clustering spatial-temporal data. *Data and Knowledge Engineering*, 60(1):208-221.

Cao L and Krumm J, 2009, From GPS Traces to a Routable Road Map, In: *Proceedings of 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Seattle, USA, 3-12.

Laube P, Imfeld S and Weibel R, 2005, Discovering relative motion patterns in groups of moving point objects. In: *International Journal of Geographic Information Science*, 19(6):639-668.

Steiner I, Clemens B, Werfelli S, Dell'Omo G, Valenti P, Troester G, Wolfer DP, Lipp H-P: A GPS logger and software for analysis of homing in pigeons and small mammals. *Physiology & Behavior*, 2000, 71:589-596.

Tietbohl A, Bogorny V, Kuijpers B and Alvares L, 2008, A clustering based approach for discovering interesting places in trajectories. In: *Proceedings of ACM Symposium on Applied Computing, Advances in Spatial and Image-Based Information Systems Track*.

Wolf J, Guensler R and Bachman W, 2001, Elimination of the travel diary: an experiment to derive trip purpose from GPS travel data. In: *Proceedings of the Transportation Research Board 80th annual meeting*, Washington, DC.

Zheng Y, Li Q, Chen Y, Xie X and Ma W Y, 2008, Understanding Mobility based on GPS data. In: *Proceedings of UbiComp'08*, Seoul, Korea, 312-321.