



UvA-DARE (Digital Academic Repository)

Effects of Position and Time Bias on Understanding Onsite Users' Behavior

Hashemi, S.H.; Hupperetz, W.; Kamps, J.; van der Vaart, M.

DOI

[10.1145/2854946.2855004](https://doi.org/10.1145/2854946.2855004)

Publication date

2016

Document Version

Final published version

Published in

CHIIR'16

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Hashemi, S. H., Hupperetz, W., Kamps, J., & van der Vaart, M. (2016). Effects of Position and Time Bias on Understanding Onsite Users' Behavior. In *CHIIR'16: proceedings of the 2016 ACM Conference on Human Information Interaction and Retrieval : March 13-17, 2016, Carrboro, NC, USA* (pp. 277-280). The Association for Computing Machinery. <https://doi.org/10.1145/2854946.2855004>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Effects of Position and Time Bias on Understanding Onsite Users' Behavior

Seyyed Hadi Hashemi Wim Hupperetz Jaap Kamps Merel van der Vaart

University of Amsterdam, Amsterdam, The Netherlands
{hashemi|W.M.H.Hupperetz|kamps|M.J.vanderVaart}@uva.nl

ABSTRACT

The existence of different biases in logged users' behavior makes it difficult to extract realistic topical and social information from users' interaction logs (e.g., query logs). To understand users' behavior and their interests in the cultural heritage domain, we have logged onsite user interaction logs of visits in a museum. This prompts the question on the reliability of the social information being gathered from the onsite logs: How does the position of museum objects affect users' behavior in the museum? How does order of visiting point of interests affect their dwell-time in front of each point of interest? How do different users' characteristics affect their behavior in the museum? In short, what are different kinds of biases that should be considered in the onsite logs? Our main findings are the following: First, there is a considerable position bias, which is due to the design of the exhibition and should be considered during extraction of social signals from the log. Second, there is a bias in the amount of time that users spend for interacting with the point of interests and the order of picking them to visit. This shows a fatigue on users' interactions while they are reaching to the end of the exhibition. Third, we find out some variations among the users' visit, which shows context is an important factor to consider while using onsite logs for different purposes.

Categories and Subject Descriptors: H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—*Query formulation, Search process, Selection process*

Keywords: Behavioral dynamics; Bias; Internet of things; On-site logs

1. INTRODUCTION

In modern search, interaction logs are one of the main sources of information about user behavior, form a key feature for training ranking algorithms, and are crucial for online and offline evaluation. However, due to existence of different kinds of biases in different kinds of human information interaction logs, extracting realistic behavioral information

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHIIR '16, March 13–17, 2016, Carboro, NC, USA

© 2016 Copyright held by the owner/author(s). Publication rights licensed to ACM. ISBN 978-1-4503-3751-9/16/03...\$15.00

DOI: <http://dx.doi.org/10.1145/2854946.2855004>

from the logs and understanding users' actual behavior is so challenging. Due to the adoption of mobile devices and the 'internet of things,' interaction logs have gained a physical component posing additional constraints on observed user behavior.

In the search domain, biases in query logs has been studied for a long time [2, 4]. Eye tracking studies for understanding how users interact with search results show that users tend to click on the top of the rank list (i.e., position rank bias) and not continuing search after finding a relevant page [2]. Trust bias is another kind of bias, which is studied for using clickthrough data as implicit feedback [4]. Biases have also been studied in user click models studies, helping to better understand users' behavior [5].

In this paper, we introduce new kinds of biases in human information interaction logs that should be considered while using onsite logs to understand users' behavior and use it for different purposes, like for improving the museum's collection search engine or for unseen object recommendation at the museum. These new onsite biases are different from the ones introduced above. Moreover, unlike the users' behavior in search task where they end their search session after finding a relevant page, museum visitors tend to keep on exploring the museum, willing to visit as many interesting objects as possible.

We observe a number of biases in the onsite logs that might affect on understanding users' behavior. The first one is walk-through bias, in which visitors usually follow a path from a check-in station to a check-out station. The second one is the position rank bias, in which visitors tend to spend more time in front of the first object being shown to them in the museum. The last one is the time rank bias, in which visitors tend to spend more time in front of objects they decided to visit first, regardless of the position of those objects. We have also observed variation in users' behavior based on the context of their visit, which is important to be considered as well.

In this paper, our main aim is to study the question: *What are different kinds of biases in onsite logs that might affect on extraction of users' interests based on their onsite behavior?* Specifically, we answer the following research questions:

1. *How do position of objects and time of the visit affect on users' behavior?*
 - (a) *How do users tend to walk through real objects in museums?*
 - (b) *How does a position rank bias affect users dwell-time in front of each point of interest?*

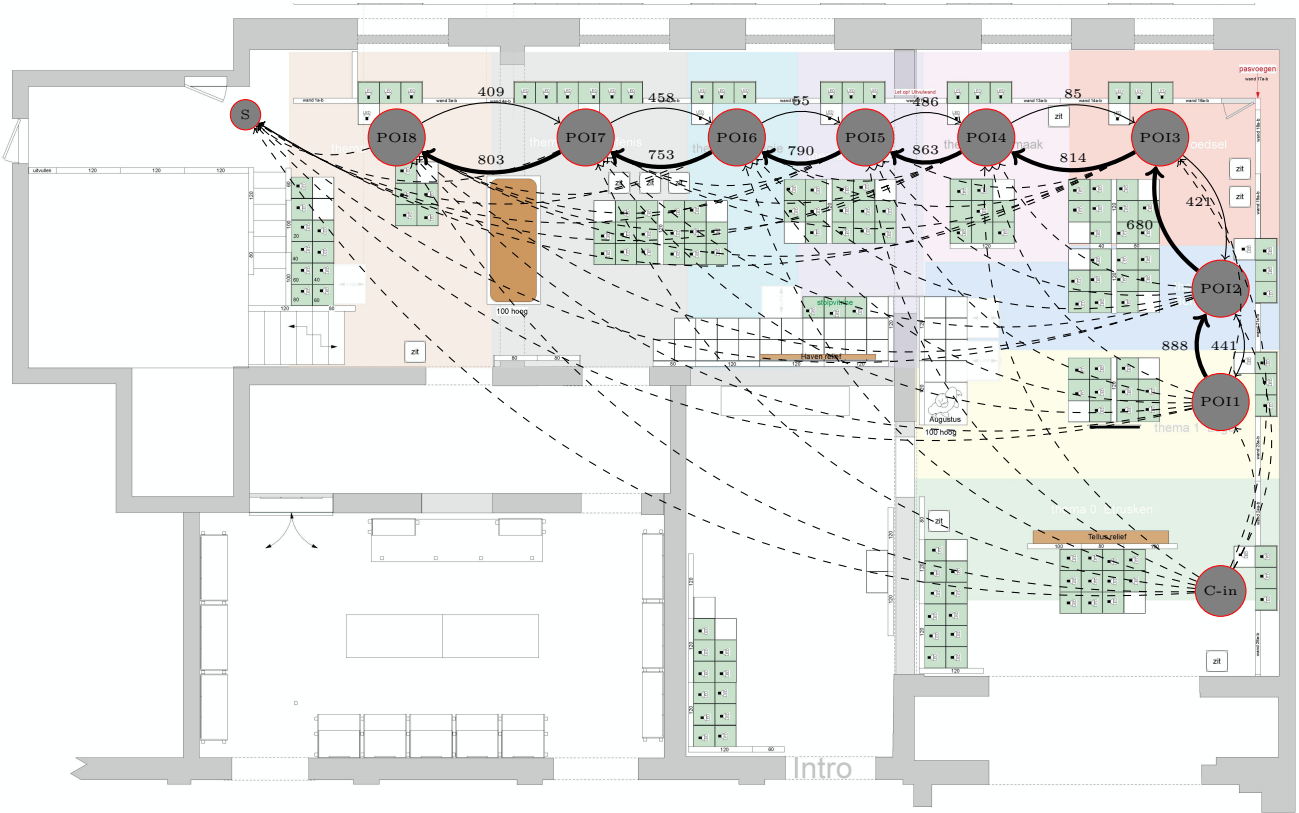


Figure 1: Walk through position bias: dominant transitions from check-in ($C\text{-in}$) to check-out (S). Numbers on the edges show count of users' movements between point of interests.

(c) How does a time rank bias affect users dwell-time in front of each point of interest?

2. What is the effect of users' profile and their visit context on the position bias?

The rest of the paper is organized as follows. Section 2 details the experimental data being used in this research. In Section 3, we introduce biases in the onsite logs. Section 4 is devoted to studying users' variation and the effect of biases on them. Finally, we present the conclusions and future work in Section 5.

2. EXPERIMENTAL DATA

In our archeological museum, RFID cards are provided as a key to access to some additional information about objects being shown in the museum. Visitors enter their preference at the beginning of the museum exhibition in order to personalize the content being shown in all of the point of interests. These preferences are perspectives of the narratives, language and the user's age range.

After checking in, users are free to put their keys on RFID readers of point of interests to unlock contents being shown about point of interests. They are free to interact with point of interests in any order. They can watch short movies, interacting with 3D photos of point of interests' objects, or read contents about objects being shown at POIs. At last, users might check out in a summary station, in which they might leave their name, birth date and email. In this paper,

5 months onsite logs of the museum with more than 21,000 sessions is used.

3. BIASES IN ONSITE LOGS

This section studies biases in onsite logs, aiming to answer our first research question: *How do position of objects and time of the visit affect on users' behavior?*

3.1 Walk Through Position Bias

We first look at the question: *How do users tend to walk through real objects in museums?* Specifically, we study visitors walking path by tracking users' activity inside the museum. To this aim, we logged users' interaction with each system using RFID cards. In this experiment, we filtered sessions without interacting with the summary station, and we use more than 5,000 out of 21,000 sessions of the logs. According to Figure 1, there are many moves between different point of interests, but the most frequent ones are bold in the Figure. Although users are free to visit any POI at any order during their visit, they tend to visit point of interests one after the other from check-in to check-out stations. The second most frequent transition is to the previous POI, consistent with location proximity as a driving force. This makes a considerable bias in users' behavior extracted from onsite logs. More positively, the bias in onsite logs makes visitors behave more closely to simple click models [1].

The users walk through bias causes two other kinds of biases, namely position rank bias and time rank bias, which will affect on understanding users' behavior and their interests based on their dwell-time in front of POIs.

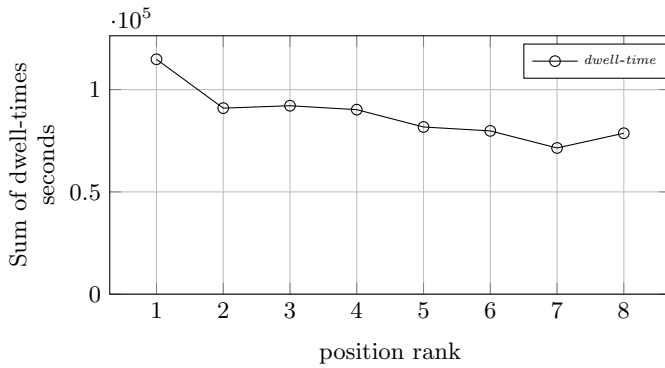


Figure 2: Position rank bias: dwell-time over POIs in exhibition order.

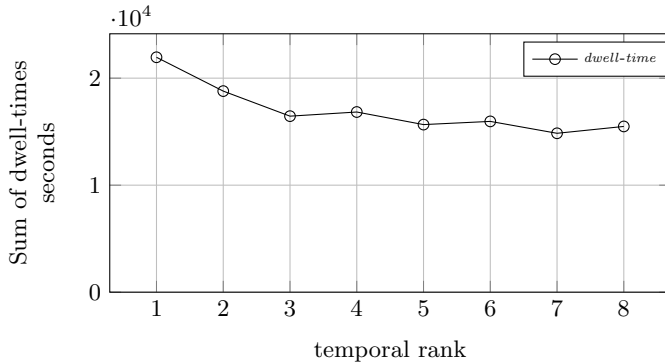


Figure 3: Time rank bias: dwell-time over POIs in order of visit.

3.2 Position Rank Bias

We next look at the question: *How does a position rank bias affect users dwell-time in front of each point of interest?* One of the important social signals in the onsite logs is the dwell-time of each user in front of each point of interests. This information indicates the degree of interest in objects being presented at each point of interests. However, walk-through bias causes a position rank bias, in which users spending more time in front of point of interests installed in the beginning of the exhibition, and the dwell-time decreases by reaching step by step to the end of the exhibition. As it is shown in Figure 2, dwell-time is inverse-proportional to the position rank of the objects in the museum, and the highest dwell-time is for the first object after check in station. Therefore, this bias should be considered while dwell-time is being used as a social signal to understand users' interests. According to this experiment, position of objects affects on the dwell time of users in front of them, which shows the dwell-time should be used with some care.

3.3 Time Rank Bias

As it is mentioned in previous experiments, users are free to visit any point of interest at anytime of their visits. As a result, they are free to not follow objects based on their position order. In this experiment, we analyzed dwell-time of users in each rank of point of interest based on users' visit order. In order to do this experiment, we filter all the sessions that do not interact with all of the point of interests.

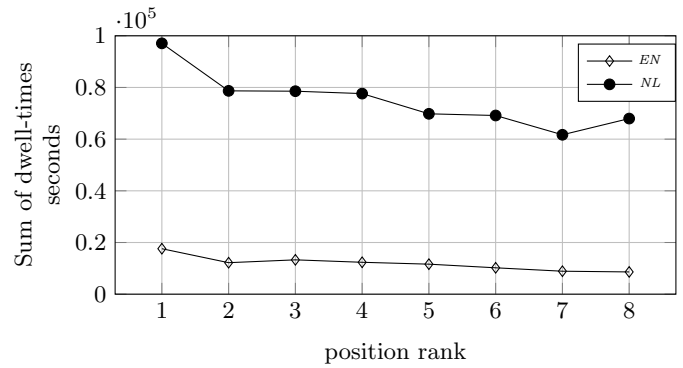


Figure 4: Position rank bias: English versus Dutch language.

According to Figure 3, users tend to spend more time in their earliest stage of their visit. In fact, the dwell-time is decreasing while the time rank of objects are increasing. It is also observed that users tend to spend less time at the end of exhibition, and users' fatigue of using technology in the museum affect on the dwell-time of their visit.

This experiment shows that spending less time at the last stage of the exhibition does not show that users are less interested in the last points of interests in comparison to the first ones; rather, they are less interested in interacting with technology or staying at museums.

4. IMPACTS OF CONTEXTS AND PROFILES

This section answer our second research question: *What is the effect of users' profile and their visit context on the position bias?*

As it is shown in the previous section, dwell-time of users in front of each object has a position and time rank bias. However, it is questionable that visits of which type of users are more affected by the position or time rank bias. To this aim, we stored gender, age, language, time and some other contexts in each session. In the following, some interesting observations based on users' profile are detailed.

Language We first look at differences between behaviors of users who decided to see Dutch content in comparison to users preferred English contents. According to Figure 4, we see a considerable difference in time spent by Dutch and English visitors (with identical content shown in each case) possibly due to foreign visitors spending more time with museum objects. People who preferred to see Dutch contents are less interested in POI7, which is about death. On the other hand, among sessions with English content preference, as expected by the position rank bias, people were more interested in spending time at POI7 rather than POI8. This experiment indicates that different variation of contents being prepared to be shown in deferent contexts affect differently on position rank bias.

Age We now study contributions of age groups in position rank bias. We log 2 different values for age in this log data, namely, adult and child. Figure 5 shows that children do not like POI7 (which is about death) and spend less time in front of death point of interest. Figure 4 and 5 indicate that children and Dutch content of POI7 contribute more on the lower dwell-time of POI7 in comparison to the POI8.

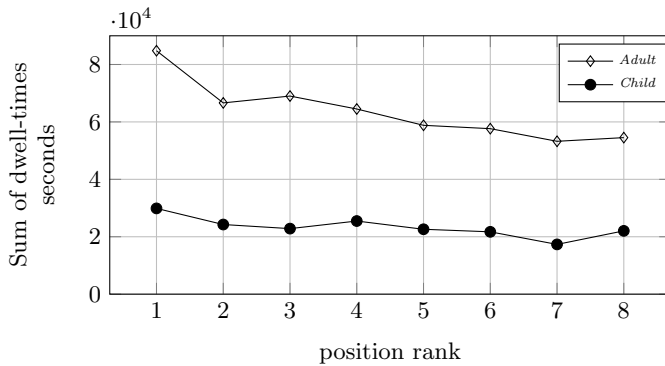


Figure 5: Position rank bias: adults versus children.

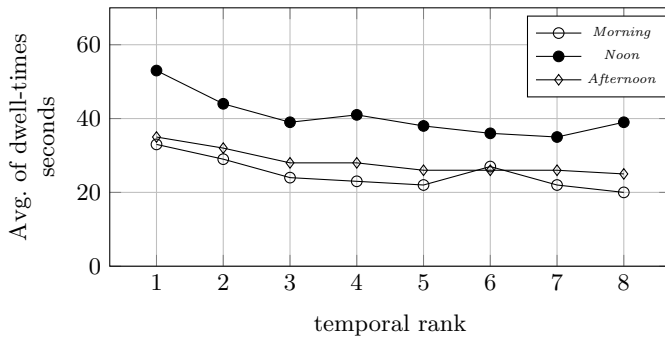


Figure 6: Time rank bias: visit time of day.

This experiment indicates that although adults’ dwell-time at each POI shows a position rank bias in their behavior, children’s behavior is less predictable and their dwell-time are less affected by the position rank bias.

Visit Time At last, we look at behavior of users in different time of a day. Basically, we answer the question: how long users in average tend to stay in front of each point of interests at different time of a day (i.e., morning, noon and afternoon)? We consider morning as a time from 10:00 to 12:00, noon as a time from 12:00 to 14:00, and afternoon as a time from 14:00 to 17:00. As it is shown in Figure 6, users are staying longer and interacting more with systems at noon in comparison to visits in the morning or afternoon. This experiment shows that due to museums opening hours in the afternoon and next activities users might plan to do after their visit in the morning, users might like to finish their visit earlier in morning or afternoon in comparison to others visits at noon.

Moreover, while users are willing to spend plenty of time in the museums, they are less affected by the time rank bias. As it is shown in the diagram of sessions at noon, users in average stayed longer in their last choice of museum object visit in comparison to their 3rd, 5th, 6th and 7th choices. However, this behavior is not true for morning and afternoon sessions.

Variations of users’ behavior in different contexts show that the context of information interaction is an important factor to consider in using the onsite social information. However, similar to other research on using contextual information for search and recommendation [3], using contextual

information increases the variation of users’ behavior and makes it difficult to have a system satisfy different users’ information needs in all contexts, and consequently complicates the evaluation of contextual search systems. In fact, different users behave differently in different contexts, which leads to different degree of biases in each context in the onsite interaction logs. As it is shown in this section, context of visits has some consequences on the time or position rank bias in the onsite logs. Therefore, contexts should be considered as a factor in the rank bias smoothing process for understanding users’ behavior, and evaluation of the contextual suggestion systems going to be used in museums.

5. CONCLUSIONS

We have studied different kinds of biases in onsite logs that could affect on the information extracted for understanding users’ behavior for different purposes like post-visit online search tasks. Specifically, we analyzed how users tend to walk through the exhibition based on a logging their interactions at different POIs.

An analysis based on more than 21,000 sessions shows that they tends to follow point of interests from entrance to the end of the exhibition. In addition, we have analyzed how position rank bias affects dwell-time of visitors in front of each point of interest. An analysis based on more than 5,000 sessions indicates that users tend to interact with the system at the first point of interests more than the others, and the dwell-time is inverse-proportional to the order of objects being shown to the users. Moreover, effects of time rank bias on the dwell-time is also studied, which shows users tend to spend less time in front of objects when they are reaching to the end of their visits, and users’ fatigue is an important factor in understanding users’ behavior based on the onsite logs. Finally, we looked at the effect of context and users’ profile on users’ behaviors. We observed that children are less sensitive to the position bias. We also observed that time of the day play an important role in the amount of time users tend to spend for interacting with point of interests.

Our general observation is that we cannot use dwell-time information of onsite logs without care, and we should consider position and time rank bias as two important factors for smoothing their effects on the dwell-time as a source of evidence of users’ interests.

Acknowledgments This research is funded by the European Community’s FP7 (project meSch, grant # 600851).

References

- [1] N. Craswell, O. Zoeter, M. Taylor, and B. Ramsey. An experimental comparison of click position-bias models. In *WSDM*, pages 87–94, 2008.
- [2] L. A. Granka, T. Joachims, and G. Gay. Eye-tracking analysis of user behavior in www search. In *SIGIR*, pages 478–479, 2004.
- [3] S. H. Hashemi, C. L. Clarke, A. Dean-Hall, J. Kamps, and J. Kiseleva. On the reusability of open test collections. In *SIGIR*, pages 827–830, 2015.
- [4] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR*, pages 154–161, 2005.
- [5] Y. Zhang, W. Chen, D. Wang, and Q. Yang. User-click modeling for understanding and predicting search-behavior. In *KDD*, pages 1388–1396, 2011.