Context Incorporation in Cultural Path Recommendation Using Topic Modelling

Konstantinos Michalakis[1 ✉][0000−0002−5943−6613], Georgios Alexandridis[1,2][0000−0002−3611−8292], George Caridakis[1][0000−0001−9884−935X], and Phivos Mylonas[3][0000−0002−6916−3129]

[1] Cultural Technology Department, University of the Aegean, Mytilene, Greece
{kmichalak, gealexandri, gcari}@aegean.gr,
[2] School of Electrical & Computer Engineering, National Technical University of Athens, Zografou, Greece,
[3] Department of Informatics, Ionian University, Corfu, Greece
fmylonas@ionio.gr

Abstract. Even though path recommendation is a subject that has been vigorously studied, the majority of related work has been predominantly focused on travel and routing topics, with relatively minimal incorporation of cultural context. The latter issue is addressed in the current contribution through the proposition of a personalized, context-aware cultural path recommendation system, aiming at achieving an enhanced cultural experience for its users. More specifically, topic modelling is used to represent the landmarks, where each location is modeled as a distribution of latent topics that eventually describe its characteristics. The initial approach is subsequently extended through the fusion of contextual aspects that include visitor profile, their behavior during the visit and other environmental parameters that might affect the cultural experience. In this work, a subset of contextual aspects, consisting of the type of visited location and the time the visit occurred, is considered. The overall system is evaluated on a benchmark dataset in order to assess the effect of the contextual dimension in the produced recommendations.

Keywords: Personalized cultural user experience · Cultural path recommendation · Context-aware recommendation · Topic modelling.

1 Introduction

Cultural user experience is a subject that has recently gained enough popularity, despite requiring complex procedures of formalization and evaluation [15]. Apart from the standard complexity incurred by the delivery of a personalized experience, the content of cultural spaces is often enriched with characteristics associated with its origin and whose correlation to the experience itself is very tight. The extraction of those underlying aspects has not been adequately explored and while some route recommendation systems have been adapted for cultural visits, the cultural aspects are usually not integrated into the process.

Context integration, on the other hand, results in an enhanced perception of the current situation by the system, with data not directly related to the objects of interest. Typical contextual parameters include *location*, *time*, *type* of recommended object and *environmental conditions*, allowing for a more insightful interpretation of the surrounding environment. Most *context-aware* recommender systems apply context-driven querying and search approaches that require the matching of contextual data with resource metadata. On the contrary, contextual filtering and modeling are sparsely used [26].

This work addresses the aforementioned issues by introducing a personalized, context-aware and topic sensitive cultural path recommendation system. The proposed architecture combines content modeling and context-awareness into a unified approach that analyzes user behavior and enhances the recommendation process with the contextual parameters of time and location. At the core of the presented methodology lies *topic modeling*; a theoretical abstraction that conceptualizes the aspects of user visits to *Points of Interest* (POIs). Context-awareness is introduced into the model by formulating the relationship between POIs and time as one contextual parameter and the correlation between user and POI category as another.

2 Related Work

Recommender Systems (RSs) process user preference data in order to propose items ranging from products to paths or actions. RSs have been applied to a variety of domains and incorporate additional information sources, when available (e.g. from social networks [4]). In this sense, a popular extension is the inclusion of contextual data, usually in the form of spatial, environmental and behavioral parameters. This process adds context-awareness to RSs, resulting in a further optimized and personalized user experience, based on the current situation. Context-awareness may be combined in route RSs in location-aware environments, where a path of actions or sites to be visited is proposed. Research on such spatiotemporal modeling and prediction has been performed for both travelers [5, 27] and drivers [16].

Route RSs usually rely on mining techniques in order to discover usable information, such as user behavior and trajectory patterns [11], fastest path and route optimization based on user-specified destinations [20] and personalized route recommendation extracted from big trajectory data [9]. The recommendation process often requires the dynamic modelling of users (e.g. normal schedule, activity recognition); such functionality is incorporated in [22], where interactive multi-criteria techniques are adopted on personalized tours that combine user profile, preference and area characteristics. Multimodal information fusion may also be used in RS in order to enrich the acquired knowledge; e.g. a route recommendation utilizing geotagged images in an effort to probabilistically model user behavior is adopted in [17]. Research on the integration of social and crowdsourcing techniques for the improvement of recommendation performance has been conducted in [10, 24].

Context-awareness may be introduced to RSs either in a pre-filtering or a post-filtering fashion, depending on whether contextual processing occurs before or after the application of the recommendation algorithm. The two approaches are compared in [3], where it is concluded that no method outperforms the other and that their suitability is highly dependent on the application domain. A further classification of contextual information in RSs is proposed in [21], where the lack of studies in the non-representational views of context is also illustrated. Recent advances on *computational intelligence* have also fueled more efficient and more complex RSs that apply context-awareness with the use of *neural networks*, *fuzzy sets* and other similar machine learning techniques [1].

In general, context-aware RSs in the cultural heritage domain have found little applications so far and context is usually limited to spatial characteristics. In [6], a framework that manages heterogeneous multimedia data gathered from various web sources is suggested, which results in a context-aware recommendation process. *SmartMusem* [23] is a mobile RS that uses ontology-based reasoning, query expansion and context knowledge, achieving optimized performance. Finally, an ontology based pre-filtering and contextual processing post-filtering hybrid technique is used to provide optimized tour recommendations in [7].

3 User Modeling

A *user model* is a theoretical concept that tries to formulate a person's interests (motion patterns in-between POIs or landmarks in this case). Various approaches to user modeling exist, with some of them having already been presented in Section 2. In sites of cultural interest (museums, archaeological sites, cities, etc), the userbase and the available options are usually constrained and as a result the collected data tend to be of relatively small volume, especially when compared with the massive userbase and itemsets of other domains (movie or music recommendation). In the former cases, statistical models seem to be a good starting point for user modeling and this approach has been followed in this work. More specifically, *topic modelling*, a statistical approach to user modelling is presented in Section 3.1, while the aforementioned technique is extended through the fusion of contextual information in Section 3.2.

3.1 Topic Modeling

A *topic model* is a hierarchical probabilistic model that quantifies the relationship between users (*visitors*) and items (*landmarks* they have visited) through the notion of *topics* [8]. In this setting, visitor interest is expressed as mixture of topics, while each topic is modelled as a probability distribution over the landmarks. Of course, topics are not known in advance; in fact their estimation is the objective of the algorithm and for this reason they are considered to be the *latent features* of the model. In general, topic models have been used in the area of *information retrieval* [13] and in preference modelling and personalization [14, 2]. The most notable topic modelling techniques are *Latent Dirichlet Allocation*

(LDA) [8], *Latent Semantic Analysis* (LSA) and *Probabilistic* LSA (PLSA) [12], which is also the technique of choice in this contribution.

More formally, let L be the set of *points of interest* (POIs) of a cultural location, which can be of any type; constrained withing a building (e.g. museum), an open space (e.g. cultural site) or even a broader place (e.g. the historical center of a city). In either of these cases, POIs could be exhibits, landmarks or distinct places.

The set of visitors (users) is denoted as U and every visitor $u \in U$ is primarily characterized by a record $h^u \equiv (l_0, l_1, \ldots, l_{t-1})$, which is the sequence of POIs visited by u up until time $t-1$. Therefore, the objective of the model and the RS in general, is to predict the POIs the user is going to visit next.

Topic models address this issue by making the basic assumption that visits to future POIs are conditionally independent from the current record (history) h^u of u. Consequently, the probability $P(l_t|h^u)$ of u visiting POI l_t at time t, given h^u, is approximated according to Equation 1

$$P(l_t|h^u) = \sum_{z \in Z} P(l_t|z)P(z|h^u) \tag{1}$$

where $P(z|h^u)$ expresses the extend to which the "hidden" topic z is of interest to u while $P(l_t|z)$ quantifies how much POI l_t is described by topic z. In more simple words, $P(z|h^u)$ models visitor interest in this topic and $P(l_t|z)$ represents the coverage ("trend") of topic z.

As both probabilities of the right-hand side of Equation 1 cannot be determined analytically, they are approximated through *Expectation-Maximization* (EM). EM is an iterative procedure particularly useful in determining the maximum likelihood in statistical models dependent upon "latent" parameters, like this one. A typical EM iteration consists of two steps, the *Expectation Step*, where the posterior probability of each latent topic z is computed, given the parameters l_t, h^u of the model (Equation 2)

$$P(z|l, h^u) = \frac{P(z|h^u)P(l|z)}{\sum_{z' \in Z} P(z'|h^u)P(l|z')} \tag{2}$$

and a *Maximization Step*, where model parameters are updated (Equations 3-4) in order to maximize the likelihood (Equation 2)

$$P(z|h^u) \propto \sum_{l \in L} N(l, h^u)P(z|l, h^u) \tag{3}$$

$$P(l|z) \propto \sum_{u \in U} N(l, h^u)P(z|l, h^u) \tag{4}$$

where $N(l, h_u)$ designates the number of times POI l occurs in the history (record) of u. $P(z|h_u)$ and $P(l|z)$ are initialized to some random values. Finally, the steps described in Equations 2-4 are repeated until convergence (that is, when $P(z|h_u)$ and $P(l|z)$ reach an equilibrium).

3.2 Context Modelling

Context modelling is the inclusion of contextual information in the recommendation process. Context, in a RS framework, is predominantly linked to the *location* of the users (e.g. home, workplace, public place) and the *time* the recommendations are either requested or produced, ranging from hours (morning, evening), to days (workdays, weekends) and beyond (holiday periods) [3]. This representation of location and time implies that both quantities are described by a set of *characteristic attributes*; therefore let X be the set of spatial attributes and Y the set of temporal attributes, respectively.

The most straightforward way of fusing spatial attributes in RS is to model them as an extra multiplicative term in Equation 1, yielding Equation 5 below

$$P_s(l_t|h^u) = P(l_t|h^u)\frac{N(x|u)}{C(h^u)} \tag{5}$$

where $P(l_t|h_u)$ is the posterior probability of the contextless model, $P_{sp}(l_t|h_u)$ is the adjusted posterior probability of the context-aware model, $N(x|u)$ is the frequency of spatial attribute $x \in X$ associated with location l being visited by user u at time t and finally $C(h_u)$ is a normalization factor that ensures Equation 5 remains a probability distribution.

In a similar fashion, Equation 5 may be further extended by yet another multiplicative term that models temporal attributes, as in Equation 6 below

$$P_{s,t}(l_t|h^u) = P_s(l_t|h^u)\frac{P(f|l)}{C(l)} \tag{6}$$

where $P_s(l_t|h^u)^w$ is the adjusted posterior probability of context-aware model of Equation 5, $P_{s,t}(l_t|h^u)$ is the posterior probability of the new context-aware model combining spatial and temporal features, $P(f|l)$ is the probability of location l to be visited at time f by all users and $C(l)$ is a normalization factor with a similar role to $C(h^u)$.

4 Experiments & Results

4.1 Dataset

The aforementioned models, contextless and context-aware, have been evaluated on the `Flickr User-POI Visits Dataset` [25], which is comprised of user visits to various Points of Interest (POIs) in eight cities. Table 1 summarizes the dataset characteristics

Entries in the dataset correspond to geotagged photos uploaded by users on Flickr4, an image and video hosting service. Apart from the photo and user ids, each entry contains other useful metadata, such as the photo timestamp, the id and theme (category) of the photographed POI, the frequency this specific location has been photographed (visited) by other users in the dataset and

[4] https://flickr.com/

Table 1. General dataset characteristics

Characteristic	Value
Number of cities	8
Number of visits	153,208
Number of paths	21,186
Number of users	5,595
Average user-based path length	27.38
Average sequence-based path length	7.23
POI Categories (themes)	18

finally a sequence id. Sequence ids group photographs uploaded by the same user together, based on their timestamp; more specifically, photographs taken by a single user within a time frame of 8 hours are considered to belong to the same sequence. In addition to the metadata presented above, each POI is also characterized by its name, its coordinates (latitude and longitude) and a matrix containing the distances in-between POIs of the same city.

This specific dataset has been previously used in [18], where a tour recommendation framework is implemented and the *PersTour* algorithm is introduced (exhibiting better performance than the baseline algorithms of *Greedy Nearest* and *Greedy Most Popular*). The described system recommends a complete route to the visitor based on his/her previous sequences, while also exploiting the cost/benefit of each proposed route. Expanding on this idea, the authors in [19] group visitors of similar POIs, recommending tour itineraries for groups rather than isolated users. To the best of our knowledge, our approach is among the first to explore the contextual dimensions of this dataset.

4.2 Data Preprocessing

Path types and length According to the description of the dataset above, visitor paths may be determined in two ways; on a user basis or on a sequence basis. The latter option seems to be more natural, as sequences incorporate the temporal dimension of each visit, in the sense that geotagged photos of a certain user may span several days. However, sequence-based paths in the dataset are extremely short. This is evident both in Table 1, where the average sequence-based path length is just above 7 and in Figure 1, which depicts the number of paths of a given path-length; the overwhelming majority of sequence-based paths has a length between 1 and 3.

The same observation holds true for user-based path lengths as well (Figure 2). In this case, however, there exists a certain fraction of users whose path lengths are well above the short-path margins of the previous case. Therefore, in our analysis we considered user-based paths.

It could be argued that in considering user-based paths, the temporal context of the recommendations is overlooked. While there is a certain validity in this argument, it should also be stressed out that such an assumption does not hurt

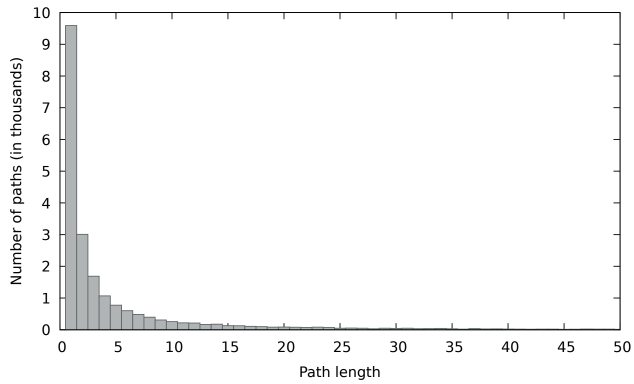

Fig. 1. Distribution of sequence-based paths in the dataset

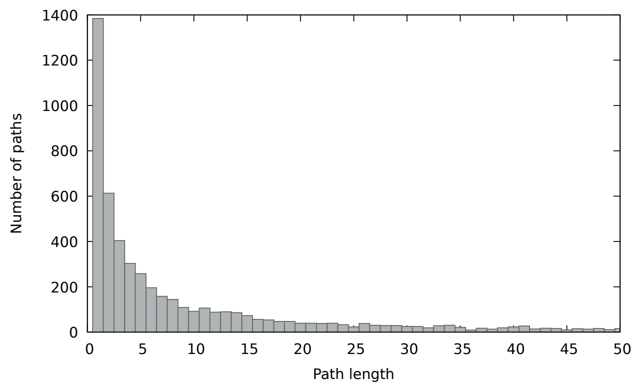

Fig. 2. Distribution of user-based paths in the dataset

the performance of the RS, as it is still able to predict new places that the users might have not visited in their previous visits.

Having fixed the way paths are determined (user-based instead of visitor-based paths), a decision should be made on the minimum path length. It is obvious that paths of short length are of no practical use to a RS and therefore need to be filtered out. This is achieved by defining a threshold on the minimum path length. Based on the reasoning above and after some experimentation, the optimal value of the minimum path length has been set to 5.

Number of latent features The number of latent features (parameter z) of the models described in Equations 1, 5-6 (Section 3) affects the performance of the respective RSs. Figure 3 depicts the performance of the topic model-based approach (Equation 1) on the visitors of the city of `Budapest`, Hungary.

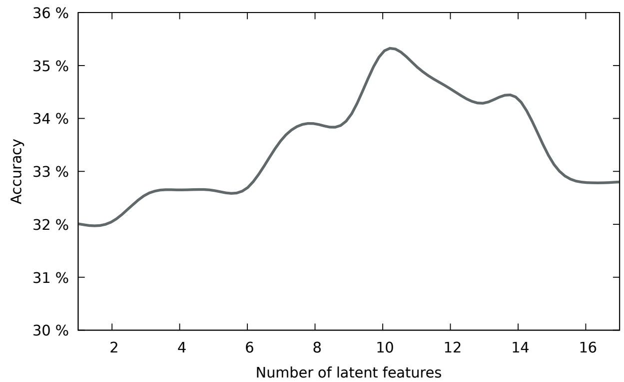

Fig. 3. The influence of the number of latent features on the performance of the topic modelling approach on the `Budapest`, Hungary subset of the dataset

It can be seen that the performance gradually rises, along with the dimensionality of the latent feature space, until it reaches a climax. Increasing z beyond that point is of no benefit to the model, as it cannot generalize. Therefore, determining the optimum value for z is a hyperparameter of the overall approach, dependent on the specific technique used and the data. By definition, the latent feature space is smaller than the quantities it models (users and number of POIs in this case). As the number of POIs per city ranges between 26 (in `Perth`, Australia) and 40 (in `Budapest`, Hungary), the optimum value for z has been sought in the range of $[2, 20]$ and it was found to be between 8 and 12 for all cities in the dataset.

4.3 Experiments

Having fixed the parameters and the hyperparameters of the proposed approach, the specifics of the experimental procedure need to be addressed. Since there is no connection (geographical or otherwise semantic) in-between the 8 cities of the dataset, 8 distinct sub-experiments were performed, pertaining to the data of each city. For each sub-experiment, four different approaches were considered; the contextless topic model of Equation 1, the inclusion of either the spatial or the temporal contextual information (Equation 5) and finally the incorporation of both contextual dimensions (Equation 6).

The experimentation protocol followed has been the *leave-one-out cross-validation*. At each iteration of the protocol one user is picked as the test user and the path s/he follows is split into two parts; the *training part* and the *test part*. This step is necessary for the models in order to approximate the distribution of the test user's interest in the latent topics (Equation 3). After experimenting with various train and test set sizes, we ended up using the first 25% of the path as the training set and the rest 75% as the test set. Consequently, when training was over, the RS proposed POIs to the test user and those recommendations were compared with the POIs in the test set in order to estimate system performance.

Context integration The dataset included two contextual parameters that could be exploited by the context-aware methodologies proposed in this work. The first one is a spatial parameter, the category or "theme" of the visited POI. It is a categorical variable that can take 18 distinct values, which are related to the cultural aspect of the visited POI (e.g. "Museum", "Historical", "Cultural"). Themes are indicative of the type of POIs each visitor prefers and therefore their modelling can be used to provide more personalized cultural recommendations. In this approach, and according to Equation 5, the frequency of the visited theme is initially calculated, adjusting the posterior probability $P_s(l_t|h_u)$ of u visiting landmark l_t next. After this adjustment, the model recommends the next POI to be visited, hopefully having gained some insight on the user's cultural preferences and achieving better performance.

The second parameter is a temporal one, namely the time the visit took place. The time is deduced from the timestamps of the photographs and it can be exploited to add contextual information about the periods within the day that each specific POI is accessed (for example, park visits usually occur during the day time). In the proposed approach, the time period corresponding to a day is split into time windows and then the frequency of the visits within each time window is measured for each POI. Then the posterior probability of u visiting landmark l_t next is adjusted according to the current time window and the aforementioned frequency. After experimenting with various time window sizes, the optimal value has been determined to be 8 hours. Finally, the two contextual parameters discussed above are integrated in the unified procedure of Equation 6.

Results Figure 4 summarizes the results of the experimental procedure for all 4 models and for the 8 cities. An initial observation is that the addition of the contextual information increases the accuracy of the recommendations in all cases by a margin of at least 5%. This is especially evident in the case of `Perth`, Australia, `Glasgow`, Scotland and `Osaka`, Japan and it is a clear indication that the incorporation of context is a process that enhances the overall quality of the recommendations. Additionally, a qualitative analysis on the cities where the RSs exhibited better results reveals that the visitors in these cases were associated with longer paths, which permitted the context-aware process to more thoroughly affect the functionality of the model.

5 Conclusions

In this work, a methodology of integrating contextual elements in cultural path recommendation has been outlined and experimentally evaluated on a dataset

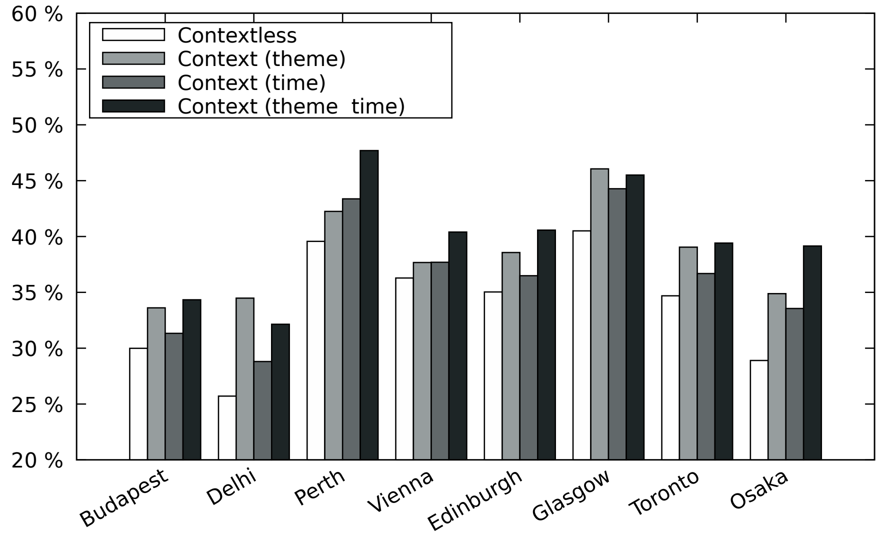

Fig. 4. Performance results of all models on all cities

of geotagged photos taken by visitors in eight cities. More specifically, the POIs visited by the users were modeled according to topic modeling, enabling the enhancement of the recommendation process with cultural location awareness. Furthermore, implicitly inferred context has been extracted from the dataset (POI theme and the time window within the day that each POI is visited). The measured performance demonstrated an increase of accuracy, when the contextual parameters were integrated.

The experimental results indicate that the inclusion of context-awareness into the recommendation process enhances the ability of the system to optimize its predictions. The achieved increase in accuracy is a promising result, pointing out that measuring and incorporating contextual data can greatly impact the overall system accuracy. Adding further and more explicit contextual data, such as user profile information (age, educational level, etc.) or environmental parameters (weather conditions, crowd density, etc.) is expected to result in even better recommendations.

References

1. Abbas, A., Zhang, L., Khan, S.U.: A survey on context-aware recommender systems based on computational intelligence techniques. Computing **97**(7), 667–690 (2015)
2. Abel, F., Gao, Q., Houben, G.J., Tao, K.: Analyzing user modeling on twitter for personalized news recommendations. In: International Conference on User Modeling, Adaptation, and Personalization. pp. 1–12. Springer (2011)
3. Adomavicius, G., Tuzhilin, A.: Context-aware recommender systems. In: Recommender systems handbook, pp. 217–253. Springer (2011)

4. Alexandridis, G., Siolas, G., Stafylopatis, A.: Accuracy Versus Novelty and Diversity in Recommender Systems: A Nonuniform Random Walk Approach, pp. 41–57. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-14379-8_3, https://doi.org/10.1007/978-3-319-14379-8_3
5. Ashbrook, D., Starner, T.: Using gps to learn significant locations and predict movement across multiple users. Personal and Ubiquitous computing **7**(5), 275–286 (2003)
6. Bartolini, I., Moscato, V., Pensa, R.G., Penta, A., Picariello, A., Sansone, C., Sapino, M.L.: Recommending multimedia visiting paths in cultural heritage applications. Multimedia Tools and Applications **75**(7), 3813–3842 (2016)
7. Benouaret, I., Lenne, D.: Personalizing the museum experience through context-aware recommendations. In: Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. pp. 743–748. IEEE (2015)
8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of machine Learning research **3**(Jan), 993–1022 (2003)
9. Dai, J., Yang, B., Guo, C., Ding, Z.: Personalized route recommendation using big trajectory data. In: Data Engineering (ICDE), 2015 IEEE 31st International Conference on. pp. 543–554. IEEE (2015)
10. Doytsher, Y., Galon, B., Kanza, Y.: Storing routes in socio-spatial networks and supporting social-based route recommendation. In: Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks. pp. 49–56. ACM (2011)
11. Giannotti, F., Nanni, M., Pinelli, F., Pedreschi, D.: Trajectory pattern mining. In: Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 330–339. ACM (2007)
12. Hofmann, T.: Probabilistic latent semantic indexing. In: Proceedings of the 22Nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 50–57. SIGIR '99, ACM, New York, NY, USA (1999). https://doi.org/10.1145/312624.312649, http://doi.acm.org/10.1145/312624.312649
13. Hofmann, T.: Collaborative filtering via gaussian probabilistic latent semantic analysis. In: Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval. pp. 259–266. ACM (2003)
14. Iwata, T., Watanabe, S., Yamada, T., Ueda, N.: Topic tracking model for analyzing consumer purchase behavior. In: IJCAI. vol. 9, pp. 1427–1432 (2009)
15. Konstantakis, M., Michalakis, K., Aliprantis, J., Kalatha, E., Caridakis, G.: Formalising and evaluating cultural user experience. In: Semantic and Social Media Adaptation and Personalization (SMAP), 2017 12th International Workshop on. pp. 90–94. IEEE (2017)
16. Krumm, J., Horvitz, E.: Predestination: Inferring destinations from partial trajectories. In: International Conference on Ubiquitous Computing. pp. 243–260. Springer (2006)
17. Kurashima, T., Iwata, T., Irie, G., Fujimura, K.: Travel route recommendation using geotagged photos. Knowledge and Information Systems **37**(1), 37–60 (Oct 2013). https://doi.org/10.1007/s10115-012-0580-z
18. Lim, K.H., Chan, J., Leckie, C., Karunasekera, S.: Personalized tour recommendation based on user interests and points of interest visit durations. In: IJCAI. vol. 15, pp. 1778–1784 (2015)
19. Lim, K.H., Chan, J., Leckie, C., Karunasekera, S.: Towards next generation touring: Personalized group tours. In: ICAPS. pp. 412–420 (2016)

20. Lu, E.H.C., Lee, W.C., Tseng, V.S.: Mining fastest path from trajectories with multiple destinations in road networks. Knowledge and information systems **29**(1), 25–53 (2011)
21. Panniello, U., Tuzhilin, A., Gorgoglione, M., Palmisano, C., Pedone, A.: Experimental comparison of pre-vs. post-filtering approaches in context-aware recommender systems. In: Proceedings of the third ACM conference on Recommender systems. pp. 265–268. ACM (2009)
22. Rodriguez, B., Molina, J., Perez, F., Caballero, R.: Interactive design of personalised tourism routes. Tourism Management **33**(4), 926 – 940 (2012). https://doi.org/https://doi.org/10.1016/j.tourman.2011.09.014
23. Ruotsalo, T., Haav, K., Stoyanov, A., Roche, S., Fani, E., Deliai, R., Mäkelä, E., Kauppinen, T., Hyvönen, E.: Smartmuseum: A mobile recommender system for the web of data. Web semantics: Science, services and agents on the world wide web **20**, 50–67 (2013)
24. Su, H., Zheng, K., Huang, J., Jeung, H., Chen, L., Zhou, X.: Crowdplanner: A crowd-based route recommendation system. In: Data engineering (ICDE), 2014 IEEE 30th International Conference on. pp. 1144–1155. IEEE (2014)
25. Thomee, B., Shamma, D.A., Friedland, G., Elizalde, B., Ni, K., Poland, D., Borth, D., Li, L.J.: The new data and new challenges in multimedia research. arXiv preprint arXiv:1503.01817 **1**(8) (2015)
26. Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., Duval, E.: Context-aware recommender systems for learning: a survey and future challenges. IEEE Transactions on Learning Technologies **5**(4), 318–335 (2012)
27. Zheng, Y., Zhang, L., Xie, X., Ma, W.Y.: Mining interesting locations and travel sequences from gps trajectories. In: Proceedings of the 18th international conference on World wide web. pp. 791–800. ACM (2009)