Social media profiling for personalized touristic recommendations

Jonathan A. Orama[#*1], Antonio Moreno[#2]

[#]*Intelligent Technologies for Advanced Knowledge Acquisition (ITAKA) Research Group, Escola Tècnica Superior d'Enginyeria, Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Av. Països Catalans, 26, 43007 Tarragona, Spain*

[1]orama.jonathan@gmail.com, [2]antonio.moreno@urv.cat

[*]*Barcelona Supercomputing Center, Plaza Eusebi Güell, 1-3, 08034 Barcelona, Spain*

[1]jonathan.ayebakuro@bsc.es

***Keywords*— recommender systems, user profiling, cluster analysis, association rule mining, tourist flows, mobility patterns**

I. EXTENDED ABSTRACT

A. Purpose

Tourists' interests and tastes guide their choice of activities at a destination. They seek out attractions and traverse routes that align with their preferences and travel habits. With an ever growing tourism demand, it has become difficult for tourists to plan trips due to an increase in potential things to do at any destination.

Tourism recommender systems offer a solution to this problem by making personalized suggestions considering user´s preferences and their contexts [1]. The development of accurate tourism recommender systems is an open topic in the tourism field. Tourism recommender systems have considered the context of the trip [2], the groups tourist are part of [3], the order in which they should visit attractions [4], and many other important aspects that may affect the accuracy of recommendations.

Social media platforms have emerged as a crucial data source for tourism analysis due to their widespread usage and accessibility by tourists seeking to share experiences online [5]. Recommender systems can also profit from the data provided by tourists on these social media platforms to learn their preferences, and gain insights into their travel habits [6]. These goals were achieved in our work by combining clustering-based social media profiling with association rule mining to suggest a set of attractions that align with a tourist's activity, travel, popularity, and temporal preferences. These suggested attractions were then ordered to provide a sequence of activities that maximizes the enjoyment of the tourist's experience [7].

Furthermore, the analysis of social media data can provide useful insights into the *mobility* of tourists within a destination, which can help destination management organizations improve the tourism experience, and curb negative effects like overtourism. We have shown that clustering tourists by their travel mobility is an effective way of providing knowledge to support decision making in destination management organizations [8].

B. Methodology

Geo-located data collected from Twitter were cleaned by removing users with a negligible amount of posts, and posts from local citizens.. A tweet was considered to be from a tourist if Barcelona was not among his/her home locations. These home locations were specified explicitly in the individual's Twitter profile, or they were inferred from his/her tweets. Concretely, a place was considered to be a user's home location if he/she had posted daily tweets in a 20-day period. Open Street Map was then used to categorize posts under 9 main categories of activities. 6066 tourists were represented with a numeric vector modelling their interests in the activity categories, their trip length, their mobility around the destination, their interests in popular or unpopular attractions, and their preferred touring times. Finally, tourists were profiled by clustering them using their vector representation, obtaining 25 clusters.

Next, association rule mining was used to select attractions for recommendation, ensuring that tourists were suggested attractions appropriate for the members of their cluster. This was done by mining frequent item sets from the sets of attractions visited by tourists within a day. Afterwards, association rules were identified. These rules were then ranked by their support, confidence, lift, and their alignment to the tourist's popularity and activity preferences. A final selection of a set of 10 attractions was carried out, preserving the causality in the rules they were selected from. An evolutionary algorithm was finally used to optimally order suggestions [7]. To evaluate the recommendations, metrics like precision, recall, coverage, similarity, personalization, popularity, and diversity were analyzed, to capture the main aspects of the recommender system.

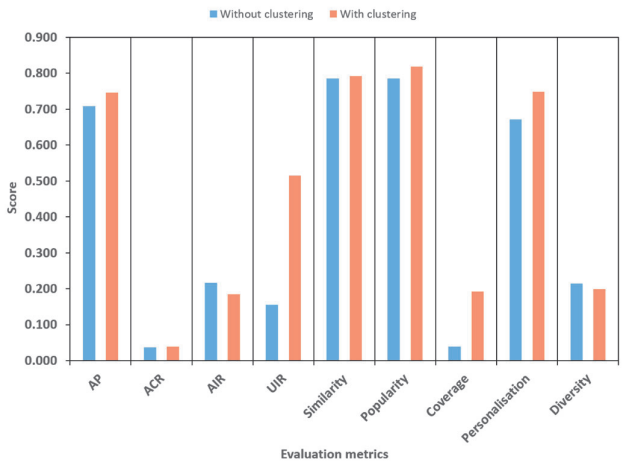

Fig. 1 A bar plot of the average performance scores of the recommender system across the three experiments. Cases with and without clustering are shown in orange and blue, respectively.

Partially supported by Spanish Thematic Network on Recommender Systems (Action RED2022-134302-T funded by MCIN/ AEI/10.13039/501100011033)

C. Findings

Three experiments were conducted to compare the effect of the clustering-based social media profiling on the association rule mining selection of attractions for recommendation. 80% of 6066 tourists were used for profiling and 20% were used to test the recommendations, computing the evaluation metrics. In almost all metrics, clustering provided better results, especially in precision, similarity, popularity, and personalization. In figure 1, the evaluation metrics are displayed. Recall metrics showed the worst performances, due to the constraints put on the selection process. These constraints also affected diversity, because the clustering process reduced the pool of suggestable attractions.

Furthermore, using the clusters to study the mobility of tourists at a destination showed that it is possible to understand the main themes of a destination. Destinations with more focus on specific attraction types will see more tourist flows in those kinds of attractions. It was also shown that it is possible to understand tourists' mobility based on their country of origin and differentiate tourist classes based on their visits. A visual study of the mobility within clusters showed types of tourists willing to travel longer distances to experience all the destination has to offer, as seen in figure 2.

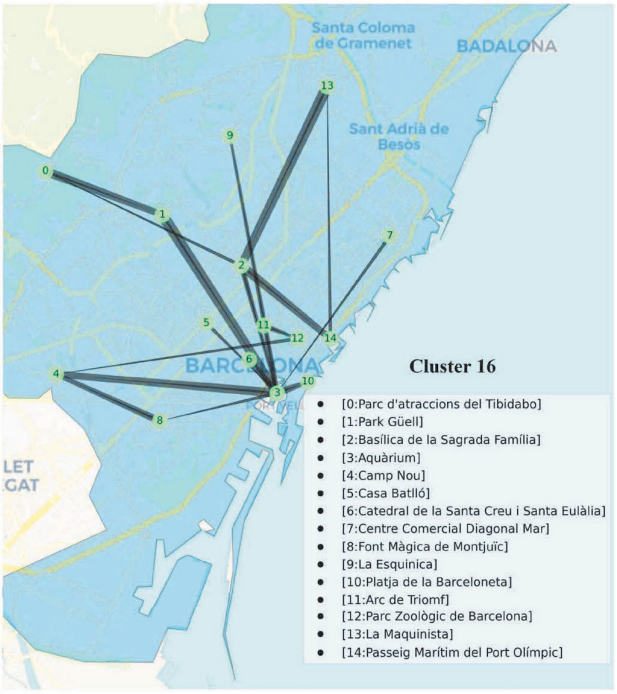

Fig. 2 Tourist mobility pattern between attractions in Barcelona for cluster 16. It shows a class of tourists willing to travel all around the city, mainly comprising of tourists that travelled long distances to visit Barcelona.

D. Research Implications and Limitations

Although recommender systems are very useful in the tourism industry, because of the diversity of tourism destinations and the vast number of factors that could affect the accuracy of recommendations, the findings in this research may not hold in some other destinations. This study only focused on attractions in Barcelona and used social media posts published in the year 2019. These facts notwithstanding, using the methods in this research, it is possible to make useful recommendations if the methodology is properly adapted to a new destination. Also, destination marketing organizations and city planners can benefit from using the methods in this research, to study the flow of tourists at a destination. Refined bus routes, designated tourist areas, and the reduction of the effect of overtourism are some of the results that could be achieved.

E. Research Originality and Value

This research implemented a novel way of combining cluster-based profiling with association rule mining for recommending sequences of touristic activities. These ideas were described with more detail in the previous publications [6], [7] and [8].

ACKNOWLEDGMENTS

This work was funded by URV projects 2022PFR-URV-41 and 2023PFR-URV-00114 and the Departament de Recerca i Universitats of Generalitat de Catalunya (Consolidated research group 2021 SGR 00114). J.Orama was a fellow of Eurecat's "Vicente López" PhD grant program. The support of the Spanish Research Network on Recommender Systems (ELIGE-IA, RED2022-134302-T) is also acknowledged.

REFERENCES

[1] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," Expert Syst. Appl., Vol. 41, No. 16, Pp. 7370–7389, Nov. 2014.

[2] K. Haruna, M. A. Ismail, S. Suhendroyono, D. Damiasih, A. C. Pierewan, H. Chiroma, T. Herawan, "Context-Aware Recommender System: A Review of Recent Developmental Process and Future Research Direction," Appl. Sci., Vol. 11, No. 14, Pp. 1211–1211, Dec. 2017.

[3] S. Dara, R. C. Chowdary, C. Kumar, "A survey on group recommender systems," J. Intell. Inf. Syst., Vol. 54, Pp. 271–295, Jan. 2020

[4] M. Quadrana, P. Cremonesi, D. Jannach, "Sequence-Aware Recommender Systems," ACM Comput. Surv., Vol. 51, No. 4, Pp. 1–36, Jul. 2018.

[5] P.Sánchez, A.Bellogín. "Point-of-Interest Recommender Systems based on Location-Based Social Networks: a survey from an experimental perspective". ACM Computing Surveys 54 (11), pp. 1-37, 2022.

[6] J. A. Orama, J. Borràs, A. Moreno, "Combining Cluster-Based Profiling Based on Social Media Features and Association Rule Mining for Personalised Recommendations of Touristic," Appl. Sci., Vol. 11, Nov. 14, Pp. 6512–6512, Jul. 2021.

[7] J. A. Orama, A. Moreno, J. Borràs, "Multi Objective Genetic Algorithm for Optimal Route Selection from a Set of Recommended Touristic Activities," Front. Artif. Intell. Appl., Vol. 356, Pp. 9–12, Oct. 2022.

[8] J. A. Orama, A. Huertas, J. Borràs, A. Moreno, and S. A. Clavé, "Identification of Mobility Patterns of Clusters of City Visitors: An Application of Artificial Intelligence Techniques to Social Media Data," Appl. Sci., Vol. 12, No. 12, Pp. 5834–5834, Jun. 2022.