Towards sustainability optimization in touristic route recommendation

Pablo Sánchez*

Instituto de Investigación Tecnológica (IIT), Universidad Pontificia Comillas, C. Alberto Aguilera 25, E28015, Madrid, Spain.
**E-mail: psperez@icai.comillas.edu*

Alejandro Bellogín

Universidad Autónoma de Madrid, Madrid, Spain.
E-mail: alejandro.bellogin@uam.es

While route recommendation plays a vital role in digital tourism services, conventional methods tend to be inherently complex, as they need to consider several constraints at once: user preferences, geographical and scheduling information, etc. Besides, recent interests from society and institutions have motivated a greater emphasis on sustainability across multiple sectors, including tourism. This work proposes to use reranking techniques as a plain and efficient method to generate personalized routes for users, while considering sustainable goals in the procedure. We believe that these results could lead to new research directions regarding sustainable route recommendation tecniques.

Keywords: Route recommendation; reranking; green tourism

1. Introduction and motivation

For several years, tourism has witnessed an exponential growth, becoming a vital economic force for many countries [1]. As Location-Based Social Networks (LBSNs) such as Foursquare or Gowalla gain traction, users can easily share their trips and check-ins at the different Points-of-Interest (POIs) they have visited. This abundance of geotagged content on LBSNs has gained considerable attention among users who are interested in finding new venues to visit when they are in a specific city [2].

When recommending POIs to users, most current studies tend to focus on the accuracy of recommendations, by analyzing how many relevant venues we are recommending to the target user [3]. However, in the classical recommendation domain, many researchers have alerted about the necessity of evaluating recommendations on other aspects. Novelty and diversity are examples of beyond accuracy dimensions, and although they have been explored in the traditional recommendation problem, they could be especially relevant in the POI recommendation domain (and other connected problems). Exploring these additional dimensions could greatly benefit users, as it allows them to uncover less popular POIs, and also for the venue owners, as they might benefit financially from these recommendations. By exploring these beyond accuracy aspects, both route and POI recommenders could offer users a wider array of candidate venues to visit, favoring a more complete exploration of the city.

Besides, in the Information Retrieval, Recommender Systems, and Machine Learning communities, there is a growing emphasis on concepts such as novelty, diversity, and fairness. Nevertheless, this is normally limited to statistical biases that can be obtained from data, like popularity bias [4], or linked to measurements on sensitive attributes [5]. In any case, these trend matches some of the Sustainable Development Goals (SDGs)[a] stated by the United Nations in 2015, such as reaching gender equality (SDG 5), lessen inequalities (SDG 10), or responsible production and consumption (SDG 12). Thus, SDGs have attracted attention to particular societal problems that require collective action. In the context of tourism, besides those previously mentioned concerns, SDG 11 is especially relevant: "make cities and human settlements inclusive, safe, resilient and sustainable". Certain studies have already tackled these and related topics within the realm of the so-called *green tourism* [6–8] or *sustainable tourism* [9]. However, as far as we know, there is no link with the development of personalized routes.

In this context, this work considers the route recommendation problem from a sustainable viewpoint. More precisely, our approach involves the use of reranking techniques to generate route recommendations based on traditional POI recommendation methods. These strategies are capable of optimizing different criteria, including considerations such as the proximity of recommended venues and their categories, in order to encourage sustainable tourism by suggesting eco-friendly destinations. Our approach offers several benefits, such as straightforward implementation, effectiveness, and clarity, which could benefit both users and other stakeholders involved in the process. Moreover, the followed evaluation includes both classical and novelty metrics, providing a thorough comprehension of the recommendation effectiveness when addressing the varied needs of the users.

[a] `https://sdgs.un.org/goals`

2. Exploiting reranking methods for sustainable optimization

In this work, we propose to use item reranking techniques in order to create valuable sequences of venues. Our approach offers two main advantages with respect to the use of complex route generation algorithms. First, we can generate different rerankers depending on the dimension we aim to maximize (or minimize). Secondly, the rerankers enable us to derive routes even from recommendation models that may not inherently incorporate any sequential component. Now, Section 2.1 presents a reranking framework for route recommendation that will be used in our work. Then, Section 2.2 introduces our proposal for a sustainable reranker in this domain.

2.1. *Reranking framework*

We follow the same formulation as in our previous work [10], based on the original formulation from [11]. Beginning with an initial pool of candidate items for a particular user, generated by any recommendation model, our objective is to reorder them to identify a route that best satisfies the user's interests. As we select each candidate item and include it within the final recommendation list R, we simultaneously delete it from the candidate pool. The formulation is shown in Equation 1.

$$f_{obj}(u, i, R) = \lambda \cdot f_{rec}(u, i) + (1 - \lambda) \cdot f_{seq}(u, i, R) \tag{1}$$

Equation 1 defines the objective function $f_{obj}(u, i, R)$ considered in our greedy reranking process. Based on the score obtained in Equation 1, we select the venue i that obtains the highest value from the current set of candidate venues. These are the venues present in the original recommendation list that have not been chosen yet for the final itinerary recommended to the user.

In this case, as it is typical in the field [11], the objective function combines the score assigned to the user u and venue i pair by a recommendation algorithm, represented by the $f_{rec}(u, i)$ component, and the score provided by the sequence-aware reranker component, represented by $f_{seq}(u, i, R)$. Here, R represents the final recommendation list being generated. The sequence-aware reranker can take into account, if necessary, the previously selected candidate items that already belong to the final recommended route for user u.

Finally, a parameter λ is used to balance the trade-off between the original score of the recommender and the sequential reranker. A lower value of λ indicates a greater influence of the reranker, and therefore more possible changes to the original recommendation provided.

Our previous work[10] defined seven different formulations for the sequence-aware reranker component f_{seq}, classified in three major families: independent reranker (the score is solely determined by the user-item pair under consideration), for example, based on scores produced by Random or Recommender-based strategies (`rec`); based on the preceding item (the score assigned to item i is only influenced by the last item included in list R, denoted as i_{n-1}), such as the rerankers based on Distance (`dist`), Feature-based Markov Chain (`feat`), or Item-based Markov Chain (`item`); and dependent on the entire sequence (the score assigned to an item is determined by considering the entire sequence that has been generated so far, that is, the current list R and the potential next item i), using algorithms based on Longest Common Subsequence (`lcs`) or Suffix Tree (`stree`).

2.2. *A reranker for sustainable route recommendation*

We propose a new reranker that corresponds to the last family described before (dependent on the whole sequence), named **Outdoor Reranker**. For this reranker, we identify a target distribution T_{Dist} of the categories of the venues to be recommended to the user, so the next candidate to be included in the route should be the one minimizing the difference between the target category distribution T_{Dist} and the categories of the route being generated so far, that is, the categories of R.

To promote more sustainable recommendations, the target distribution T_{Dist} is defined based on categories tailored for sustainability, such as parks and botanical gardens or, in general, outdoor POIs[8]. By doing so, we aim to provide a more balanced and nuanced set of recommendations that align with the user's preferences (by considering the recommendations coming from the original algorithms) while also promoting the exploration of more sustainable categories (through the use of the Outdoor Reranker, `outd`).

3. Experiments

3.1. *Settings*

In this section, we will describe the evaluation methodology followed in our experiments. We use the Trip builder dataset from[12], that combines two data sources, Flickr and Wikipedia, by building trajectories using the geotagged photos taken by the users in three Italian cities, although we focus on the city with a higher number of check-ins, Rome. This dataset, once processed, provides 7,954 users who created 27,252 routes (trajectories) on 394 venues, producing 61,330 interactions between the users and the venues.

We chose the last itinerary of users with a minimum of three distinct routes as long as they have at least 4 POIs in the last route to create the test set. The remaining data will be retained for training the recommenders. The evaluation methodology, in line with recommendations from the research community [13], selects the candidate items from the training set, to ensure a more realistic setting. This means that for each user in the test set, we will consider as candidate POIs all venues that appear but not been visited by the user in the training set.

3.2. *Recommendation approaches*

To cover classical algorithms and methods more tailored for the tourism domain, we have included in our experiments a varied combination of methods, which shall be classified as classic and geographical approaches. In the first family, we include a popularity recommender (Pop), two nearest-neighbor methods [14] (one based on users: UB, and another on items: IB), and two matrix factorization models (one based on Alternate Least Squares: HKV [15], and another on Bayesian Personalized Ranking from [16]: BPRMF). In the second family, the algorithms exploit the geographical coordinates of the items, such as AvgDis (that suggests the closest POIs to the user's average location), IRenMF (a weighted Matrix Factorization technique proposed in [17] that considers the geographical influence between the target POI with respect to its neighboring venues and the region it belongs to), and RankGeoFM (a matrix factorization model proposed in [18] that incorporates the geographical influence between neighboring POIs while also capturing user geographical preferences by integrating an extra latent matrix).

3.3. *Results*

We now present the performance results obtained when applying the reranking framework described before to the recommendation algorithms listed in the previous section. Specifically, for accuracy, we report $\mathrm{nDCG_s}$, which stands for nDCG incorporating the sequential component with LCS penalization, as defined in [10]. For novelty, we use the Expected Popularity Complement (EPC), which assumes that a POIs is novel if its popularity is low [11]. Higher values in all metrics indicate a better performance of the models, indicating a higher accuracy or novelty. Moreover, unless specified

Table 1. Results of standard recommenders. **Bold** denotes the best results for each metric, in *italic* the (next) best result in each family (classical/geographical) is emphasized.

	Pop	UB	IB	HKV	BPRMF	AvgDis	IRenMF	RankGeoFM
nDCG	**0.508**	0.469	0.473	0.425	*0.508*	0.362	0.401	*0.486*
nDCG_s	**0.447**	0.420	0.419	0.386	*0.447*	0.334	0.367	*0.427*
EPC	0.848	0.865	0.864	*0.894*	0.848	**0.965**	*0.937*	0.865

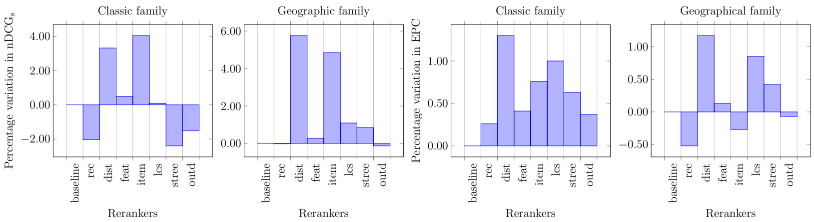

Fig. 1. Variation in nDCG_s (left) and EPC (right) with respect to no reranker (see reranker abbreviations in Sections 2.1 and 2.2).

otherwise, these metrics will be reported on rankings of length 10.

Table 1 shows a strong popularity bias, since the popularity recommender outperforms other models in terms of accuracy (nDCG). This is a well-known fact in the community for other domains[4], but in this case, factors such as high sparsity and a concentration of check-ins in a short number of venues, reinforce this behavior. The trend in performance stays the same when incorporating sequence in accuracy measurement (nDCG_s), where Pop and BPRMF obtain the best results among the classical approaches, whereas RankGeoFM is the best geographical algorithm in terms of accuracy. EPC, as a measure of novelty, prompts a balance with the accuracy metrics, as the best methods now (AvgDis and HKV) correspond to the worst methods according to nDCG_s or nDCG.

Let us now analyze the impact of the reranking framework. Figure 1 shows the performance variation when using such framework on the best method of each family. These figures take as baseline that method, which is then used by the reranker to reorder its first 20 recommended items, and apply Equation 1 with $\lambda = 0.5$ exploiting each reranker capabilities. We observe that distance-based and Item-based Markov Chain rerankers are the only ones that consistently improve accuracy and novelty in both recommendation families. Regarding the outdoor reranker, it is remarkable that, by promoting more sustainable venues, it is able to gain novelty at the expense of accuracy in the classic family, whereas for the geographic family the observed changes are negligible.

Hence, we conclude that, by optimizing for sustainability through reranking, recommendation performance either stays the same, or, at least, one of the analyzed evaluation dimensions (accuracy or novelty) improves.

4. Conclusions and future work

In this paper, we have extended a reranking framework proposed previously and analyzed how sensitive the recommendation methods were when promoting outdoor activities. By doing so, we have observed which algorithms and rerankers are more effective in promoting more sustainable suggestions and healthier recommendations for users. In particular, our proposed reranker has shown promising results in these directions. However, promoting only outdoor POIs, suggests the necessity of combining our approach with other models to maximize its usefulness.

A promising line of future work would involve extending the reranking framework for multi-objective route recommendation. This would allow us to combine and optimize several rerankers at once, each of them tailored to a different goal. Such an approach holds particular interest in tour planning, where it could be seamlessly integrated [19]. Related with the previous idea, we believe the proposed reranker could be used to further sustainability objectives – such as reducing CO2 emissions [8] – and, in general, other principles of fair recommendation with an emphasis on sustainability [9].

Acknowledgments

Work supported by Grant PID2022-139131NB-I00 funded by MCIN/AEI/ 10.13039/501100011033 and by "ERDF A way of making Europe."

References

1. H. Rasool, S. Maqbool and M. Tarique, The relationship between tourism and economic growth among brics countries: a panel cointegration analysis, *Future Business Journal* **7**, p. 1 (Jan 2021).
2. L. W. Dietz, A. Sen, R. Roy and W. Wörndl, Mining trips from location-based social networks for clustering travelers and destinations, *J. Inf. Technol. Tour.* **22**, 131 (2020).
3. P. Sánchez and A. Bellogín, Point-of-interest recommender systems based on location-based social networks: A survey from an experimental perspective, *ACM Comput. Surv.* **54**, 223:1 (2022).
4. A. Bellogín, P. Castells and I. Cantador, Statistical biases in information retrieval metrics for recommender systems, *Inf. Retr. Journal* **20**, 606 (2017).
5. E. Amigó, Y. Deldjoo, S. Mizzaro and A. Bellogín, A unifying and general account of fairness measurement in recommender systems, *Inf. Process. Manag.* **60**, p. 103115 (2023).

6. S. kwon Hong, S. il Kim and J. hyun Kim, Implications of potential green tourism development, *Annals of Tourism Research* **30**, 323 (2003).
7. G. Divsalar, A. Divsalar, A. Jabbarzadeh and H. Sahebi, An optimization approach for green tourist trip design, *Soft Comput.* **26**, 4303 (2022).
8. A. Susanty, N. B. Puspitasari, S. Saptadi and S. Prasetyo, Implementation of green tourism concept through a dynamic programming algorithm to select the best route of tourist travel, *IOP Conference Series: Earth and Environmental Science* **195**, p. 012035 (dec 2018).
9. D. Streimikiene, B. Svagzdiene, E. Jasinskas and A. Simanavicius, Sustainable tourism development and competitiveness: The systematic literature review, *Sustainable Development* **29**, 259 (2021).
10. P. Sánchez and A. Bellogín, Applying reranking strategies to route recommendation using sequence-aware evaluation, *User Model. User Adapt. Interact.* **30**, 659 (2020).
11. M. Kaminskas and D. Bridge, Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-accuracy objectives in recommender systems, *TiiS* **7**, 2:1 (2017).
12. D. Monti, E. Palumbo, G. Rizzo and M. Morisio, Sequeval: A framework to assess and benchmark sequence-based recommender systems, *CoRR* **abs/1810.04956** (2018).
13. A. Said and A. Bellogín, Comparative recommender system evaluation: benchmarking recommendation frameworks, in *Eighth ACM Conference on Recommender Systems, RecSys '14, CA, USA - October, 2014*, (ACM, 2014).
14. X. Ning, C. Desrosiers and G. Karypis, A comprehensive survey of neighborhood-based recommendation methods, in *Recommender Systems Handbook*, (Springer, 2015) pp. 37–76.
15. Y. Hu, Y. Koren and C. Volinsky, Collaborative filtering for implicit feedback datasets, in *Proceedings of the 8th IEEE International Conference on Data Mining, December 15-19, Pisa, Italy*, (IEEE Computer Society, 2008).
16. S. Rendle, C. Freudenthaler, Z. Gantner and L. Schmidt-Thieme, BPR: bayesian personalized ranking from implicit feedback, in *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, Montreal, QC, Canada, June 18-21, 2009*, (AUAI Press, 2009).
17. Y. Liu, W. Wei, A. Sun and C. Miao, Exploiting geographical neighborhood characteristics for location recommendation, in *Proceedings of the 23rd International ACM CIKM Conference, Shanghai, China, November 3-7, 2014*, (ACM, 2014).
18. X. Li, G. Cong, X. Li, T. N. Pham and S. Krishnaswamy, Rank-geofm: A ranking based geographical factorization method for point of interest recommendation, in *Proceedings of the 38th International ACM SIGIR Conference, Santiago, Chile, August 9-13, 2015*, (ACM, 2015).
19. S. Choachaicharoenkul, D. W. Coit and N. Wattanapongsakorn, Multi-objective trip planning with solution ranking based on user preference and restaurant selection, *IEEE Access* **10**, 10688 (2022).