Characterization of Traveler Types Using Check-in Data from Location-Based Social Networks

Linus W. Dietz, Rinita Roy, and Wolfgang Wörndl

Technical University of Munich, Department of Informatics, Garching, Germany
linus.dietz@tum.de, rinita.roy@tum.de, woerndl@in.tum.de

Abstract. Characterizing types of travelers can serve as a foundation for tourism recommender systems. In this paper, we present an approach to identify traveler types by analyzing check-in data from location-based social networks. We show how to segment 33 million Foursquare check-ins of 266,909 users into 23,340 foreign trips based on mobility patterns of travelers. Then we apply hierarchical clustering to identify distinct groups of travelers using features such as travel duration, countries visited, and distance from home. Our results revealed four clusters of traveler types that can be interpreted.

Keywords: Data mining, Cluster analysis, Human mobility patterns, Tourism, Recommender systems

1 Introduction

Traveling is a highly emotional and personal endeavor which is of great importance to a person's education, personal well-being, and also social status. Choosing a destination and planning, however, is a complex task for an inexperienced person involving many unknowns, such as the expected costs, the quality of experience and the risks involved. Traditionally, travel agencies supported users in the selection and booking progress, but nowadays users rather inform themselves using online information sources and 75% of millennials prefer to travel independently on their own itinerary, rather than in a packaged tour [1]. This gives rise to destination recommender systems that support the user in making informed decisions where to travel to based on their budget and the personalized preferences [9]. These, however, need to be elicited efficiently – to provide a nice user experience – and effectively – to actually improve recommendation accuracy. Traditionally, destination recommender systems often perform a coarse-grained preference elicitation, e.g., by binary indications of interest for activities [15]. Others leverage more sophisticated models, such as the *Seven Factor Model* [21] of tourists' behavioral patterns that can be used to provide high-quality, personalized recommendations [26]. In this paper, we propose an alternative model for traveler characterization by analyzing their mobility patterns derived from location-based social network (LBSN) data. By analyzing the itineraries of tourists, we can derive their travel mobility patterns and therefore get an impression of what kind of traveler they are. But what kind of travelers are there? Literature provides several answers to that question, but none of these are backed by actual observation. Motivated by this, we analyze the check-in behavior of Foursquare users over the time of 18 months. We segment the check-in stream of individual users into trips, which we characterize solely using spatio-temporal information derived from the underlying mobility pattern we captured using the check-ins. Through cluster analysis, an unsupervised learning method, we find distinct, coherent groups of trips, each corresponding to one type of travel.

The main contributions of this paper are therefore the characterization of trips generated from LBSN check-in data, the selection of features for the clustering, and finally, the description of the identified trip types. Thus, we provide an alternative characterization of trips and travelers that is backed with data. This work provides novel insights into the travel and tourism domain and has practical applications in the preference elicitation and user modeling of tourist recommender systems. The approach is also applicable to data stemming from other sources than LBSN data. The data model employed is suited to work with fine-grained mobility data, such as raw GPS trajectories [10]. Our focus is nevertheless on LBSN data, since the granularity is sufficient for destination recommendation and it is more readily available.

The remainder of this paper is structured as follows: After an overview of existing literature in Section 2, we describe our methods focusing on the feature selection process for clustering in Section 3. Then, we describe the identified trip types in Section 4 and draw our conclusions not without pointing out future work in Section 5.

2 Related Work

The goal of this research is to find distinct trip types to augment the personalization of destination recommender systems. In this section, we provide some background on tourist recommender systems, approaches to characterize travelers and research on the analysis of human mobility.

2.1 Tourist Recommender Systems

Recommender systems for tourism are a well-established tool for prospective travelers due to the substantive complexities of planning an independent trip and the huge economic importance of the travel and tourism domain. The biggest economic players in the area of tourist recommender systems recommend hotels, restaurants, and flights, while the academic community is rather concerned with destination recommendation on different levels, i.e., tourist packs, single attractions, and composite trips [5]. Since these items are often not as well defined as hotels or restaurants, collaborative filtering methods are not commonplace; instead, content-based and knowledge-based recommendation techniques are employed [6]. To facilitate the content-based paradigm, a mapping of user needs to item attributes is required. This is usually done using a common categorization, which allows to calculate a similarity measure for the ranking of recommendations. It is a non-trivial task, since it requires reliable information about both the user and the items. Another option is to characterize users based on their past trips [11], or to find a more elaborate mapping between classes of users and destinations [26].

2.2 Tourist Roles

Tourists are assigned to various roles in literature. In 1963, Bhatia [4] characterized travel into two basic types – for leisure, or for business. Cohen [8] labeled tourists with four different social roles, i.e., *'the organized mass tourist'*, *'the individual mass tourist'*, *'the explorer'*, and *'the drifter'*. Pearce [24] used fuzzy set theory to yield 15 different travel roles. McKercher [20] classified tourists based on the importance of cultural motives when deciding which destination to visit and the depth of cultural experience gathered by the tourist. Yiannakis and Gibson [32] observed which roles – they identify 17 – are enacted by people when they go on travel associated with different psychological needs.

Observing this collection of various tourist roles in literature, it is quite unclear which characterization of tourist roles would be useful to improve destination recommendation. More importantly, none of these categorizations has been validated with data [21], so it is unclear if they are actually observable on travelers. Facing this challenge, Neidhardt et al. [21] developed the *Seven Factor Model* of tourist behavioral patterns based on the *Big Five Factor Model* [19] from personality psychology and a factor analysis of Gibson and Yiannakis' 17 tourist roles [12,32]. Having a destination recommender system in mind, they approached the preference elicitation process using a set of classified images. Instead of asking the user of their destination recommender to fill out a questionnaire, they presented her a selection of pictures from which she should pick the most appealing ones. The classification of these pictures was derived from the preferences of travelers, whose classification along the *Seven Factors* has been previously determined using a questionnaire. The user's selection of images defines a personalized mixture of taste model that enables content-based recommendation of points of interest, which have been rated by experts along the *Seven Factors*. Continuing this strand of research, Sertkan et al. [26] used unsupervised learning for clustering 561 tourist destinations based on 18 motivational and seven geographical attributes from a rich commercial data set. Using an expert mapping of the *Seven Factors* to these destinations, they could distill associations between destination attributes and the *Seven Factor Model* that refers to travel behaviors. While being scientifically sound, the *Seven Factor Model* relies heavily on expert knowledge, which is a drawback if this information is not available or hard to obtain. Also, we feel that these behavioral patterns are important when people talk about tourism, but there is no evidence how they relate to users' actual travel behavior in the field. Based on these considerations, we employ a different approach to determine behavioral patterns of tourists. We investigate whether it is possible to find evidence for a number of distinct tourist types analyzing the mobility patterns of travelers.

2.3 Human Mobility Analysis

The analysis of human mobility gives insights into various aspects of humans' everyday life and is especially useful when characterizing humans while performing activities that involve movement, such as travel. Traditionally, data sources like mobile phone communication records [13,30], Wi-Fi usage [33] and raw GPS trajectories [35] were used. Given the availability of LBSN data, which enriches the pure mobility trace with further information, such as users' posts and social network, much research has been done analyzing mobility via Twitter, Foursquare and others [16].

One research objective is the predictability of human mobility. Song et al. identified that the individual mobility patterns follow reproducible scaling laws [27] and also described the limits of the extent human mobility can be predicted [28]. More recently, Ouyang et al. [23] analyzed mobility data to predict human trajectories to understand human mobility patterns using a deep learning framework.

Correlating mobility with social activity using LBSN data provides interesting insights into our society. Noulas et al. [22] used Twitter and Foursquare data to form a relationship between location data and human activities. Cheng et al. [7] reveal recurring daily and weekly patterns of activity, while Wang et al. [29], find a positive pairwise correlation between social connectedness, i.e., the strength of interactions, and mobility.

LBSN mobility data has been used to improve recommender systems [3]. Zheng et al. [34] used spatial co-occurrences that can also be used to identify similar users and generate implicit ratings for collaborative filtering algorithms. Hawelka et al. [14] found similar patterns of tourists globally and then compared the result with worldwide tourism statistics and commonly used human mobility models, whereas Bao et al. [2] matched travelers in a foreign city to local experts based on their respective home behavior to improve the accuracy of a point of interest recommender. Finally, Hsieh et al. [17] use past LBSN data to recommend trips within cities determining the popularity, proper time of day to visit, the transit time between venues and the best order to visit the places. Recently, Dietz et al. [10] proposed a metric-based approach to derive foreign trips from LBSN data and analyze tourist mobility patterns with the goal of investigating the popularity and the co-occurrences of tourist destinations in composite trips. We follow this approach to mine the trips and use clustering algorithms to identify distinct groups of trips and travelers.

3 From Check-ins to Human Mobility Patterns

LBSN data, i.e., geo-tagged posts, gives an incomplete view of a user's mobility, since we only know a user's location when she decides to share it. However, given the prevalent use of LBSNs on mobile devices, users often leave a continuous spatio-temporal trace behind them. Combining them to a trajectory, we can characterize the mobility of a user with innate metrics.

3.1 Trip Mining from LBSN Data

For example, if a user tweets using a mobile device and decides to enable the *"Tweet with a location"* feature, an indication of her presence at a specific location is established. Similarly, if a user checks in at Foursquare venues, it is known that she has been at the venue at a specific time. Since we are interested in the user's behavior when traveling abroad, i.e., outside her home country, we determine the user's home country using the *"Plurality"* strategy, i.e., the country with the highest number of check-ins [18]. Now we can segment the check-in stream into periods in the home country and travel. A trip is defined as consecutive check-ins abroad before returning home . We omit all trips that are shorter than a week to eliminate the typical business trips. Furthermore, we discard trips where the user has not checked-in often enough by requiring a minimum check-in density of 0.2, meaning that on average the user must have checked-in at least once in five days [10]. Using this methodology, we mine 23,340 trips from a Foursquare dataset of 266,909 users from April 2012 to September 2013 [31].

3.2 Choosing Features for Clustering

These trips are characterized by several features derived from the check-in stream.

1. **Travel duration:** number of days between the first and the last check-in of the trip
2. **Countries visited:** number of countries visited
3. **Check-in frequency:** number of check-ins made during the trip divided by the travel duration. Note that this is different from the aforementioned check-in density, which only counts the days with a check-in
4. **Check-in duration:** the mean time between two consecutive check-ins
5. **Check-in distance:** the mean distance between two consecutive check-ins
6. **Radius of gyration:** the mean distance between the mean location of the trip and all check-in locations
7. **Displacement:** distance between the users home location and the mean position for the places visited during the trips
8. **Check-in time:** the average time of day of the check-ins, signifying a time of activity
9. **Day length:** the mean length of days, i.e., the time the sun being above the horizon for the place at the specific date
10. **Activities:** The types of venues according to Foursquare's categorization[1]

Since we want to capture the underlying phenomenon of travel behavior, we discard the check-in frequencies and the check-in duration from the clustering, as these metrics capture the quality of the data, instead of providing insights in the underlying mobility of the traveler. Although the information about activities could potentially give interesting insights into what the tourists do during their trips, we also exclude them because these features are very biased toward the

[1] See `https://developer.foursquare.com/docs/api/venues/categories`

Table 1: Feature statistic for 23340 trips

Feature	Mean	Median	Maximum	St. Dev.
Travel duration	16.44	10.00	429.00	21.27
Countries visited	1.61	1.00	24.00	1.13
Check-in distance	394.87	25.44	18203.37	1081.23
Radius of gyration	721.91	65.27	13762.70	1540.85
Displacement	5002.42	3708.58	19933.28	4308.79
Check-in time	14.84	14.99	23.58	2.65
Day length	12.71	12.62	18.84	2.00

check-in behavior of Foursquare users, and are not generally available in other data sources. This leaves us with seven candidates features listed in Table 1. Besides the day length, the check-in time and the mean displacement all features follow an exponential distribution with a long tail, which limits the interpretability using mean and median.

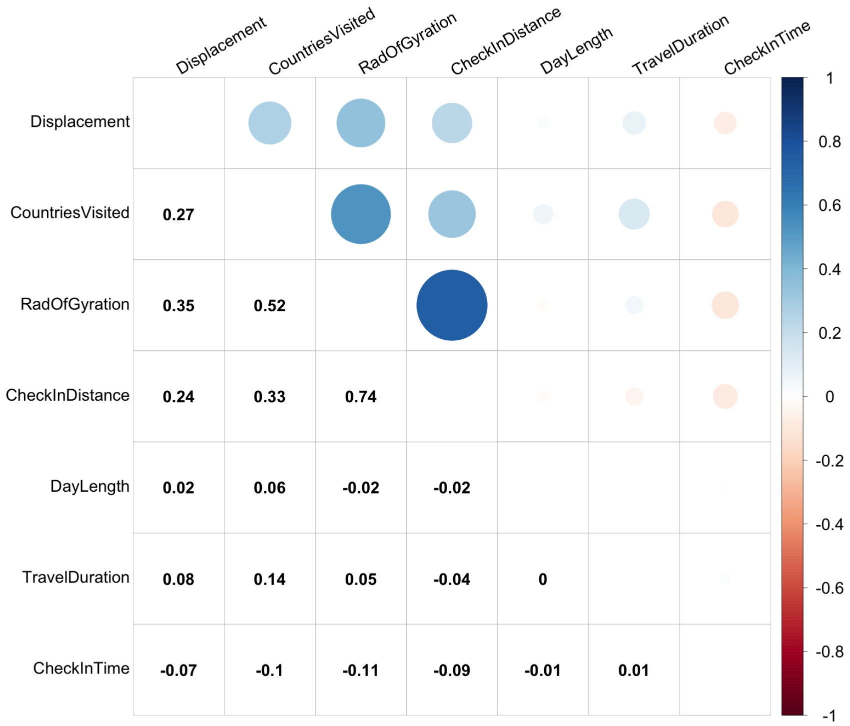


Fig. 1: Correlation analysis of candidate features

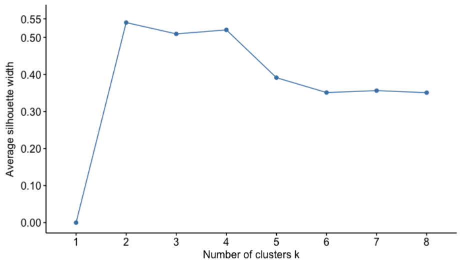
Before running the actual clustering algorithm, we perform a correlation analysis of the remaining candidate features. As can be seen in Figure 1, we obtain a strong positive relationship between the radius of gyration and the check-in distance with a coefficient of 0.74. Therefore, we remove the latter as the radius of gyration is more independent of the check-in behavior of the user. The other notable correlation with 0.52 is between the radius of gyration and the number of countries visited. However, in this case, we keep both, since a medium correlation is actually ideal for a clustering algorithm to find structure in the data. Day length and the check-in time are barely correlated ($\leq \pm 0.11$) to the other features. Therefore, we also exclude them, as they will not improve the segregation in the clustering algorithm. This leaves us with four remaining features: travel duration, countries visited, displacement, and the radius of gyration.

4 Clustering Results

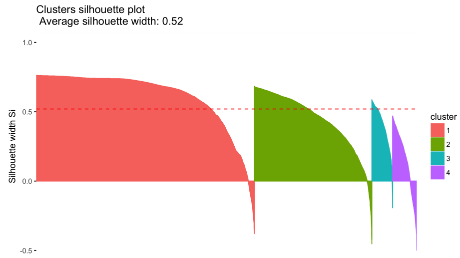
Cluster analysis is the task of finding groups of data objects, where each group comprises similar objects, whereas the groups themselves are dissimilar to each other. It is useful to uncover structure within unlabeled data and is therefore categorized as unsupervised learning. Since our features are all numeric, we are not restricted in the usage of the mainstream clustering algorithms, such as k-means, k-medoids, or hierarchical clustering. To avoid bias due to different value ranges of the features, we normalize the features using min-max scaling. Further, the distance measure in all experiments is the Euclidean distance in the four-dimensional feature space.

We have experimented with these algorithms to find a suitable number of clusters. By varying the number of expected clusters between two and eight, we evaluated the quality of the determined clusters by the within-cluster sums of squares and the average silhouette [25]. The silhouette measures how well a data object fits into its labeled cluster as opposed to all other clusters. Therefore, it is a robust and easy to interpret method that gives a broad overview of the overall solution quality, but also information of each data object. In our analysis, the general tendency throughout all algorithms was that 3–5 clusters seemed to be a reasonable choice with clusters in the area of 0.4–0.5 silhouette width.

A closer analysis reveals hierarchical clustering to achieve the highest average silhouette widths, followed by k-means and k-medoids. Thus, we use the hierarchical clustering algorithm to find the optimal number of clusters. As can be seen in Figure 2a, the average silhouette of the hierarchical clustering is in the area of slightly above 0.5 for two to four clusters and plummets below 0.4 at five clusters and more. Based on this, we select four clusters as our final solution. The evaluation using within-cluster sums of squares also support this result, although the evidence for four clusters is not as clear as with the silhouette method. A quick glance at the silhouette plots of the four clusters in Figure 2b, shows that they are mostly in the positive area, with few trips per cluster being in the negative area. This confirms a reasonably good clustering result.



(a) Average silhouette width



(b) Silhouette plot for the four clusters

Fig. 2: Choosing the right number of clusters

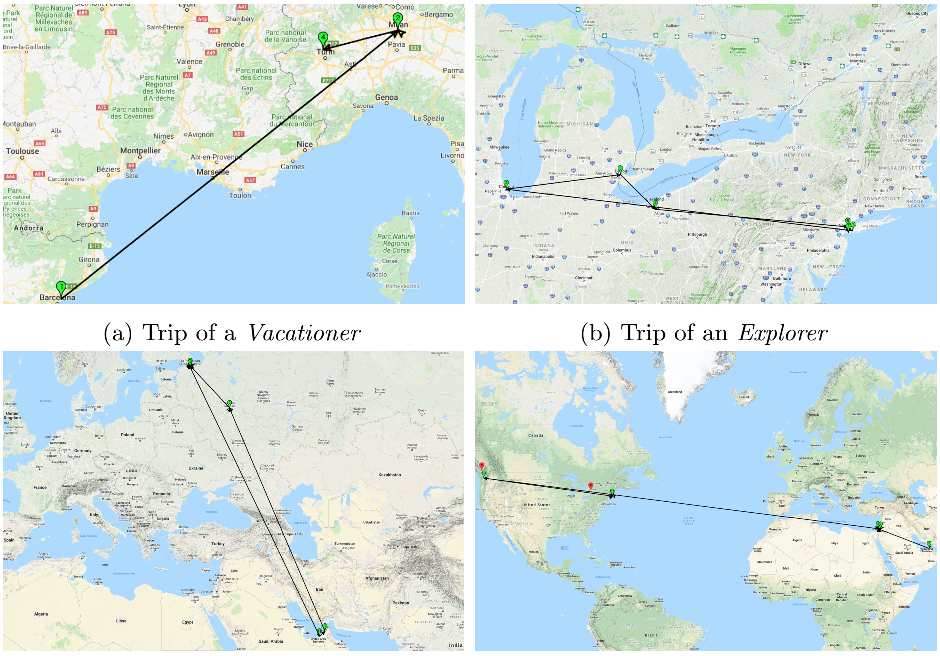
Table 2: Mean values and standard deviations of features per determined cluster

Feature	**Vacationers**	**Explorers**	**Voyagers**	**Globetrotters**
Trips	13373	7213	1277	1477
Silhouette width	0.62	0.44	0.39	0.15
Travel duration	15.96/20.97	16.68/23.23	18.73/17.65	17.60/16.04
Countries visited	1.36/0.72	1.66/1.09	1.87/1.25	3.30/2.22
Radius of gyration	310/733.15	418/586.50	1535/1725.77	5230/2348.73
Displacement	1798/1508	8450/1851.82	14594/1555.88	8887/2878.50

The four identified clusters each represent a distinct type of trip. Unsurprisingly, the numbers of trips per cluster is not uniformly distributed, with 57% of the trips residing in the first cluster, 31% in the second, 5% in the third, and 6% in the fourth cluster. Furthermore, the travel duration was relatively equal among the clusters with a mean value between two and three weeks. This was somewhat expected, as it showed relatively low correlations (up to 0.14) in the correlation plot of Figure 1. We nonetheless used this feature, because the duration of travels is important to consider even with the rather low correlation.

Table 2 illustrates the mean values of the trips per cluster. Recalling the existing research on tourist roles from Section 2, it seems that we have identified a novel characterization of trip types. To get an impression of the clusters, we describe their characteristics and visualize them using a trip taken from close proximity of the cluster center.

"Vacationers" (cluster 1) and *"Explorers"*, (cluster 2) contain trips to one or two neighboring countries. The mean radius of gyration is very low, with 310km and 418km, respectively. The notable feature that distinguishes them is that the former is near to the traveler's home, whereas the latter is on another continent. This can be seen in the mean displacement of 8450km, which is about the distance from Cyprus to Montréal, Canada. The trip of a *Vacationer*, visualized in Figure 3a, was a seven days trip by a tourist from Turkey covering Barcelona, Spain, and Milan, Italy before flying out from Turin airport. Conversely, an



(a) Trip of a *Vacationer*

(b) Trip of an *Explorer*

(c) Trip of a *Voyager*

(d) Trip of a *Globetrotter*

Fig. 3: Characteristic trips for each cluster. The numbered green markers show the consecutive check-ins of the trip. The red markers show check-ins within the user's detected home country.

Explorer's trip was an eight-day round trip by a Russian tourist to New York City, USA the Great Lakes area visiting Cleveland, Chicago, and Detroit on the way. The overall mean displacement was large, but the radius of gyration was relatively small, as the whole trip revolved within the north-east of the United States of America (see Figure 3b).

"Voyagers", the third cluster, comprises extremely distant trips, traveling slightly less than two countries on average. The displacement from home is on average as far as the distance between Cyprus and Melbourne, Australia. Also, the travelers move slightly more at their destinations with a radius of gyration of about 1500km. The exemplary trip is of an Australian, who traveled to Russia with stopovers in the United Arab Emirates. The travel duration was ten days, of which at least five were spent in Russia for visiting major attractions within the inner cities of Moscow and St. Petersburg. Figure 3c visualizes this itinerary.

Finally, *"Globetrotters"* are also distant trips to a different continent, but not as far from home as those of the previous cluster. This cluster stands out, because with a mean value of 5230km, the radius of gyration spans a very large area. Since these trips also comprise of an average of 3.3 countries, we can truly speak of world travels here. Our exemplary traveler from Vancouver, Canada went on travel for 20 days, covering four countries. Flying in to the Near East via the United Arab Emirates, she continues her travel after three days to visit Israel and Jordan for one week before arriving to Seattle the day after. However, the trip seemingly continues four days later with several check-ins at Harvard University, Boston, USA. In this case, we assume that this trip should have been split into two trips. The one to the Near East and the trip to Boston. Since there were only four days between these trips and no further check-ins at home, the algorithm couldn't perform this segmentation. Figure 3d shows the trip itinerary.

5 Conclusions & Future Work

In this paper, we have described an approach for finding tourist types using LBSN data. We segment the raw check-in stream of users into foreign trips and characterize it via features that capture the mobility patterns of the trip. Having a global destination recommender system in mind, we use these features to discard trips that are not of our interest. To filter out typical business travelers, this includes those that are shorter than seven days, and trips with insufficient data quality, for example, because the check-ins are not occurring frequently enough. To find a number of distinct trip types, we perform a cluster analysis of the remaining trips. Comparing several clustering algorithms, we achieve the best results with hierarchical clustering at a cutoff of four clusters. Evaluating the resulting clusters using the silhouette method, we find a reasonably good segregation of trips that we describe using exemplary visualizations of characteristic trips.

This method can be used to characterize travelers by aggregating the types of their past trips. Thus, it constitutes a contribution to preference elicitation and user modeling within the context of recommenders in tourism. Moreover, the analysis requires no user interaction, which is good for the user experience, and it is also computationally cheap. However, it requires access to the user's check-in history. This can be achieved by asking the user for access to their timeline in a LBSN they have been using, e.g., through a third-party Facebook or Twitter application. Obtaining the data in such a way, we can use a classifier that was trained with this paper's approach to classify the current user, thus providing more personalized recommendations.

In future, it would be interesting to expand this approach with more data stemming from different data sources. We think it would be worthwhile to compare mobility patterns and the resulting traveler types from different LBSNs, but also fine-grained GPS data. Since the cluster analysis strongly depends on the input data, deriving further features would be interesting. For example, we tried to infer the travel season using the day length feature, i.e., the time between sunrise and sunset. However, the cluster analysis showed that this feature is not useful for discriminating trips due to the low correlation with other features. Extending this with actual climate data of the check-in locations and dates, however, might be a fruitful approach. To further diversity the traveler types, one could also try to infer the cost of a trips from third party sources. Finally, the effects of the presented approach on recommender systems, i.e., the recommendation accuracy and user satisfaction, should be measured in a user study.

References

1. Airbnb. Airbnb and the rise of millennial travel. https://www.airbnbcitizen.com/airbnb-millennials-study-travel-more-important-than-saving-for-a-home, November 2016.
2. Jie Bao, Yu Zheng, and Mohamed F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '12, pages 199–208, New York, NY, USA, November 2012. ACM.
3. Jie Bao, Yu Zheng, David Wilkie, and Mohamed Mokbel. Recommendations in location-based social networks. *GeoInformatica*, 19(3):525–565, February 2015.
4. Arjun K. Bhatia. *International Tourism Management.* Sterling Publishers, 1963.
5. Joan Borràs, Antonio Moreno, and Aida Valls. Intelligent tourism recommender systems. *Expert Systems with Applications*, 41(16):7370–7389, November 2014.
6. Robin Burke and Maryam Ramezani. *Matching Recommendation Technologies and Domains*, pages 367–386. Springer US, Boston, MA, October 2011.
7. Zhiyuan Cheng, James Caverlee, Kyumin Lee, and Daniel Z. Sui. Exploring millions of footprints in location sharing services. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media.* AAAI, 2011.
8. Erik Cohen. Towards a sociology of international tourism. *Social Research*, 39(1):164–182, 1972.
9. Linus W. Dietz. Data-driven destination recommender systems. In *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*, UMAP '18, New York, NY, USA, July 2018. ACM.
10. Linus W. Dietz, Daniel Herzog, and Wolfgang Wörndl. Deriving tourist mobility patterns from check-in data. In *Proceedings of the WSDM 2018 Workshop on Learning from User Interactions*, Los Angeles, CA, USA, February 2018.
11. Linus W. Dietz and Achim Weimert. Recommending crowdsourced trips on wOndary. In *Proceedings of the RecSys Workshop on Recommenders in Tourism*, Vancouver, Canada, October 2018.
12. Heather Gibson and Andrew Yiannakis. Tourist roles: Needs and the lifecourse. *Annals of Tourism Research*, 29(2):358–383, April 2002.
13. Marta C. González, César A. Hidalgo, and Albert-László Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, June 2008.
14. Bartosz Hawelka, Izabela Sitko, Euro Beinat, Stanislav Sobolevsky, Pavlos Kazakopoulos, and Carlo Ratti. Geo-located twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, 41(3):260–271, February 2014.
15. Daniel Herzog and Wolfgang Wörndl. A travel recommender system for combining multiple travel regions to a composite trip. In *CBRecSys@RecSys*, volume 1245, pages 42–48, Foster City, CA, USA, 2014.
16. Andrea Hess, Karin Anna Hummel, Wilfried N. Gansterer, and Günter Haring. Data-driven human mobility modeling. *ACM Computing Surveys*, 48(3):1–39, December 2015.
17. Hsun-Ping Hsieh, Cheng-Te Li, and Shou-De Lin. Exploiting large-scale check-in data to recommend time-sensitive routes. In *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*, UrbComp '12, pages 55–62, New York, NY, USA, August 2012. ACM.

18. Ankit Kariryaa, Isaac Johnson, Johannes Schöning, and Brent Hecht. Defining and predicting the localness of volunteered geographic information using ground truth data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing System*, CHI'18. ACM, April 2018.
19. Robert R. McCrae and Oliver P. John. An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215, June 1992.
20. Bob McKercher. Towards a classification of cultural tourists. *International Journal of Tourism Research*, 4(1):29–38, January 2002.
21. Julia Neidhardt, Leonhard Seyfang, Rainer Schuster, and Hannes Werthner. A picture-based approach to recommender systems. *Journal of Information Technology & Tourism*, 15(1):49–69, September 2014.
22. Anastasios Noulas, Salvatore Scellato, Cecilia Mascolo, and Massimiliano Pontil. An empirical study of geographic user activity patterns in Foursquare. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. AAAI, 2011.
23. Xi Ouyang, Chaoyun Zhang, Pan Zhou, and Hao Jiang. Deepspace: An online deep learning framework for mobile big data to understand human mobility patterns. *CoRR*, abs/1610.07009, 2016.
24. Philip L. Pearce. *The Social Psychology of Tourist Behavior*. International Series in Experimental Social Psychology, 1982.
25. Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Computational and Applied Mathematics*, 20:53–65, 1987.
26. Mete Sertkan, Julia Neidhardt, and Hannes Werthner. Mapping of tourism destinations to travel behavioural patterns. In Brigitte Stangl and Juho Pesonen, editors, *Information and Communication Technologies in Tourism*, pages 422–434, Cham, CH, December 2017. Springer.
27. Chaoming Song, Tal Koren, Pu Wang, and Albert-László Barabási. Modeling the scaling properties of human mobility. *Nature Physics*, 6(10):818–823, September 2010.
28. Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. Limits of predictability in human mobility. *Science*, 327(5968):1018–1021, February 2010.
29. Dashun Wang, Dino Pedreschi, Chaoming Song, Fosca Giannotti, and Albert-László Barabási. Human mobility, social ties, and link prediction. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD'11, New York, NY, USA, August 2011. ACM.
30. Amy Wesolowski, Nathan Eagle, Andrew J. Tatem, David L. Smith, Abdisalan M. Noor, Robert W. Snow, and Caroline O. Buckee. Quantifying the impact of human mobility on Malaria. *Science*, 338(6104):267–270, October 2012.
31. Dingqi Yang, Daqing Zhang, and Bingqing Qu. Participatory cultural mapping based on collective behavior data in location-based social networks. *ACM Transactions on Intelligent Systems and Technology*, 7(3):1–23, January 2016.
32. Andrew Yiannakis and Heather Gibson. Roles tourists play. *Annals of Tourism Research*, 19(2):287–303, January 1992.
33. Yan Zhang, Lin Wang, Yi-Qing Zhang, and Xiang Li. Towards a temporal network analysis of interactive WiFi users. *Europhysics Letters*, 98(6), June 2012.
34. Yu Zheng and Xing Xie. Learning travel recommendations from user-generated GPS traces. *ACM Transactions on Intelligent Systems and Technology*, 2(1):1–29, January 2011.
35. Yu Zheng, Lizhu Zhang, Xing Xie, and Wei-Ying Ma. Mining interesting locations and travel sequences from GPS trajectories. In *Proceedings of the 18th International World Wide Web Conference*, WWW'09, New York, NY, USA, April 2009. ACM.