AGILE: GIScience Series, 5, 25, 2024. https://doi.org/10.5194/agile-giss-5-25-2024
Proceedings of the 27th AGILE Conference on Geographic Information Science, 4–7 Sept. 2024.
Editors: Alison Heppenstall, Mingshu Wang, Urska Demsar, Rob Lemmens, and Jing Yao.
This contribution underwent peer review based on a full paper submission.

Exploring User Semantic Annotation from Trajectories in the Scenario of Shared Locations

Jens Golze[1], Monika Sester[1], Udo Feuerhake[1], and Claus Brenner[1]

[1]Institute of Cartography and Geoinformatics, Leibniz University Hannover, Appelstraße 9A, Hannover 30167, Germany

Correspondence: Jens Golze (golze@ikg.uni-hannover.de)

Abstract. Over the past decade(s), collecting spatio-temporal data has become easier due to technological advancements and more user-friendly collection processes. Additionally, government agencies, companies, and open data projects have made general environmental data, such as points of interest or land use coverage, more freely available. Scientific studies can combine this spatio-temporal and non-spatial data to analyze different types of human mobility data. The results of these studies are relevant to transportation and urban planning, as similar information is typically collected by means of surveys. However, deriving relevant information from Global Navigation Satellite System (GNSS) trajectories remains challenging due to inaccuracies in the positioning and the unavailability of groundtruth information regarding individual user location semantics (e.g. *home place*, *work place*, *leisure place* or others). This work presents a semantic location annotation approach based on a Hidden Markov Model and the Viterbi optimization algorithm. The model includes location emissions to account for the general usage of a particular location. The annotations are applied to the clustered stop points that identify regions of special interest to individual users in a trajectory data set. The proposed approach demonstrates that the adapted Viterbi optimization can assign the most probable and meaningful semantic labels to the user's sequences and provides insights on the underlying regions of special interest.

Keywords. location of interest, semantic place annotation, GPS trajectories, clustering, HMM, viterbi optimization, spatio-temporal analysis

1 Introduction

The availability of smartphones and reliable connections to Global Navigation Satellite Systems (GNSS) has led to the creation and use of location-based services (LBS). LBS is now ubiquitous in everyday life, with applications in navigation, social media, and many smartphone-based services.

Due to the widespread use and easy accessibility of location-based information through various devices, the collection of spatio-temporal data has reached its peak in terms of quantity. The type and amount of data collected depend on the method used for a particular application. For instance, geo-tagged images or text posts, such as social media contributions, are used to share geo-locations with smaller or larger social networks. Location check-ins, although fixed in the spatial domain, can provide sequences of visited locations for individuals, as well as insights into user preferences or group dynamics. Another widespread source of spatio-temporal data is Global Positioning System (GPS) trajectories. Due to the ease of use, particularly of smartphones as low-cost sensors, large quantities of GPS data are collected daily worldwide. Companies are interested in location information of users to gain insights into people's movement behaviors, connections with other individuals and general user interests.

Furthermore, the trajectory data can be used for the purpose of automatic spatial information acquisition in the form of context generation. This can be, for example, the detection of changes in road networks or general road network updates (Gao et al., 2021; Tang et al., 2019). Beyond that, the detection and classification of road intersection regulator rules (Zourlidou et al., 2022a; Golze et al., 2020). Additional types of information can be, for example, the detection of road roughness (Hiremath et al., 2021; Wage and Sester, 2021) or the determination of the traffic flow (Tu et al., 2021; Li et al., 2021) or the quantification of traffic flow delays caused by traffic accidents (Golze et al., 2021).

Large trajectory sets, which contain many users with many trips over an extended period of time, can provide valuable insights not only on where individuals have moved, but also enable predictions of their future movements. Thus, it could be possible to uncover spatio-temporal behaviors of the users. Examples in this context are the understanding of (typical) commuting patterns (Qin et al., 2018) or using restricted user groups for the extraction of the spatio-temporal behavior of tourists in certain cities (Yao et al., 2021).

In this paper, the approach of (Golze and Sester, 2024) is extended to allow that the same location could give rise to different activities. Imagine a shopping center, which triggers the activity *shopping* for users who pursue this activity, however, for a worker at the shopping center, it would trigger the activity *working*.

The paper is structured as follows. In Section 2, an overview of the related work is presented. Section 3 describes the methodology of this work. In Section 4, the results are presented. The discussion, the main findings and the outlook are given in Section 5.

2 Related Work

GPS trajectory data is being used to reveal movement patterns and to highlight frequently visited locations. A common first step is to identify and extract these places from the GPS trajectories, known as *stop points* or *interesting places* (Feuerhake et al., 2011).

The semantics of interesting places are assigned to an artificial point, e.g. a cluster center (point) of a collection of stop points, rather than to an actual measured GPS point. Furthermore, semantic information is directly attached to user trajectories, called *semantic trajectories*. Ying et al. (2010) propose a similarity measure for these semantically enriched trajectories embedded in the context of travel recommendation services by analyzing the semantic movement behavior of users. The work of Lin et al. (2018) explores the temporal characteristics of trajectories to extract semantic mobility profiles to detect dwell regions. They present a similarity measure to compare different users for further investigation. Another approach for working with the semantics of spatial points is presented in the *habit2vec* framework by Cao et al. (2020). Their approach is based on the principle of vector representation of the Natural Language Processing (NLP) domain. They encode the semantic and temporal information using representation learning to study typical user habits of the tracked user base. Andrienko and Andrienko (2018) take a different approach using a graph-based representation of location region (states), overlaid with semantics as an additional layer of information. The work of Chen and Poorthuis (2021) presents the implementation of an R package that provides four different approaches from the literature to determine home locations from spatio-temporal data. They also address privacy considerations and research ethics when using LBS data.

Lv et al. (2016) present a hierarchical clustering approach to merge trajectory stops enriched with temporal and spatial descriptive features to assign pre-defined semantic type labels to these stop clusters. The work of Yang et al. (2018) aims extracting home and work locations. Therefore, they use clustering to aggregate stop points derived from GPS trajectories. These are additionally enriched with temporal signatures, assuming a regular movement behavior per user (e.g. full-time worker). Recently, published work by Cheng et al. (2022) presents an approach using semantic location annotation based on Point of Interest (POI) and Area of Interest (AOI). Furthermore, they include the temporal domain in their approach represented by features such as *length of stay* and *visiting time*. Li et al. (2023) introduce a life pattern clustering approach that identifies groups of users with similar life patterns, called *meta-graph*. They implemented a rule-based system for the detected interesting places of the users, highlighting places such as home places, work places, night places, day places, or other places.

The work of Wu et al. (2021) performs pattern mining on historical trajectory data to extract social relationships and preferences of users. They use features derived from the points of interest (context specific) and the behavior of interest (user specific). Another work by Hosseinpoor Milaghardan et al. (2021) focuses on the identification of trajectory activity patterns by combining geometric clusters extracted from stop points with their associated activity sequence. The geometric data are provided as large trajectory data sets in Korea. While spatio-temporal analysis of movement patterns is of great interest, especially in the field of traffic and commuting smaller spatial and temporal frames can also be considered for example, in football analysis (Feuerhake, 2016). Moreover, spatio-temporal pattern mining is applied to non-trajectory data such as origin-destination (OD) data. Recently, publications dealing with OD pattern mining focus on shared mobility patterns of e-scooters and their general usage (Heumann et al., 2021).

The work of Xu et al. (2021) proposes a Hidden Markov Model (HMM) to assign the most probable semantic annotation labels (*home place* and *work place*) to clusters of stop points extracted from GPS trajectories using OPTICS. The model employs a dynamic radius for selecting points of interest (POIs) based on their influence radius at a given time of day. This way, the approach indirectly incorporates temporal aspects into their analysis. A similar approach is presented by Golze and Sester (2024), while focusing on a broader list of features from the contextual, temporal and network domain.

3 Methodology

The primary goal of the presented method is to determine the hidden states, *living*, *working*, *shopping*, *leisure* and *unknown*, under the consideration of observed features: POI categories, weekday, arrival time, stop duration.

The methodology is structured in multiple parts. In case the data sets are based on raw GPS trajectories, the locations of interest first have to be extracted using stop point extraction and interesting places retrieval, as described in

(Golze and Sester, 2024). Subsequently, contextual features are extracted at these stop points, which relies on a selection of POIs in the vicinity of the stop point. Those locations are potential locations which the user has visited for his/her activities, which ultimately, we are interested in. In this process, POIs with a larger distance to the stop point get a lower weight, as calculated with the IDW (inverse distance weighting).

Overall, the semantic annotation regarding the POI categories is applied on three generalization levels. The first level (Level-1) is the most general, while the subsequent levels include more detailed information (e.g. Level-1: *Public Service*, Level-2: *Supply*, Level-3: *Money*). Table 1 provides an overview of the categories included in each of the generalization levels. Level-1 semantic categories were used for the annotations in this work.

Table 1. Enumeration of the semantic POI categories (from OSM) for each level of detail.

Level	Semantic Categories
Level-1	Recreation, Commercial, Residential, Transportation, Public Services
Level-2	Water, Greenland, Touristic, Activity, Agriculture, Industry, Shopping, Food, Office, Accommodation, Nightlife, Living, Memorial, Religion, Railway-System, Street-Network, Airport, Education, Security, Social, Health, Supply
Level-3	River, Lake, Park, Garden, Camping, Forest, Zoo/Park, Attraction, Museum, Sport, Entertainment, Agriculture, Industry, Supermarket, Mall, Retail, Restaurant, Fast-Food, Office, Hotel, Motel, Apartment, Bar, Disco, Kiosk, XXX, Living, Graveyard, Church, Train-Station, Railway, Waterway, Harbour, Road, Junction, Parking, Busstop, Airport, University, School, Kindergarden, Library, Police, Fire Department, Military, Law, Community, Hospital, Doctor, Energy, Water, Post & Telecommunication, Environment, Money

The temporal features describe the time-related characteristics of the stop points. Due to the nature of time, a discretization into periods of the day is defined as given in Table 2. These periods of the day are used to determine the temporal feature *period of day* for each stop point. Additionally, the features extracted from the timestamps include the *weekday* and *stop duration* at the respective stop point.

3.1 Semantic Label Prediction

In the presented approach a Hidden Markov Model is used to estimate the (hidden) semantic labels (living, working, shopping, leisure and unknown) while the observations are the estimated features for each stop point. The general HMM consists of transition and emission probability matrices and is used by e.g. Viterbi optimization. These matrices are populated by observations and common knowledge about human mobility behavior. The transition matrix describes the probability of the most likely next state X_2, given a current state X_1. The emission matrix reflects the probabilities of the most likely observed feature characteristics, given a certain state X. Table 3 shows an example for the emission matrix of weekdays, indicating, e.g. that leisure activities have a higher probability at the weekends.

Table 2. Discretization of time of the day into periods of the day.

Hour of day	Period of day
01 - 04	Night 1
05 - 08	Morning 1
09 - 11	Morning 2
12 - 14	Noon
15 - 17	Afternoon
18 - 21	Evening
22 - 00	Night 2

Table 3. An example emission matrix for the weekdays, starting with Monday (coded as 0) and the hidden activity labels.

	0	1	2	3	4	5	6
living	.15	.15	.15	.15	.15	.13	.12
working	.18	.18	.18	.18	.18	.09	.01
shopping	.13	.13	.13	.13	.2	.3	.005
leisure	.1	.1	.1	.1	.1	.2	.3
unknown	.1	.1	.1	.1	.1	.25	.25

The HMM and the annotated stop sequences are the input for the semantic labeling process. This process is implemented using the Viterbi algorithm (Viterbi, 1967). The Viterbi algorithm is a widely used method for solving optimization problems involving hidden states. It is based on the principles of dynamic programming.

The methodology adjusts the HMM in the way that location-specific activity emissions (L_As) are introduced to reflect related activities at a specific location l. These emissions, L_As, are added to the product of the emissions obtained by the observed features. The following formula shows that the L_A is an additional factor dependent on the location l used in the HMM.

$$\hat{P}(X) = \prod_{i}^{n} P(X_i \mid X_{i-1}) \times \prod_{i}^{n} P\left(feat._i \mid X_i\right) \times L_A(l) \tag{1}$$

In order to differentiate them from the hidden semantic label X, they are given different names, e.g. *Home*, *Work*, *Shop*, *Leisure*, *Unknown*. The procedure in this work is following an alternating estimation approach which consists of three phases:

1. **Initial Phase**:
   In the initial phase, the Viterbi algorithm is applied to

predict the most likely hidden semantic labels (X) for each stop in the given user sequences, while storing the state probabilities of each stop point at a given location l, referred to as $I_A(l)$. The location emissions (L_As) in the initial phase are set as equal for each possible hidden state. In the course of the next steps, these emissions are adjusted.

2. **Update Phase**:
   The update phase collects all I_As according to their associated location l to re-estimate the location emission L_A based on the following formula:

$$L_A(l) = \sum_{i}^{m} \frac{I_{A_i}(l)}{m} \tag{2}$$

3. **Alternation Phase**:
   The updated L_As are used in the alternation phase, where the Viterbi algorithm is used again to repeat the prediction of the hidden semantic labels X and to estimate the new I_As.

When the predicted semantic labels X are the same as in the previous iteration for all stop points in all sequences, the whole process ends. The output contains the most probable labels for each sequence (user specific) as well as the L_A of each location l reflecting the general usage of the location given by the probabilities of the activities.

4 Experiment and Results

4.1 Data sets

Synthetic Data Set To analyze the effects of the modeling structure developed in this paper, a synthetic data set was manually created. This data set allows for different activities at certain locations and was generated by simulating the daily activity behavior of people who frequently visit known locations. The data set consists of three users, two of whom work full-time and one who works part-time. User 1 and 2 are friends who occasionally visit each other's homes. User 1 shops at the workplace of user 3. Additionally, all users visit different places for shopping and during leisure time. The data set includes the movement behavior of the three users over a typical week. Due to the nature of the synthetic data set, the groundtruth activities at all locations are known for all users.

GPS Trajectory Data Set. The GPS trajectory data set consists of 700 time-ordered sequences of GPS points collected by a single user from December 2017 to March 2019 in Hanover, Germany. The trajectories, ranging from 5 to 14 kilometers in length, represent the daily routines of a full-time working family member.

4.2 Experiments

The methodology was first applied to the synthetic data set, which was already in the required input form for the optimization method, without the need for pre-processing. The results indicate that the introduction of location emissions and the alternating approach affects the probabilities for activities at locations with multiple purposes through visits from different users (see Figure 2). Obviously, in the course of the iterations, most likely assignments are adjusted, e.g. most likely activity *Shop* (green) changed to *Work* (orange).

As exact groundtruth information for the synthetic data set is available, it can be compared to the most likely assignment given by the Viterbi. The number of correctly and incorrectly classified locations are shown in the context of locations shared with other users (see Table 4).

Table 4. Overview of the correctly (T) and incorrectly (F) predicted sequence labels, broken down by the different users (U1, U2, U3), and highlighting the locations shared with other users in the synthetic data set. On the diagonal are the locations that were shared only with itself.

User	# of Visits	Shared w U1 **T** | **F**	Shared w U2 **T** | **F**	Shared w U3 **T** | **F**
U1	29	**6** | 1	**18** | 2	**0** | 2
U2	26	**18** | 2	**6** | 0	0 | 0
U3	26	**5** | 0	0 | 0	**17** | 4

Combining this both information from Figure 2 and Table 4 it can be concluded that most wrongly assigned labels occur due to sharing of the location with another user performing a different activity at the same location. This can be seen in Figure 1, where the center location (red dot) reflects this scenario.

The presented approach was also applied to the GPS trajectory data set containing only one person. It can be observed from the location emissions that the state probabilities are either slightly strengthened or weakened with an increasing number of iterations. When comparing the predicted semantic labels for the locations with the groundtruth information it is shown that most predictions are correct - except for location 3 where *Shop* was assigned instead of *Leisure*, and location 8 where *Home* was assigned instead of *Leisure* - still, *Leisure* also received the second largest prediction value. Table 5 presents the results of the previous work by Golze and Sester (2024), demonstrating that the new version performs similarly, although the label *Leisure* was split into *Shop* and *Leisure*. However, the new version also provides the probabilities of the resulting predictions as a way for reasoning.

In summary, the advantage of the new approach is that the predicted labels can incorporate different, individual-dependent, activities at one location, which are stored and updated in the probabilities given by the L_As.

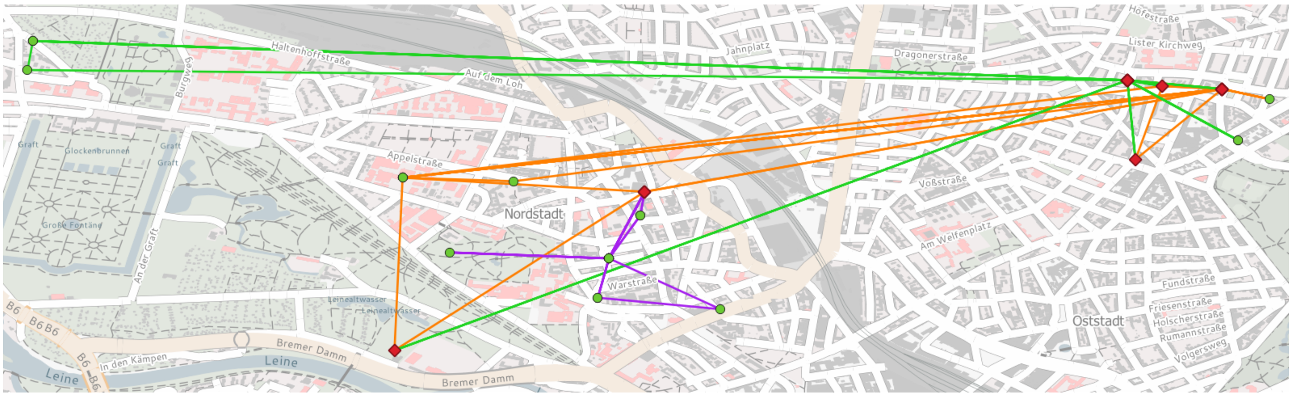

Figure 1. Synthetic data set with the three users (green, orange and purple) The locations are indicated, green circles if only one user or nor activity conflict and red squares if otherwise. Basemap: © GeoBasis-DE / BKG CC BY 4.0

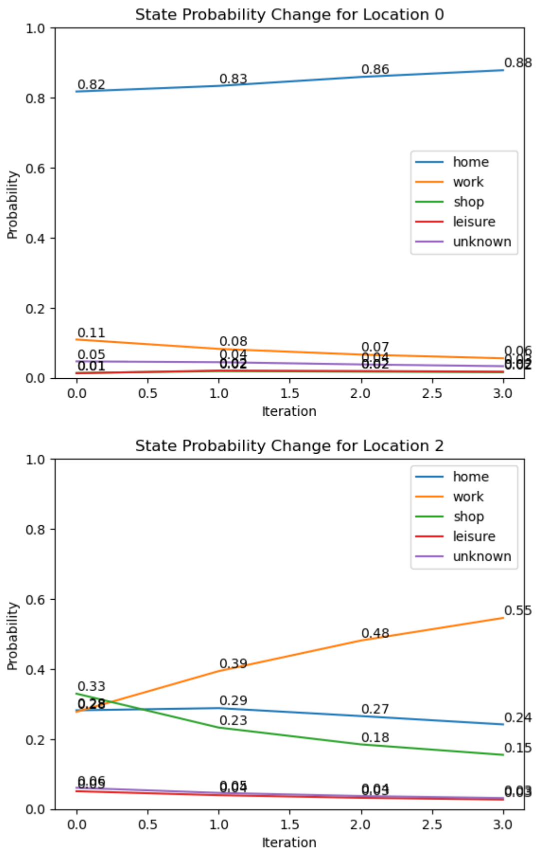

Figure 2. State probability change for L_A of location 0 (top), only visited by a single user, and location 2 (bottom), visited by multiple users of the synthetic data set.

Table 5. Comparison of the top three L_As for the locations (highest value highlighted in boldface) compared with the groundtruth labels of the GPS trajectory data set and the results of the work by Golze and Sester (2024).

Location	groundtruth	[1]	L_A
0	Shop	Leisure time	$Work$ 0.05 **Shop 0.86** $Leisure$ 0.05
1	Shop	Leisure time	$Home$ 0.06 $Work$ 0.28 **Shop 0.59**
2	Home	Home	**Home 0.94** $Work$ 0.02 $Unknown$ 0.02
3	Leisure	Work	$Home$ 0.07 $Work$ 0.18 **Shop 0.72**
4	Shop	Leisure time	$Home$ 0.06 $Work$ 0.03 **Shop 0.85**
5	Work	Work	$Home$ 0.05 **Work 0.92** $Shop$ 0.01
6	Work	Work	$Home$ 0.06 **Work 0.88** $Shop$ 0.02
7	Leisure	Leisure time	$Home$ 0.20 $Work$ 0.13 **Leisure 0.62**
8	Leisure	Leisure time	**Home 0.70** $Work$ 0.07 $Leisure$ 0.17

[1] Work of Golze and Sester (2024); here "leisure time" also contains the activity *shopping*.

4.3 Data and Software Availability

The original full GPS trajectory data set is not available due to privacy restrictions. An anonymized version of the data set is available by Zourlidou et al. (2022b). The developed code (containing the synthetic data set) can be provided on request.

5 Discussion and Outlook

Possible wrong assignments in the results for the GPS trajectory data could be due to different factors: it may be affected not only by GPS inaccuracy but also by the selection of the parameters for stop point extraction and the clustering method used. The resulting stop point clusters may cover a smaller or larger area, introducing potential ambiguity due to a wider range of possible activities.

In future work, the level of POI categories will be increased to provide more detailed information of the environment (compare Table 1. This way, the emission matrix for the POI categories will reflect these details by providing a more granular view of what can be expected to be observed in a certain state. Similarly, the set of hidden semantic labels can be provide a finer differentiation for more insights into the users' activities.

Another aspect is estimating the transition and emission probability matrices using the Baum-Welch algorithm (Welch, 2003). This learning procedure, however, would require a large amount of groundtruth annotations, which are rare for GPS trajectories. To address this issue, we will first use a synthetic dataset generated through simulation, which has a much larger scope than the previous synthetic dataset. The simulated data is derived from a user survey on user mobility behavior in Germany.

In the future, the approach will be applied on a larger GPS trajectory data set that fulfills the same prerequisites as the synthetic data set, e.g. multiple users that eventually interact with the same locations for the same or different activities.

Author contributions. The first author contributed to conceptualization, methodology, investigation, validation, formal analysis, writing-original draft preparation and visualization. The second author contributed to supervision, writing-original draft preparation and funding acquisition. The third and fourth author contributed in discussion and modeling of the methodology.

Competing interests. The authors declare no conflict of interest.

References

Andrienko, N. and Andrienko, G.: State Transition Graphs for Semantic Analysis of Movement Behaviours, Information Visualization, 17, 41–65, https://doi.org/10.1177/1473871617692841, 2018.

Cao, H., Xu, F., Sankaranarayanan, J., Li, Y., and Samet, H.: Habit2vec: Trajectory Semantic Embedding for Living Pattern Recognition in Population, IEEE Transactions on Mobile Computing, 19, 1096–1108, https://doi.org/10.1109/TMC.2019.2902403, 2020.

Chen, Q. and Poorthuis, A.: Identifying home locations in human mobility data: an open-source R package for comparison and reproducibility, International Journal of Geographical Information Science, 35, 1425–1448, https://doi.org/10.1080/13658816.2021.1887489, 2021.

Cheng, J., Zhang, X., Luo, P., Huang, J., and Huang, J.: An unsupervised approach for semantic place annotation of trajectories based on the prior probability, Information Sciences, 607, 1311–1327, https://doi.org/10.1016/j.ins.2022.06.034, 2022.

Feuerhake, U.: Recognition of Repetitive Movement Patterns—The Case of Football Analysis, ISPRS International Journal of Geo-Information, 5, https://doi.org/10.3390/ijgi5110208, 2016.

Feuerhake, U., Kuntzsch, C., and Sester, M.: Finding interesting places and characteristic patterns in spatio-temporal trajectories, in: 8th LBS symposium, 2011.

Gao, S., Li, M., Rao, J., Mai, G., Prestby, T., Marks, J., and Hu, Y.: Automatic Urban Road Network Extraction From Massive GPS Trajectories of Taxis, pp. 261–283, Springer International Publishing, Cham, https://doi.org/10.1007/978-3-030-55462-0_11, 2021.

Golze, J. and Sester, M.: Determining user specific semantics of locations extracted from trajectory data, Transportation Research Procedia, 78, 215–221, https://doi.org/10.1016/j.trpro.2024.02.028, 25th Euro Working Group on Transportation Meeting, 2024.

Golze, J., Zourlidou, S., and Sester, M.: Traffic Regulator Detection Using GPS Trajectories, KN - Journal of Cartography and Geographic Information, 16, 321, https://doi.org/10.1007/s42489-020-00048-x, 2020.

Golze, J., Feuerhake, U., Koetsier, C., and Sester, M.: IMPACT ANALYSIS OF ACCIDENTS ON THE TRAFFIC FLOW BASED ON MASSIVE FLOATING CAR DATA, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B4-2021, 95–102, https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-95-2021, 2021.

Heumann, M., Kraschewski, T., Brauner, T., Tilch, L., and Breitner, M. H.: A Spatiotemporal Study and Location-Specific Trip Pattern Categorization of Shared E-Scooter Usage, Sustainability, 13, https://doi.org/10.3390/su132212527, 2021.

Hiremath, R., Malshikare, K., Mahajan, M., and Kulkarni, R. V.: A Smart App for Pothole Detection Using Yolo Model, in: ICT Analysis and Applications, edited by Fong, S., Dey, N., and Joshi, A., pp. 155–164, Springer Singapore, Singapore, 2021.

Hosseinpoor Milaghardan, A., Ali Abbaspour, R., Claramunt, C., and Chehreghan, A.: An activity-based framework for detecting human movement patterns in an urban environment, Transactions in GIS, 25, 1825–1848, https://doi.org/10.1111/tgis.12749, 2021.

Li, J., Boonaert, J., Doniec, A., and Lozenguez, G.: Multi-models machine learning methods for traffic flow estimation from Floating Car Data, Transportation Research Part C: Emerging Technologies, 132, 103 389, https://doi.org/10.1016/j.trc.2021.103389, 2021.

Li, W., Zhang, H., Chen, J., Li, P., Yao, Y., Shi, X., Shibasaki, M., Kobayashi, H. H., Song, X., and Shibasaki, R.: Metagraph-Based Life Pattern Clustering With Big Human Mobility Data, IEEE Transactions on Big Data, 9, 227–240, https://doi.org/10.1109/TBDATA.2022.3155752, 2023.

Lin, Z., Zeng, Q., Duan, H., and Lu, F.: Finding similar users from GPS data based on assignment problem, in: Proceedings of the 4th International Conference on Communication and Information Processing, edited by Ben-Othman, J., Yu, H., Unger, H., and Arai, M., pp. 283–288, ACM, New York, NY, USA, https://doi.org/10.1145/3290420.3290470, 2018.

Lv, M., Chen, L., Xu, Z., Li, Y., and Chen, G.: The discovery of personally semantic places based on trajectory data mining, Neurocomputing, 173, 1142–1153, https://doi.org/10.1016/j.neucom.2015.08.071, 2016.

Qin, T., Shangguan, W., Song, G., and Tang, J.: Spatio-Temporal Routine Mining on Mobile Phone Data, ACM Trans. Knowl. Discov. Data, 12, https://doi.org/10.1145/3201577, 2018.

Tang, J., Deng, M., Huang, J., Liu, H., and Chen, X.: An Automatic Method for Detection and Update of Additive Changes in Road Network with GPS Trajectory Data, ISPRS International Journal of Geo-Information, 8, https://doi.org/10.3390/ijgi8090411, 2019.

Tu, W., Xiao, F., Li, L., and Fu, L.: Estimating traffic flow states with smart phone sensor data, Transportation Research Part C: Emerging Technologies, 126, 103 062, https://doi.org/10.1016/j.trc.2021.103062, 2021.

Viterbi, A.: Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, IEEE Transactions on Information Theory, 13, 260–269, https://doi.org/10.1109/TIT.1967.1054010, 1967.

Wage, O. and Sester, M.: JOINT ESTIMATION OF ROAD ROUGHNESS FROM CROWD-SOURCED BICYCLE ACCELERATION MEASUREMENTS, ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-4-2021, 89–96, https://doi.org/10.5194/isprs-annals-V-4-2021-89-2021, 2021.

Welch, L. R.: Hidden Markov Models and the Baum-Welch Algorithm, IEEE Information Theory Society Newsletter, 53, 2003.

Wu, S. X., Wu, Z., Zhu, W., Yang, X., and Li, Y.: Mining Trajectory Patterns with Point-of-Interest and Behavior-of-Interest, in: 2021 IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–6, https://doi.org/10.1109/ICCWorkshops50388.2021.9473612, 2021.

Xu, H., Ye, Z., and Jiao, M.: A POI classification semantic annotation algorithm based on HMM, in: International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2021), edited by Wu, F., Liu, J., and Chen, Y., p. 93, SPIE, https://doi.org/10.1117/12.2631298, 2021.

Yang, M., Cheng, C., and Chen, B.: Mining Individual Similarity by Assessing Interactions with Personally Significant Places from GPS Trajectories, ISPRS International Journal of Geo-Information, 7, 126, https://doi.org/10.3390/ijgi7030126, 2018.

Yao, Q., Shi, Y., Li, H., Wen, J., Xi, J., and Wang, Q.: Understanding the Tourists' Spatio-Temporal Behavior Using Open GPS Trajectory Data: A Case Study of Yuanmingyuan Park (Beijing, China), Sustainability, 13, https://doi.org/10.3390/su13010094, 2021.

Ying, J. J.-C., Lu, E. H.-C., Lee, W.-C., Weng, T.-C., and Tseng, V. S.: Mining User Similarity from Semantic Trajectories, ACM Conferences, ACM, New York, NY, https://doi.org/10.1145/1867699, 2010.

Zourlidou, S., Golze, J., and Sester, M.: Traffic Regulation Recognition using Crowd-Sensed GPS and Map Data: a Hybrid Approach, AGILE: GIScience Series, 3, 22, https://doi.org/10.5194/agile-giss-3-22-2022, 2022a.

Zourlidou, S., Golze, J., and Sester, M.: GPS Trajectory Dataset of the Region of Hannover, Germany, https://doi.org/10.25835/9BIDQXVL, 2022b.