Learning Pedestrian Profiles from Movement Trajectories

David Jonietz

Institute for Cartography and Geoinformation, ETH Zurich

Extended Abstract

Pedestrians are highly heterogeneous with regards to their physical capabilities and preferences, which in turn determine their individual infrastructural needs (Saelens et al. 2003, Millonig 2006). Other than in current pedestrian navigation systems, therefore, the possibility to compute personalized walking routes would be desirable, a precondition of which, however, is to derive detailed user profiles (Gartner et al. 2011, Jonietz 2016a).

In the context of personalized routing, historical GPS-trajectories of car drivers have been used to mine their individual preferences and restrictions (e.g. Letchner et al. 2006, Yang et al. 2015). This work, however, builds on a previous study which proposed to infer the infrastructural needs of individual pedestrians by extracting selected environmental properties of their previously visited areas, as identified from movement trajectories (Jonietz 2016a). Here, the focus is placed on gradually refining a user profile based on new input movement data, in particular the problem of handling new value inputs which, depending on their plausibility, might entail different procedures:

- an adaptation of the user profile (valid refinement or update of infrastructural need)
- an adaption of the context model (changed infrastructural need, however likely related to temporary contextual conditions, e.g. the weather or time of day)
- no action (implausible input value likely caused by measurement error)

LBS 2016
13th Conference on Location-Based Services
Vienna, 14–16 November 2016

Published in "Proceedings of the 13th International Conference on Location-Based Services", edited by Georg Gartner and Haosheng Huang, LBS 2016, 14-16 November 2016, Vienna, Austria.

Based on Jonietz (2016a), table 1 lists relevant pedestrian infrastructural needs which can be computed for each walking trip based on its trajectory and, if rated plausible, should be stored in a user profile for future personalized route calculation.

Feature	Description
maxBenchDist	max. walking distance between potential resting places
minBuildingDist	min. space needed for walking between buildings or other obstacles
maxVisibleAesthBuildings	visual exposure to aesthetical buildings (max. number of simultaneously visible aesthetical buildings)
maxLightDist	max. acceptable distance to the nearest streetlight
minCyclelaneDist	min. safety distance buffer kept to cycle lanes
minStreetDist	min. safety distance buffer kept to streets
maxHeight	max. step height (relative height difference)
maxSlope	max. acceptable slope
maxVisibleStreet	visual exposure to street traffic (normalized max. amount of visible street)
minSurface	min. acceptable surface quality type of footpaths
maxVisibleTrees	visual exposure to urban greenery (max. number of simultaneously visible trees)

Table 1. Features to be Extracted from Movement Trajectories

For refining a user profile, checking the plausibility of new value inputs is non-trivial. A new input value with regards to a hard restriction (e.g. the maximum tolerated slope), for instance, can either represent a valid new information about the user's capabilities, or simply be the result of uncertain positioning or errors in the underlying spatial data layers which, if erroneously being represented in the user profile, could result in the recommendation of routes which are in fact inaccessible. Furthermore, explicitly identifying the influence of temporary contextual factors, such as the time of day with regards to the feature *maxLightDist*, on behavioral adaptions would be desirable for a later context-aware routing.

For these purposes, we propose the general procedure as illustrated in figure 1. First, in a heuristics-based noise filtering step, the focus is put merely on the geometric characteristics of the raw input tracking data. In a velocity-based filtering, the travel speed between consecutive tracking points is computed, and, if it exceeds a predefined realistic threshold value for pedestrian movement, filtered accordingly (c.f. Zheng 2015). In a second filtering step, an overlay procedure is employed in a Geographic Information System (GIS) to identify and delete all tracking points which are located outside of the walkable area, and therefore clearly due to positional error.

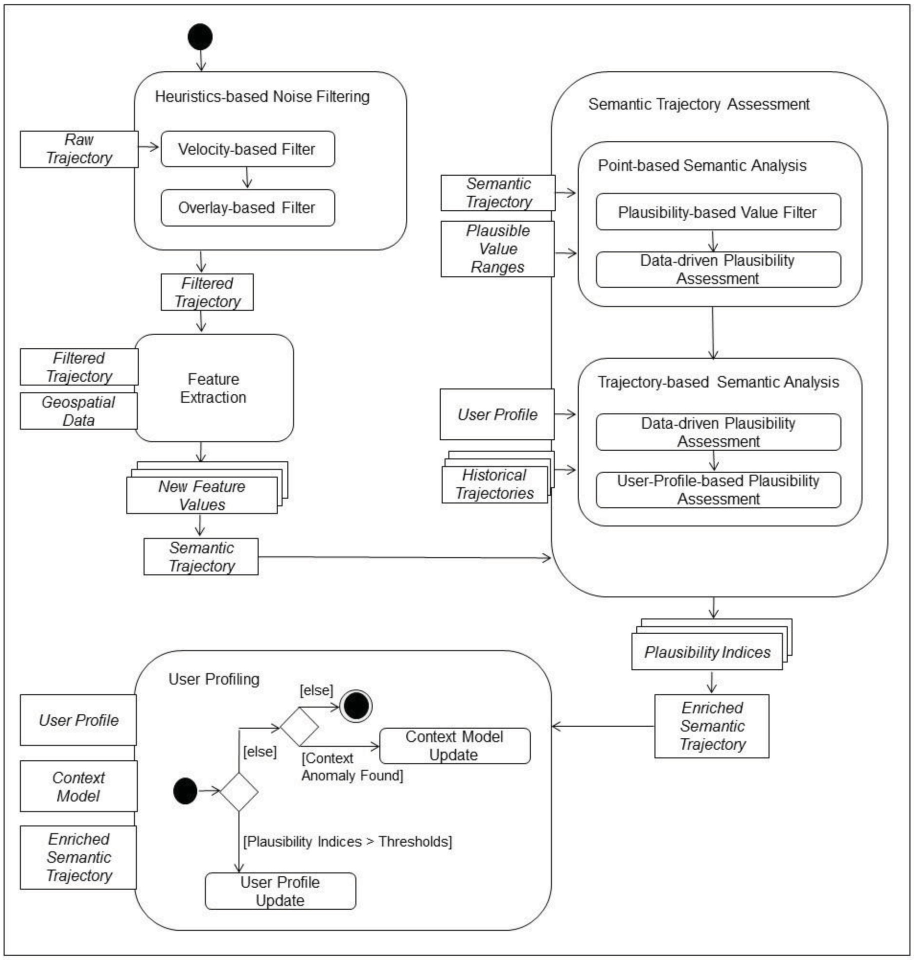

Figure 1. General Framework for Handling Input Data

Then, as described in Jonietz (2016a), the new feature values are calculated by overlaying with pre-computed raster layers for each environmental attribute as listed in table 1, extracting the underlying values and attributing the property values to the tracking points. These semantically enriched trajectories are then assessed for their plausibility. This procedure is based on the following assumptions:

- the new value must lie within a predefined range of plausible domain values

- the larger the deviation of a new value from the corresponding values of the temporal neighbors of its respective tracking point, the lower its plausibility

- the larger the deviation of a new value from all previously recorded values for this user (historical trajectories), the lower its plausibility

- the larger the deviation of a new value from the one currently stored in the user profile, the lower its plausibility

The first assumption guides the implementation of a plausibility-based value filter, which filters tracking points with feature values outside of predefined plausible ranges, e.g. slope values above 40%. In the following step, a sliding window approach is used to compare each feature value with the corresponding values of its temporal neighbors, meaning all preceding or subsequent tracking points within a certain time interval, and compute according outlier scores (c.f. Chandola et al. 2009). Points with scores exceeding a preset threshold are filtered from the dataset. This step is motivated by the general assumption that the location of temporally adjacent tracking points should be set in comparable environmental conditions, which, however, clearly depends on the temporal resolution of the position measurements.

After the level of the individual tracking points, the trajectory as a whole is examined. First, the index values as shown in table 1 are computed. Then, in a data-driven plausibility assessment, the derived values for each trajectory are compared to the movement history of the same user, i.e. the entire set of his or her previously recorded trajectories. Comparable to the previous step, an outlier or plausibility score is calculated for each trajectory based on the relative deviation of its index values with regards to the whole set. At this stage, however, no filtering takes place, but instead, the computed plausibility scores are attributed to the respective trajectories. A second score is calculated in the final step of user-profile-based plausibility scoring, where the deviation of the index values of each trajectory with regards to the values currently stored in the user profile is assessed.

The semantic trajectories, further enriched with the plausibility scores, finally provide the input for the user profiling process, where a threshold-based decision process evaluates whether to update the user profile based on a preset ruleset, check for possible explaining anomalies in the contextual data (e.g. weather data at time of walking) and, if necessary, update the context model, a step which is currently omitted, or disregard the new value as a probable measurement error. In an iterative process, thus, the user profile can be further refined and updated with regards to new input data.

At the present stage, the described framework has been implemented in Python, and tested with a set of movement trajectories derived from an agent-based pedestrian simulation which is further described in Jonietz (2016b), and features virtual pedestrians with differing infrastructural needs. In the context of this study, the simulation model has been extended to include simulated positional errors of the tracking mechanism. The generated trajectories and user profiles are stored in a PostGIS database. In this preliminary testing scenario, the proposed method shows promising potential for filtering inaccurately positioned tracking points as well as derive the individual infrastructural needs of pedestrian agents. For future work, it is planned to implement the context model updating process, test the model with real pedestrian data, as well as embed it within a personalized pedestrian routing system. In this context, the quality of the derived user profile could be assessed by comparing the recommended routes with the actual choice a user made, and use this evaluation for further refining the sketched procedure.

References

Chandola V, Banerjee A, Kumar V (2009) Anomaly detection: A survey. ACM Computing Surveys 41(3): 15

Gartner G, Huang H, Millonig A, Schmidt M, Ortag F (2011) Human-centered mobile pedestrian navigation systems. Mitteilungen der Österreichischen Geographischen Gesellschaft 153: 237-250

Jonietz D (2016a) Personalizing walkability: a concept for pedestrian needs profiling based on movement trajectories. In: Sarjakoski T, Santos MY, Sarjakoski LT (eds) Geospatial data in a changing world. Springer, Heidelberg, 279-297

Jonietz D (2016b) From space to place: A computational model of functional place. Doctoral thesis. University of Augsburg, Germany.

Letchner J, Krumm J, Horvitz E (2006) Trip router with individualized preferences (TRIP): incorporating personalization into route planning. In: Porter B (ed) Proceedings of the 18th conference on Innovative applications of artificial intelligence - Volume 2 (IAAI'06), AAAI Press, 1795-1800

Millonig A (2006) Routennetze für mobile Fußgänger-Navigationsanwendungen: ein neuer Ansatz für die Optimierung auf Basis von quantitativen Bewegungsdaten. In: CORP 2006, Vienna, Austria

Saelens BE, Sallis JF, Frank LD (2003) Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. Annals of Behavioural Medicine 25(2):80-91

Yang B, Guo C, Ma Y, Jensen CS (2015) Toward personalized, context-aware routing. The VLDB Journal 24(2): 297-318

Zheng Y (2015) Trajectory data mining: An overview. ACM Transactions on Intelligent Systems and Technology 6(3): 29