Holistic Summarization of Recommender Systems Results

Noemi Mauro[1], Zhongli Filippo Hu[1], Liliana Ardissono[1], Sara Capecchi[1], Gianmarco Izzi[1] and Claudio Mattutino[1]

[1]Computer Science Department, University of Torino, Corso Svizzera 185, 10149, Torino

Abstract

Item-centric summaries of recommender systems results fail to provide an overall perspective on consumer satisfaction because they focus on features and aspects of suggestions. However, the appreciation of items is not only a matter of liking their properties. The whole process of their fruition, which might involve interacting with services and other actors, should be taken into account to enhance user awareness and decision-making. To address this issue, we present a visual summarization model that supports a holistic overview of search results by exploiting an explicit representation of the service underlying item fruition as a basis to measure multiple user experience evaluation dimensions. We instantiated our model on the home-booking domain. A preliminary user study has shown its usefulness as an information filtering tool to screen recommendation lists down to a small set of promising options.

Keywords

Interactive information exploration, Information visualization, Service Journey Maps

1. Introduction

In the recommender systems research [1], it is well known that the generation of focused suggestion lists causes a bubble effect that limits user awareness [2]. However, extending the diversity and number of suggestions might overload the user with too much information. Some visual presentation models are proposed to summarize consumer feedback emerging from the reviews collected by online retailer platforms [3, 4], or to explain recommendation results [5, 6]. However, these approaches describe specific aspects of items, at the risk of providing partial views of the available options, or overloading users in the attempt to be more specific. Moreover, they generate item-centric summaries that partially represent consumers' experience with items. For example, in services such as hotel booking, the room details, and the interaction with clerks, might jointly impact customer experience. To enhance decision-making, a holistic presentation of information should be provided that takes these aspects into account.

Chen et al. reported that product comparison is a crucial decision stage that buyers usually perform before they make a choice [4]. However, while detailed item features, and aspects, are relevant during the analysis of small sets of options, we point out that compact, high-level information should be provided to help users screen large item sets down to a pool of reasonably

IIR 2021 – 11th Italian Information Retrieval Workshop, September 13–15, 2021, Bari, Italy

noemi.mauro@unito.it (N. Mauro); zhonglifilippo.hu@unito.it (Z. F. Hu); liliana.ardissono@unito.it (L. Ardissono); sara.capecchi@unito.it (S. Capecchi); gianmarco.izzi@unito.it (G. Izzi); claudio.mattutino@unito.it (C. Mattutino)

relevant options, which can be analyzed with limited effort. Thus, we propose to put the user in control of the search process by characterizing items at two levels: a high-level one, which summarizes consumer experience, and a fine-grained one that deals with detailed item features and aspects. In this paper, we propose a visual model to manage the higher level. Our model supports the inspection of possibly long recommendation lists by providing a holistic summary of consumer feedback about items. For each item, it shows an interactive bar chart that provides quantitative information about previous consumers' opinions, during the whole process of item fruition, and taking multiple evaluation dimensions of experience into account. For the extraction and organization of information about items, we rely on a domain representation based on the "Service Journey Maps" [7, 8], which support the definition of high-level aspects of a service associated with different stages of fruition by its users. We tested our model within the Apartment Monitoring application [9], which personalizes the suggestion of homes for rent based on data extracted from Airbnb (https://www.airbnb.com/). A preliminary user study has shown that our model is perceived as a useful information filtering tool to overview search results and quickly identify a small set of promising options out of the available ones.

2. Background and Related Work

The Service Journey Maps (SJMs) [7, 10] are a pillar of our work. They support the design and development of physical and online products and services by focusing on the customer's viewpoint. A SJM is a visual description of the user experience with a service that models the various stages a person encounters during its fruition, in a temporal line from the start point to the end one. We use SJMs to analyze and organize consumer feedback in home-booking. Starting from the identification of the stages of this service, we derive the high-level evaluation dimensions of items that can be shown in the recommendation list to summarize user experience.

We now position our work in the literature about explanation and justification of recommender systems results. Some aspect-based recommender systems present suggestions by highlighting the features of items that match, or mismatch, the target user's preferences [11]. Other ones support feature-based item comparison [4, 12, 13], or information exploration based on the visualization of the relevance of items with respect to the keywords of search queries [14, 15]. Finally, some systems fuse the generation and the presentation of suggestions to enhance their own transparency [16, 17].

Different from these works, we aim at enabling the user to make a first-hand opinion about the proposed solutions by overviewing the recommendation list, and by efficiently inspecting consumer feedback. Previous works present items in detail, but this challenges the visualization of long lists of options. In comparison, we pursue the generation of high-level, quantitative overviews providing a holistic perspective about items, used both to rank options and to present them to the user.

In the exploratory search support research [18], faceted search interfaces empower the user to control the information filtering process by guiding the selection of item features [19, 14, 20]. However, they return items having the exact features specified by the user; e.g., the restaurants that offer outdoor seating. Therefore, they poorly address evaluation dimensions that depend on the aggregation of properties, such as product quality, or the experience in item fruition.

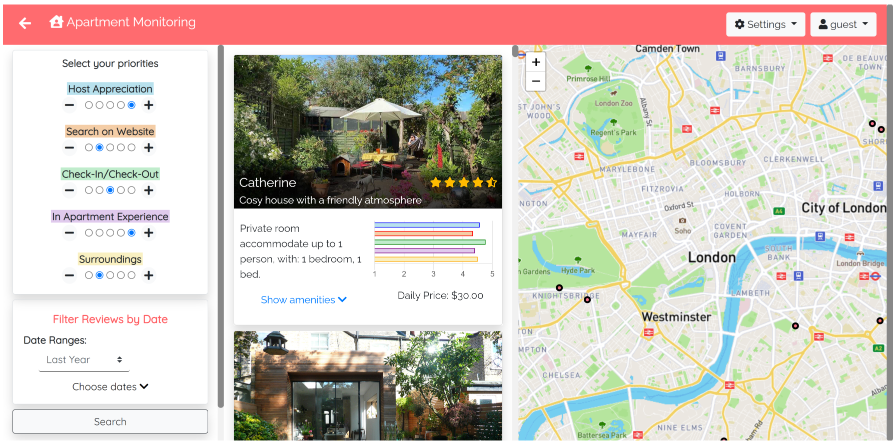

Figure 1: Current user interface of Apartment Monitoring: visualization of a recommended home with bar graphs summarizing previous guest's experience about the home. Photo and data about the home are retrieved from http://insideairbnb.com/get-the-data.html.

3. Visual Summarization of Recommendation Lists

Our model is integrated in the Apartment Monitoring application, which helps the user explore homes for rent by enriching the basic data (details, and reviews) provided by Airbnb[1] with a holistic overview of previous guests' experience with homes. This overview is based on a domain model that defines a set of high-level evaluation dimensions of the expected experience with the home, considering the overall home-booking process. In [9], Mauro et al. specified the domain model of Apartment Monitoring by defining a Service Journey Map (SJM) that describes the stages of a typical home-renting experience from the customer's viewpoint. By analyzing the literature about home and hotel-booking [21, 22, 23], these authors defined a small set of high-level evaluation dimensions associated with the stages of the SJM, meant to measure previous guests' renting experience: D = *Host appreciation* (i.e., perceptions about the host, and the interaction with her/him at any time of service fruition), *Search on website* (perceived effort in retrieving data about the home in the Airbnb website), *Check-in/Check-out* (experience at check-in, and check-out times), *In apartment experience* (experience within the home), and *Surroundings* (perceptions about the area around the home).

3.1. Overview of Previous Guests' Experience with Homes

Figure 1 shows the current user interface of Apartment Monitoring, which extends the work described in [9] with (i) the integration of a recommender system for the personalized suggestion of homes to the individual user, and (ii) the summarization of previous guests' experience with homes to help the user quickly find the relevant options within the recommendation list.

The left sidebar of the user interface displays the form through which users can specify their priorities towards the set D of experience evaluation dimensions. The right portion of the user

[1]Data can be downloaded from http://insideairbnb.com/get-the-data.html, under Creative Commons CC0 1.0 Universal (CC0 1.0) "Public Domain Dedication" license.

interface shows a scrollable list of suggested homes, ranked by estimated user rating, and their locations in a geographical map. Homes are represented as red circles on the map. The rating of a home h is computed as the weighted mean of the evaluation dimensions $d \in D$ that emerge from h's reviews. The user's priorities are employed as weights so that the most important dimensions influence rating estimation in the strongest way.

For each home h, the application shows the name of its host, the denomination of h, a picture with the overall evaluation of experience, expressed as a star list, a short description, the daily price, and the list of offered amenities. Moreover, it shows a bar graph that summarizes previous guests' renting experience with respect to the dimensions of D. By clicking on a home, the user can view detailed information about it, including its reviews.

For each home h, the value of each evaluation dimension $d \in D$ is the average opinion about d extracted from the reviews received by h. These dimensions take values in the [1, 5] interval. Their values are computed by applying standard NLP and sentiment analysis techniques to the text of the reviews. For the association of the terms occurring in the reviews to the dimensions of D, we use a set of thesauri we defined, one for each dimension. See [9] and [24] for details.

4. Preliminary User Study

We investigate the information filtering support provided by our visual model to understand if users can efficiently identify the most promising items of the recommendation list by looking at the experience summaries provided by the bar graphs of the homes.

4.1. Methodology

We recruited 11 participants from the University staff, students, and our social connections, having in mind a target of people who might be interested in searching for homes online. People joined the user study on a voluntary basis, without any compensation, and they gave their informed consent to participate in the study; see Section 4.3. The study took place live, in video calls with shared screen due to the COVID-19 pandemic.

PARTICIPANTS DATA. Gender: 54.54% women; 45.46% men. Age: between 19 and 57, mean=36.36. Education level: 27.28% attended high school, 36.36% university, 18.18% have a Ph.D, and 18.18% attended middle school. All participants regularly use the Internet.

TASK. We asked participants to rate 5 homes presented in a minimalistic format that only included the home number and its bar graph. The reason for this decision was that we wanted to minimize the presence of data that might influence the user in the evaluation task [6]. People could specify that they did not know which rating to give (opting-out).

4.2. Results

Some people did not evaluate the homes (12.73%). Moreover, a participant said that the visualized information is very useful to select the candidate homes for detailed inspection, but that (s)he would not feel confident in booking a home on the sole basis of this data. These findings are not surprising, as bar graphs do not describe detailed characteristics of homes, which are important to make rating decisions. However, two people declared that bar graphs are useful to filter out homes that do not deserve to be further analyzed due to low performance in some evaluation dimensions they care about. Moreover, three people stated that, if a home has a low *Host appreciation*, they would not consider the other dimensions to rate it. These findings suggest that the experience summaries provided by the bar graphs are useful to filter out irrelevant homes because they help users efficiently identify their pitfalls at a coarse-grained level. On a different perspective, three people declared that they did not care much about the values of the bars in the graph. Indeed, they compared the bars to each other to see on which dimensions a home was rated better, or worse, by previous guests.

4.3. Ethical Issues

In planning our user study, we followed literature guidelines on controlled experiments[2] [25]. Participants were informed about their rights: (i) the right to stop participating in the experiment, possibly without giving a reason; (ii) the right to obtain further information about the purpose, and the outcomes of the experiment; (iii) the right to have their data anonymized.

Before starting the experiment, participants were asked to: (i) read a consent form, stating the nature of the experiment and their rights, and (ii) sign it to indicate that they read and understood their rights. Moreover, we reassured them that that the objective of the experiment was to identify possible faults in the proposed service, and not to test their own ability, or intelligence. Every participant was given the same instructions before the experimental tasks.

We did not store participants' names. During the user study, and the analysis of its results, we worked with anonymous codes.

5. Conclusions

We described a visual model aimed at summarizing consumer experience with items. Our model, instantiated in the home-booking domain, employs the Service Journey Maps to present holistic information that takes the service underlying item fruition into account. Usually, in e-commerce feedback systems, there is a lack of connection between reputation statements and their context. For instance, the judgment of the product, delivery, price, and interaction with the retailer, are jointly represented by means of a 5-star claim type, plus a detailed written feedback. Thus, the resulting value is an overall evaluation that does not reflect the related contexts in which it is obtained [26, 27, 28]. Differently, the model we propose offers a detailed evaluation of each stage of item fruition. A preliminary user study has shown that participants could benefit from our model to screen down the set of homes to be inspected for a booking decision. As future work, we plan to extend the bar graphs describing consumer experience with on-demand, qualitative data supporting the evaluations they express [29]. In this way, users will be able to inspect detailed information about the items of interest. Moreover, we plan to validate our model through a larger user study.

This work has been funded by the University of Torino under grant ARDL_RILO_2019.

[2] https://www.tech.cam.ac.uk/research-ethics/school-technology-research-ethics-guidance/controlled-experiments

References

[1] F. Ricci, L. Rokach, B. Shapira, Introduction to Recommender Systems Handbook, Springer US, Boston, MA, 2011, pp. 1–35. URL: https://doi.org/10.1007/978-0-387-85820-3_1. doi:10.1007/978-0-387-85820-3_1.

[2] T. Nguyen, P.-M. Hui, F. M. Harper, L. Terveen, J. Konstan, Exploring the filter bubble: The effect of using recommender systems on content diversity, in: Proceedings of the 23rd International Conference on World Wide Web, WWW '14, Association for Computing Machinery, New York, NY, USA, 2014, p. 677–686. URL: https://doi.org/10.1145/2566486.2568012. doi:10.1145/2566486.2568012.

[3] C. Musto, M. de Gemmis, P. Lops, G. Semeraro, Generating post hoc review-based natural language justifications for recommender systems, User-Modeling and User-Adapted Interaction 27 (2020). URL: https://doi.org/10.1007/s11257-020-09270-8. doi:10.1007/s11257-020-09270-8.

[4] L. Chen, F. Wang, L. Qi, F. Liang, Experiment on sentiment embedded comparison interface, Knowledge-Based Systems 64 (2014) 44–58. URL: https://www.sciencedirect.com/science/article/pii/S0950705114001099. doi:https://doi.org/10.1016/j.knosys.2014.03.020.

[5] I. Nunes, D. Jannach, A systematic review and taxonomy of explanations in decision support and recommender systems, User Modeling and User-Adapted Interaction 27 (2017) 393–444. URL: https://doi.org/10.1007/s11257-017-9195-0. doi:10.1007/s11257-017-9195-0.

[6] N. Tintarev, J. Masthoff, Explaining recommendations: design and evaluation, Springer US, Boston, MA, 2015, pp. 353–382. URL: https://doi.org/10.1007/978-1-4899-7637-6_10. doi:10.1007/978-1-4899-7637-6_10.

[7] M. Stickdorn, J. Schneider, K. Andrews, This is service design thinking: Basics, tools, cases, Wiley, 2011.

[8] A. J. Berre, Y. Lew, B. Elvesaeter, H. de Man, Service Innovation and Service Realisation with VDML and ServiceML, 2012 IEEE 16th International Enterprise Distributed Object Computing Conference Workshops 0 (2013) 104–113. doi:http://doi.ieeecomputersociety.org/10.1109/EDOCW.2013.18.

[9] N. Mauro, L. Ardissono, S. Capecchi, R. Galioto, Service-aware interactive presentation of items for decision-making, Applied Sciences, Special Issue Implicit and Explicit Human-Computer Interaction 10 (2020) 5599. URL: https://www.mdpi.com/2076-3417/10/16/5599. doi:10.3390/app10165599.

[10] A. Richardson, Using Customer Journey Maps to improve customer experience, Harvard Business Review, 2015.

[11] K. I. Muhammad, A. Lawlor, B. Smyth, A live-user study of opinionated explanations for recommender systems, in: Proceedings of the 21st International Conference on Intelligent User Interfaces, IUI '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 256–260. URL: https://doi.org/10.1145/2856767.2856813. doi:10.1145/2856767.2856813.

[12] L. Chen, F. Wang, Explaining recommendations based on feature sentiments in product reviews, in: Proceedings of the 22nd International Conference on Intelligent User Interfaces, IUI '17, Association for Computing Machinery, New York, NY, USA, 2017, p. 17–28. URL:

https://doi.org/10.1145/3025171.3025173. doi:10.1145/3025171.3025173.

[13] M. Millecamp, N. N. Htun, C. Conati, K. Verbert, What's in a user? towards personalising transparency for music recommender interfaces, in: Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization, UMAP '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 173–182. URL: https://doi.org/10.1145/3340631.3394844. doi:10.1145/3340631.3394844.

[14] J. C. Chang, N. Hahn, A. Perer, A. Kittur, SearchLens: composing and capturing complex user interests for exploratory search, in: Proceedings of the 24th International Conference on Intelligent User Interfaces, IUI '19, ACM, New York, NY, USA, 2019, pp. 498–509. URL: http://doi.acm.org/10.1145/3301275.3302321. doi:10.1145/3301275.3302321.

[15] C. di Sciascio, V. Sabol, E. E. Veas, Rank as you go: User-driven exploration of search results, in: Proceedings of the 21st International Conference on Intelligent User Interfaces, IUI '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 118–129. URL: https://doi.org/10.1145/2856767.2856797. doi:10.1145/2856767.2856797.

[16] K. Muhammad, A. Lawlor, E. Rafter, B. Smyth, Great explanations: opinionated explanations for recommendations, in: E. Hüllermeier, M. Minor (Eds.), Case-Based Reasoning Research and Development, Springer International Publishing, Cham, 2015, pp. 244–258. doi:10.1007/978-3-319-24586-7_17.

[17] Y. Lu, R. Dong, B. Smyth, Why I like it: Multi-task learning for recommendation and explanation, in: Proceedings of the 12th ACM Conference on Recommender Systems, RecSys '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 4–12. URL: https://doi.org/10.1145/3240323.3240365. doi:10.1145/3240323.3240365.

[18] G. Marchionini, Exploratory search: from finding to understanding, Communications of the ACM 49 (2006) 41–46. URL: https://doi.org/10.1145/1121949.1121979. doi:10.1145/1121949.1121979.

[19] M. Hearst, A. Elliott, J. English, R. Sinha, K. Swearingen, K.-P. Yee, Finding the flow in web site search, Communications of the ACM 45 (2002) 42–49. URL: http://doi.acm.org/10.1145/567498.567525. doi:10.1145/567498.567525.

[20] N. Mauro, L. Ardissono, M. Lucenteforte, Faceted search of heterogeneous geographic information for dynamic map projection, Information Processing & Management 57 (2020) 102257. URL: https://doi.org/10.1016/j.ipm.2020.102257. doi:10.1016/j.ipm.2020.102257.

[21] M. Cheng, X. Jin, What do airbnb users care about? an analysis of online review comments, International Journal of Hospitality Management 76 (2019) 58 – 70. URL: http://www.sciencedirect.com/science/article/pii/S0278431917307491. doi:https://doi.org/10.1016/j.ijhm.2018.04.004.

[22] X. Xu, Y. Lu, The antecedents of customer satisfaction and dissatisfaction toward various types of hotels: A text mining approach, International Journal of Hospitality Management 55 (2016) 57 – 69. URL: http://www.sciencedirect.com/science/article/pii/S0278431916300202. doi:https://doi.org/10.1016/j.ijhm.2016.03.003.

[23] L. Ren, H. Qiu, P. Wang, P. M. Lin, Exploring customer experience with budget hotels: Dimensionality and satisfaction, International Journal of Hospitality Management 52 (2016) 13 – 23. URL: http://www.sciencedirect.com/science/article/pii/S0278431915001486. doi:https://doi.org/10.1016/j.ijhm.2015.09.009.

[24] N. Mauro, Z. F. Hu, L. Ardissono, Service-oriented justification of recommender system suggestions, in: Proceedings of 18th IFIF TC13 International Conference on Human-Computer Interaction, Springer, Bari, Italy, 2021.
[25] R. E. Kirk, Experimental design: Procedures for the behavioral sciences, SAGE Publications, Inc., 2013. Https://www.doi.org/10.4135/9781483384733.
[26] R. Alnemr, S. König, T. Eymann, C. Meinel, Enabling usage control through reputation objects: A discussion on e-commerce and the internet of services environments, J. Theor. Appl. Electron. Commer. Res. 5 (2010) 59–76.
[27] S. Capecchi, P. Pisano, Reputation by design: Using VDML and service ML for reputation systems modeling, in: 2014 IEEE 11th International Conference on e-Business Engineering (ICEBE), IEEE Computer Society, Los Alamitos, CA, USA, 2014, pp. 191–198. URL: https://doi.ieeecomputersociety.org/10.1109/ICEBE.2014.41. doi:10.1109/ICEBE.2014.41.
[28] Y. Yao, S. Ruohomaa, F. Xu, Addressing common vulnerabilities of reputation systems for electronic commerce, J. Theor. Appl. Electron. Commer. Res. 7 (2012) 1–20. URL: https://doi.org/10.4067/S0718-18762012000100002. doi:10.4067/S0718-18762012000100002.
[29] N. Mauro, Z. F. Hu, L. Ardissono, G. Izzi, A service-oriented perspective on the summarization of recommendations, in: Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, ACM, Utrecht, The Netherlands, 2021.