# ADORES: A Diversity-oriented Online Recommender System

Max Chevalier
IRIT UMR 5505 CNRS
Université de Toulouse, France
Max.Chevalier@irit.fr

Damien Dudognon
Overblog, France
damien@overblog.com

Josiane Mothe
IRIT UMR 5505 CNRS
ESPE, Université de Toulouse, France
Josiane.Mothe@irit.fr

## ABSTRACT

Browsing a content platform usually does not require a user identification. In this context, personalized approaches can not be used since no information related to the user is available. In that case, it is important to consider the variety of potential interests of users when providing recommendations. In this paper, we propose a scalable recommendation diversity-oriented model which considers solely the current visited document, the available collection and the past clickthrough documents to produce a list of diversified recommendations. A learning phase is integrated to improve recommendation relevance along time. Our proposals are validated through several experiments.

## CCS Concepts

•**Information systems** → **Recommender systems;** *Information retrieval;* World Wide Web;

## Keywords

Recommender Systems; Blog Platform; Diversity, Enterprise

## 1. INTRODUCTION

Many web content providers face the problem of increasing volumes of data and of fitting the users expectation in terms of providing them with relevant information. We consider the two sides of the same coin that exist in information provider platforms: the end-users point of view and the platform editors point of view. With regard to the users, the convenience is to be assisted in information seeking and to get interesting related articles with a limited effort. For the blog platform editors, the challenges are to attract users by offering them efficient and effective services and providing them with relevant content to keep users captive as long as possible in order to generate a more consistent traffic.



*SAC'16, April 4-8, 2016, Pisa, Italy*



http://dx.doi.org/xx.xxxx/xxxxxxx.xxxxxxx

Recommender systems (RS) have been identified as an interesting solution to make users captive. They automatically provide content that is potentially relevant to the user while maximizing the fact that a user will stay on the platform.

However, RS have also some drawbacks. First, RS can be perceived as deemed intrusive if the recommended items are not relevant [1]. It is essential for the system to identify the real interests of the user to respond adequately.

In this paper, we propose a recommendation model based on diversity which has been evaluated on an online blog platform. It is named **ADORES** for *A Diversity-oriented Online REcommender System*. We also take into account a dimension often overlooked in the academic literature: industrial constraints specific to platforms with high traffic and scalability in these environments.

## 2. ADORES MODEL

We started our research with a preliminary study to validate the assumption that various ranking functions retrieve different sets of relevant documents, and then show that their aggregation positively impacts the accuracy and diversity of final results, and increases the users' satisfaction (more details in [2]). In line with the literature of data fusion, the results from this preliminary study shows it is worth fusing various lists of results.

In our context, the real users' expectations are unknown and we have no information allowing us to infer them. Therefore, we can not make any assumption about expectations. To solve this cold start issue, we propose a three-step process.

The first step selects the best recommendations from the different ranking functions which provide documents by decreasing score order. The documents most likely to be relevant are located on top of the lists [2]. By limiting the selection to the $N$ first documents from each function, we maximize the chances of obtaining relevant recommendations and improve the representativeness of the different aggregated functions. A small number of documents for each function also helps to insure an acceptable processing time.

The second step aggregates the results from different ranking functions to produce a unique and diversified set of recommendations in order to satisfy as many users as possible. Our proposal neither relies on a combination of scores, which are not necessarily comparable and standardized, nor on coefficients intended to reflect the importance of ranking functions. In fact, we chose to select the best representatives of the various ranking functions in variable proportions, and thus ensure a greater diversity of results to satisfy a wider user panel.

Finally, in the last step the recommendations are reordered by two successive phases: documents occurring the most in the various lists are first positioned on top of the final ranking, then documents that occur the same number of times are reordered according to the *Inverse Average Inverse Rank* [5]. Thus, the documents occurring in various lists are favored insofar as they meet several interests and therefore are more likely to meet a large number of users' interests.

In order to recommend most adapted items, a learning phase completes our model. It adjusts the proportion of recommendations coming from each function to catch the dimensions to which this information lead to. The learning phase is based on clickthroughs. We determine the proportion of clicks for each function and then adapt the final recommendation list in putting more items coming from the most popular function. At least one item coming from each function is kept in order to assure a minimal diversity in the final list.

# 3. MEASURING EFFECTIVENESS

The evaluation focuses on the implementation of the proposed model on the *Overblog* platform which provides a real context to evaluate our model.

## 3.1 Protocol and metrics

For this experiment, we compared the results of our model (*fused*) with three ranking functions: *mlt* (using the most weighted terms from the visited document), *search* (based on title of the visited document) and *blog* (which considers the blog posts from the same blog).

The aggregation model is based on selecting the two best recommendations from each of the three measures. In that way, we ensure the representativeness of each ranking function and the diversity of recommandation, considering different dimensions (content, membership of a blog). The computational complexity is therefore limited.

The experiment was divided into three periods to successively evaluate the aggregation model and the learning phase.

To evaluate the performance of our model, we used the number of clicks on the proposed recommendations which is a good approximation of relevance perceived by users and often used to evaluate online systems, that is to say without reference collection [3, 4].

## 3.2 Results

In order to demonstrate the contribution of the aggregate function, we initially focused only on the comparison of the functions considered individually with our aggregate function *fused*. On average and regardless of the function used, the CTR (clickthrough rate) was 3.89%. We find that the three functions have a lower CTR (less than 4%), while our model leads to a higher CTR (4.70%). These results show the interest of the aggregating approach compared to the functions considered individually.

The advantage of the aggregation function is also shown when considering the distribution of clicks between recommendations of the constituent list of *fused*, which selects the two best results of each of the other three measures. A significant part of these recommendations (24.02%) does not come from content-based function.

To evaluate the learning phase, the initial lists of recommendations (used in the first period) have been recalculated by integrating clickthrough data. Only the function *fused* was presented to visitors. The average CTR observed reaches 5.43%, that is well above the rate of 4.7% found during the evaluation of the aggregation.

The experiment was conducted in a real context which involves a substantial volume. The collection from which the recommendations were extracted counts about 20 million posts, and during the evaluation, the blocks of recommendations have been proposed 1.5 million times to visitors. Despite this substantial volume of data, the system provides a recommendation in less than one second and thus is operationally viable on a large-scale, e.g. on a real blog platform like Overblog.

# 4. CONCLUSION

The model we proposed provides a partial solution to the cold start problem. It recommends items based on various ranking functions and considers only the visited document as a reference for the recommendation step, avoiding any assumptions about the users' interests. This is an important point for privacy considerations. Diversifying the final list of recommendations helps maximizing the user's satisfaction.

Our approach is also based on a recurring up-date which allows an adaptive weighting of contributions of each ranking function and adaptive ordering of recommendations.

The large scale evaluation conducted on our model integrated into the *Overblog* online platform validates its viability in an industrial context. CTR obtained is also higher than those of non-aggregated approaches.

# 5. REFERENCES


[1] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE TKDE*, 17(6):734–749, 2005.

[2] L. Candillier, M. Chevalier, D. Dudognon, and J. Mothe. Multiple similarities for diversity in recommender systems. *International Journal On Advances in Intelligent Systems*, 5(3 and 4):234–246, 2012.

[3] O. Chapelle and Y. Zhang. A dynamic bayesian network click model for web search ranking. In *WWW '09*, pages 1–10, 2009.

[4] T. Joachims, L. Granka, B. Pan, H. Hembrooke, and G. Gay. Accurately interpreting clickthrough data as implicit feedback. In *SIGIR '05*, pages 154–161, 2005.

[5] M. Siegler, M. J. Witbrock, S. Slattery, K. Seymore, R. E. Jones, and A. G. Hauptmann. Experiments in spoken document retrieval at cmu. In *Text REtrieval Conference*, pages 291–302, 1997.