An Analysis of User Click Behaviour in Online Hotel Search

Discussion Paper

Emanuele Cavenaghi[1,*], Lorenzo Camaione[4], Paolo Minasi[4], Gabriele Sottocornola[1], Fabio Stella[3] and Markus Zanker[1,2]

[1]*Free University of Bozen-Bolzano, Bolzano, Italy*
[2]*University of Klagenfurt, Austria*
[3]*University of Milano-Bicocca, Milano, Italy*
[4]*Bravonext SA t/a lastminute.com, Chiasso, Switzerland*

Abstract

Recommender Systems were created to support users in situations of information overload. However, users are consciously or unconsciously influenced by many factors in their decision making. Hence, we aim at exploring how these factors influence user choices in the context of online hotel search and booking. Specifically, we focused our study on (i) ranking position, (ii) price, (iii) rating, and (iv) number of reviews when analyzing users' click behavior. The results showed that there were "two elephants in the room": position and price, that heavily influenced the user decision-making and need to be taken into account when, for instance, trying to learn user preferences from offline data in order to bootstrap a recommender system.

Keywords

recommender Systems, tourism, online hotel search, data analysis, position, price,

1. Introduction

Recommender Systems (RSs) are algorithms developed for helping users to find items of interest [1]. The massive volume of information available on the web leads to the problem of information overload, which increases the need for delivering effective and timely recommendations. The main idea behind these methods is to know users' interests, based on their feedback on past interactions with items, in order to recommend new unseen items matching their preferences.

RSs are extensively applied into E-Tourism domains [2, 3] to recommend destinations/travel packages [4, 5, 6], point of interest [7, 8, 9], restaurants [10, 11] and so on. Furthermore, travelling generally includes accommodating in some place for one or more nights. In this respect, it is fundamental to suggest appropriate *properties*[1] to the users, exploiting both contextual features

IIR2022: 12th Italian Information Retrieval Workshop, June 29 - June 30th, 2022, Milan, Italy
*Corresponding author.
ecavenaghi@unibz.it (E. Cavenaghi); lorenzo.camaione@lastminute.com (L. Camaione); paolo.minasi@lastminute.com (P. Minasi); gabriele.sottocornola@unibz.it (G. Sottocornola); fabio.stella@unimib.it (F. Stella); Markus.Zanker@unibz.it (M. Zanker)
0000-0002-0235-0421 (E. Cavenaghi); 0000-0001-9983-2330 (G. Sottocornola); 0000-0002-4805-5516 (M. Zanker)

[1]With the term properties we refer to any type of accommodation like hotels, apartments, guest houses, etc.

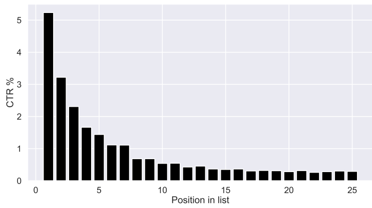


Figure 1: CTR for the top 25 positions in the recommended list. The x-axis report the position in the recommended list, while the y-axis report the CTR achieved in that position.

(such as season and place) and user's preferences. In this regard, in the last years, many RSs were developed to recommend hotels in the context of online booking [12, 13, 14].

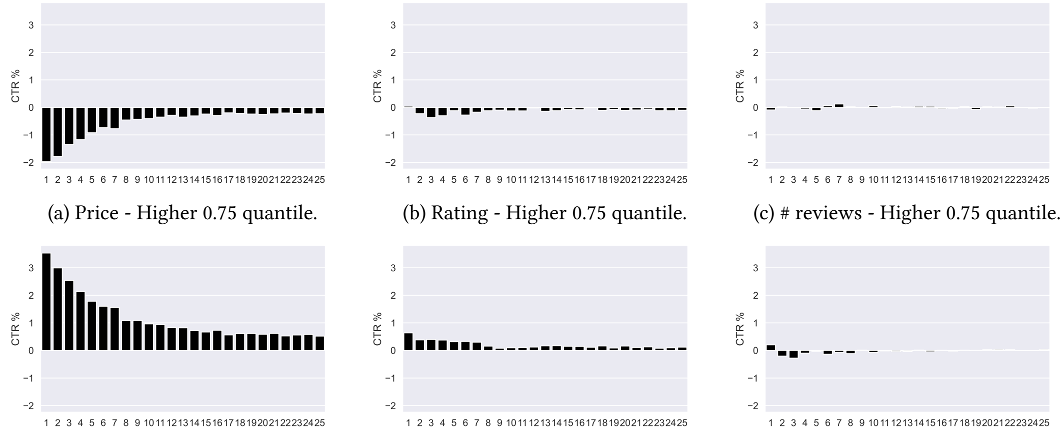
The main contribution of our work is to study the influence of several variables related to user decisions in the online hotel search context. Therefore, for each variable, we tried to answer the following research question: *"Does the variable influence the user decision-making?"* Specifically, we focus our attention on the following variables: (i) the property's position in the recommended list, (ii) the price of the property, (iii) the average rating of the property and (iv) the number of reviews of the property.

2. Data Analysis

The dataset is composed by ca. 130, 000 recommended lists, each composed by 25 properties showed in the same page, collected on a metasearch booking platform in the period between Nov 2021 - Apr 2022.

We report in Figure 1 the Click-Through Rate (CTR) computed for each of the 25 positions in the recommended list. For the remainder of the paper, we will refer to the CTR distribution shown in Figure 1 as the *a-priori CTR*. In Figure 2, we report the difference between the CTR distribution conditioned on price, rating and number of reviews and the *a-priori CTR*. For each variable, we ran six different experiments with the following filter conditions: (i) the highest value, (ii) values higher than the 0.75 quantile, (iii) values higher than the mean, (iv) values lower than the mean, (v) values lower than the 0.25 quantile and (vi) the lowest value. Each condition is referred to the values within each recommended list. For example, in Figure 2a, we computed the CTR distribution taking into account only the properties that had a price higher than the 0.75 quantile price within each recommended list, i.e. we removed from the dataset all the properties with a price lower than the 0.75 quantile in each recommended list, and we subtracted the a-priori CTR distribution to obtain the plot. Due to space constraints, we report only the experiments for values higher/lower than the 0.75/0.25 quantile. However, for the other conditions the results are similar to the reported ones. Finally, for all the plots in Figure 2, we kept the same scale to highlight the different influence of the variables.

Influence of Ranking Position It is well-know in the literature that the ranking position affects the user choice, *Feigenbaum and Simon* [15] and *Joachims et al.* [16], and our analysis confirmed this assertion. From Figure 1 it was clear that the position had an important effect on



(a) Price - Higher 0.75 quantile.
(b) Rating - Higher 0.75 quantile.
(c) # reviews - Higher 0.75 quantile.
(d) Price - Lower 0.25 quantile.
(e) Rating - Lower 0.25 quantile.
(f) # reviews - Lower 0.25 quantile.

Figure 2: Difference between conditioned CTR and a-priori CTR, showed in Figure 1, for each position in the recommended list.

the CTR. In fact, the CTR for position *k* was always higher than the CTR for subsequent positions. For example, in our dataset, the CTR for the first position was 5.23%, while the CTR for the last position was 0.29%. This means that the probability of clicking the first position was 18 times the probability of clicking on position 25. However, while for the first 10 positions this was evident, the difference in CTR for the positions between 10 and 25 was close to zero. The answer to the research question: "Does the ranking position influence the user decision-making?" is yes: the ranking position clearly influence the user choice.

Influence of Price Figure 2a and Figure 2d report the results for price when we considered only the properties having a price higher than the 0.75 quantile and lower than the 0.25 quantile, respectively. For each position in the list, the CTR was higher than the a-priori CTR if we considered properties with a price lower than the 0.25 quantile. This clearly means that lower prices positively influenced the users' propensity of clicking on a property and the opposite happened if we considered properties with a price higher than the 0.75 quantile: for higher prices the CTR was lower. The answer to the question: "Does the price influence the user decision-making?" is clearly yes. Price influences the user decision-making both positively and negatively as stated in *Lockyer* [17] and *Stávková et al.* [18].

Influence of Rating Figure 2b reports the analysis for the properties with a rating higher than the 0.75 quantile and Figure 2e for the properties with a rating lower than the 0.25 quantile. While price influenced the CTR with an absolute difference of 2-3% w.r.t. a-priori CTR, for the rating this difference was greater than 0.5% just in one case (for the first position in Figure 2e). Assuming that these differences are statistically significant, we expected higher rating to lead to higher CTR, while for lower rating we expected lower CTR. Instead, we observed the exact opposite phenomenon. In this case, the answer to the research question: "Does the rating influence the user decision-making?" was not so clear. While for position and price the effect was relevant, for the rating it was more difficult to give a unique answer: the rating seemed to have a very weak influence or no influence at all. These results are in contrast with *Gavilán et al.* [19] and *Coba et al.* [20, 21] that found that users were influenced by rating. However, these works ran user studies considering only rating and number of reviews, while for us price had a greater influence that probably confounded the effect of rating.

Influence of Number of Reviews Finally, we report in Figure 2c the analysis for the properties that had a number of reviews higher than the 0.75 quantile, and in Figure 2f the analysis for the properties with a number of reviews lower than the 0.25 quantile. In this case, it was clear that the number of reviews did not influence the user's choice at all. These results are in contrast to other studies in the literature, such as those by *Coba et al.* [20, 21], in which users tend to trust the rating of the items only if they are made up on a sufficient number of reviews. To verify the existence of a sufficient number of reviews above which users trust the rating, we compared the distribution of the CTR for properties with more and less than 35 reviews. We found that there was a positive influence for the properties above the threshold and this confirmed our claims. Given our data, the answer to the last research question: "Does the number of reviews influence the user decision-making?" was clear: the number of reviews did not influence user decision-making when the number of the reviews exceeds a certain threshold.

3. Conclusions

In this paper, we studied how different variables affect the user decision-making. Specifically, we took into account the following variables: (i) ranking position, (ii) price, (iii) rating and (iv) number of reviews, and we measured their influence by observing changes in the CTR distribution.

Our analysis confirmed that the ranking position and the price heavily influenced the user's choice: a property in the top positions had a greater probability of receiving a click than a property in the last positions and a high price generally discouraged users.

On the other hand, differently from the previous literature by *Gavilán et al.* [19] and *Coba et al.* [20, 21], we found that the average rating had a weak influence on the user's choice. Probably, this influence was heavily confounded by the influence of the price because, as stated by *De Langhe et al.* [22], average ratings are influenced by price: products tend to have a higher rating when they have a higher price. Even regarding the number of reviews, we discovered that this variable did not have a significant effect on user's choice. This was because above a threshold, about 35 reviews in our case, the user trusted the average rating, and a larger number of reviews did not change the user's perception.

References

[1] D. Jannach, M. Zanker, A. Felfernig, G. Friedrich, Recommender systems: an introduction, Cambridge University Press, 2010.

[2] J. Borràs, A. Moreno, A. Valls, Intelligent tourism recommender systems: A survey, Expert systems with applications 41 (2014) 7370–7389.
[3] F. Ricci, Recommender systems in tourism, Handbook of e-Tourism (2020) 1–18.
[4] Q. Liu, Y. Ge, Z. Li, E. Chen, H. Xiong, Personalized travel package recommendation, in: 2011 IEEE 11th international conference on data mining, IEEE, 2011, pp. 407–416.
[5] F. Lorenzi, S. Loh, M. Abel, Personaltour: a recommender system for travel packages, in: 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, volume 2, IEEE, 2011, pp. 333–336.
[6] Z. Yu, H. Xu, Z. Yang, B. Guo, Personalized travel package with multi-point-of-interest recommendation based on crowdsourced user footprints, IEEE Transactions on Human-Machine Systems 46 (2015) 151–158.
[7] T. Ruotsalo, K. Haav, A. Stoyanov, S. Roche, E. Fani, R. Deliai, E. Mäkelä, T. Kauppinen, E. Hyvönen, Smartmuseum: A mobile recommender system for the web of data, Journal of Web Semantics 20 (2013) 50–67.
[8] W.-S. Yang, S.-Y. Hwang, itravel: A recommender system in mobile peer-to-peer environment, Journal of Systems and Software 86 (2013) 12–20.
[9] C. Binucci, F. De Luca, E. Di Giacomo, G. Liotta, F. Montecchiani, Designing the content analyzer of a travel recommender system, Expert Systems with Applications 87 (2017) 199–208.
[10] A. Gupta, K. Singh, Location based personalized restaurant recommendation system for mobile environments, in: 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), IEEE, 2013, pp. 507–511.
[11] J. Zeng, F. Li, H. Liu, J. Wen, S. Hirokawa, A restaurant recommender system based on user preference and location in mobile environment, in: 2016 5th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI), IEEE, 2016, pp. 55–60.
[12] M. P. O'Mahony, B. Smyth, Learning to recommend helpful hotel reviews, in: Proceedings of the third ACM conference on Recommender systems, 2009, pp. 305–308.
[13] A. Levi, O. Mokryn, C. Diot, N. Taft, Finding a needle in a haystack of reviews: cold start context-based hotel recommender system, in: Proceedings of the sixth ACM conference on Recommender systems, 2012, pp. 115–122.
[14] M. Nilashi, O. bin Ibrahim, N. Ithnin, N. H. Sarmin, A multi-criteria collaborative filtering recommender system for the tourism domain using expectation maximization (em) and pca–anfis, Electronic Commerce Research and Applications 14 (2015) 542–562.
[15] E. A. Feigenbaum, H. A. Simon, A theory of the serial position effect, British Journal of Psychology 53 (1962) 307–320.
[16] T. Joachims, L. Granka, B. Pan, H. Hembrooke, F. Radlinski, G. Gay, Evaluating the accuracy of implicit feedback from clicks and query reformulations in web search, ACM Transactions on Information Systems (TOIS) 25 (2007) 7–es.
[17] T. Lockyer, The perceived importance of price as one hotel selection dimension, Tourism Management 26 (2005) 529–537.
[18] J. Stávková, L. Stejskal, Z. Toufarová, et al., Factors influencing consumer behaviour, ZEMEDELSKA EKONOMIKA-PRAHA- 54 (2008) 276.
[19] D. Gavilan, M. Avello, G. Martinez-Navarro, The influence of online ratings and reviews on hotel booking consideration, Tourism Management 66 (2018) 53–61.

[20] L. Coba, L. Rook, M. Zanker, P. Symeonidis, Decision making strategies differ in the presence of collaborative explanations: two conjoint studies, in: Proceedings of the 24th International Conference on Intelligent User Interfaces, 2019, pp. 291–302.
[21] L. Coba, L. Rook, M. Zanker, Choosing between hotels: impact of bimodal rating summary statistics and maximizing behavioral tendency, Information Technology & Tourism 22 (2020) 167–186.
[22] B. De Langhe, P. M. Fernbach, D. R. Lichtenstein, Navigating by the stars: Investigating the actual and perceived validity of online user ratings, Journal of Consumer Research 42 (2016) 817–833.