Performance Prediction in Recommender Systems

Alejandro Bellogín

Universidad Autónoma de Madrid, Escuela Politécnica Superior
Calle Francisco Tomás y Valiente 11, 28049 Madrid, Spain
alejandro.bellogin@uam.es

Abstract. Research on Recommender Systems has barely explored the issue of adapting a recommendation strategy to the user's information available at a certain time. In this thesis, we introduce a component that allows building dynamic recommendation strategies, by reformulating the performance prediction problem in the area of Information Retrieval to that of recommender systems. More specifically, we investigate a number of adaptations of the query clarity predictor in order to infer the ambiguity in user and item profiles. The properties of each predictor are empirically studied by, first, checking the correlation of the predictor output with a performance measure, and second, by incorporating a performance predictor into a recommender system to produce a dynamic strategy. Depending on how the predictor is integrated with the system, we explore two different applications: dynamic user neighbour weighting and hybrid recommendation. The performance of such dynamic strategies is examined and compared with that of static ones.

Keywords: recommender systems, performance prediction, query clarity, personalisation, user modelling.

1 Introduction

The aim of Recommender Systems (RS) is to help users to cope with information overload by suggesting "interesting" items from huge databases or catalogues. Three types of recommender systems are commonly recognised, based on how item suggestions are made [1]: content-based filtering (CBF), collaborative filtering (CF), and hybrid filtering (HF). CBF recommends the user items similar to the ones she preferred in the past, CF recommends the user items that people (called neighbours in the literature) with similar tastes and preferences liked in the past, and HF combines content-based and collaborative filtering approaches. In this context, although many alternatives are possible, the most common form of ground evidence of user preferences, upon which recommendations are generated, consists of explicit numeric ratings for individual items.

In the RS research area, a barely explored issue –as a problem to be addressed by systematic approaches– is how to dynamically adapt a recommendation strategy to the user's preference information available at a certain time. Let us consider the following two examples: neighbourhood building in CF, and ensemble recommenders, as a special case of HF. In both cases, most of existing recommendation approaches are not adaptive in the sense that every user is considered equally, from the system viewpoint. In user-based CF [1], the user's neighbours are only weighted according to the similarity between the user and each neighbour, not considering global characteristics of each neighbour that would make them less appropriate, such as their confidence or trustworthiness [7]. In ensemble recommenders, on the other hand, the weighting factor which aggregates the outputs from individual recommenders is usually the same for every user and item, again, not taking into account inherent properties of the different users and items.

In this thesis, we aim at addressing the following research question: **how to dynamically adapt a recommendation strategy to the available user's information?** For such purpose, we investigate the adaptation of Information Retrieval (IR) performance prediction techniques to RS, where performance prediction refers to the estimation of the performance of an IR system in response to a specific query [5]. More specifically, we draw from the notion of *query clarity* [4] as a basis for finding suitable predictors. In essence, query clarity captures the lack of ambiguity in a query, by computing the distance between the language models induced by the query and the collection. In analogy to query clarity, we hypothesise that the amount of uncertainty involved in user and item data (reflecting ambiguity in the users' tastes, and popularity patterns in the items) may correlate with the accuracy of a system's recommendations. We believe this uncertainty can be captured as the clarity of users or items by performing a reformulation of query clarity. Thus, we could introduce a performance predictor in a recommender system to produce an adaptive recommendation strategy. In this way, we would build dynamic neighbourhoods and ensemble recommenders by exploiting user or item clarity values. Moreover, our approach allows for an interpretation of the weights given to the user, in contrast to other works in the literature [3, 6]. This is possible since these weights are closely related with her ambiguity in the system, which would provide for further explanation to the user about her current performance in the system and how she could improve it.

2 Performance Prediction in Recommender Systems

The main goal of this thesis can be summarised as predicting the best way for combining the available information from a recommender system in a user-basis, such as deciding the user's weight in a hybrid recommender and when building her neighbourhood in a CF approach. For this purpose, we first need to capture the ambiguity in user's tastes, and then introduce it within the recommendation process to build an adaptive strategy.

Starting problem

Our main problem is that of considering the ambiguity in a user's tastes, so that, depending on this prediction, we may adapt the recommendation strategy. For instance, let us suppose a movie recommendation situation, where user tastes are represented by rating-movie pairs, and ratings are in a 1-5 scale. Then, let us assume user *U* with the following tastes: (*Matrix*, 5), (*Star Wars*, 2), (*Titanic*, 1), (*Pretty Woman*, 4). Moreover, let the community be defined as follows (*Matrix*, 4), (*Star Wars*, 5), (*Titanic*, 2), (*Pretty Woman*, 1). Then, it seems user *U* is an ambiguous user because she rates those movies very differently, such as *Matrix* and *Star Wars*, which are movies with tastes commonly shared by the community. Moreover, this user is not only difficult to be recommended from the system point of view, but if we use her as a neighbour in a CF recommender, the performance would probably degrade because, in user-based CF, items liked by neighbours would be recommended. In this situation, if a profile only partially matches the user *U*'s profile, she could receive unexpected suggestions, e.g., the system may recommend *Pretty Woman* if *Titanic* is low rated and *Matrix* is highly rated.

In this context, we then should solve the following problems: 1) define a proper ambiguity predictor, 2) check its predictive power, and 3) evaluate whether the use of the predictor improve the final performance or not. In the next section, we introduce the solutions being developed for the problems described above.

Proposed solution

We state that estimating the user's ambiguity by predicting her performance (in IR terms) could bring relevant contributions to the Recommender Systems community and be used to address the problems mentioned above.

In particular, we have revised the research literature on performance prediction in IR, and adapted some of its more prominent models, such as query clarity [4], for recommendation. In this way, we have defined *user clarity* as follows: $\text{clarity}(u) = \sum_{x \in X} p(x|u) \log p(x|u)/p_c(x)$, where x denotes items, users, or ratings. We assume different probability models, for instance, we may have the probability space X defined as the space of all possible rating values, then we can estimate $p_c(r)$ and $p(r|u)$ using uniform probabilities, or other distribution estimations [8] (more details in [2]).

Once a user performance predictor has been defined, we need to confirm that it has some predictive power. Thus, the first goal is to find strong correlation between the envisioned predictor and a performance metric, which depends on the final application. In this context, Pearson and Spearman correlation coefficients are the most common sources of evidence. Then, also depending on the application, we would incorporate the predictor into the recommendation process in different ways.

Our research, so far, has demonstrated that when a predictor obtains strong (or positive enough) correlation with respect to a performance metric, a significant improvement is obtained when the predictor is introduced into the recommendation process for building an adaptive strategy. For instance, in [2] we describe how a predictor can be used for dynamic neighbour weighting. In this situation, we obtained an improvement of over 5% with respect to the baseline, which in this experiment was the standard (static) CF algorithm. It is worth noticing our method outperformed the baseline, even though Pearson's correlation between the predictor and the performance metric was not very large, obtaining correlation values between 0.15 and 0.20.

Moreover, latest experiments shown consistent results for different ensemble recommenders, such as combination of CBF and CF recommenders, and different types of CF algorithms (user- and item-based). In these cases, Pearson's correlation between our clarity-based predictors and the performance metric hovered around 0.3 and 0.46, and the improvements over the baseline –in this case, a fixed hybrid recommender with the same weight for all the users– were between 3 and 12%.

The obtained results confirm that performance prediction is also possible in recommendation. Two applications have been proposed: dynamic user neighbour weighting and dynamic hybrid recommendation. In both situations, adaptive strategies outperformed non-adaptive ones.

3 Conclusions and Future Work

The main contribution of our work is the idea of inferring the user's performance within a recommender system in order to use it for building adaptive recommendation strategies, such as boosting those neighbours which are predicted to perform better, and weighting differently users or items in ensemble recommenders. Preliminary obtained results are promising and encouraging, showing that our approach is useful, improving the performance of state-of-the-art algorithms.

The main focus of our research concerns the definition of user performance predictors, as well as different applications where it can be used. In this line, this thesis needs to a) find a theoretical background about why some predictors work better than others, i.e., have stronger correlations; and b) explore other input sources apart from ratings. Regarding the first aspect, there are open issues that need further investigation, like for example analysing why some recommenders are more inclined to correlate stronger with respect to different formulations of the same predictor than others. In order to explore other input sources, the next step will be to obtain more heterogeneous datasets where not only ratings, but implicit feedback, time and social relations are available. Furthermore, new performance predictors using this data should be defined and evaluated accordingly.

Acknowledgments. This work was supported by the Spanish Ministry of Science and Innovation (TIN2008-06566-C04-02) and Dirección General de Universidades e Investigación de la Comunidad de Madrid and Universidad Autónoma de Madrid (CCG10-UAM/TIC-5877).

References

1. Adomavicius, G., Tuzhilin, A.: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering 17(6), 734–749 (2005)
2. Bellogín, A., Castells, P.: A Performance Prediction Approach to Enhance Collaborative Filtering Performance. In: Gurrin, C. et al. (eds.) Advances in Information Retrieval, LNCS, vol. 5993, pp. 382–393. Springer, Heidelberg (2010)
3. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M.: Combining Content-Based and Collaborative Filters in an Online Newspaper. In: ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Applications (1999)
4. Cronen-Townsend, S., Zhou, Y., Croft, B. W.: Predicting Query Performance. In: 25th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2002), pp. 299–306. ACM Press, New York (2002)
5. Hauff, C.: Predicting the Effectiveness of Queries and Retrieval Systems. PhD thesis, University of Twente, Enschede (2010)
6. Lathia, N., Hailes, S., Capra, L.: Temporal Collaborative Filtering with Adaptive Neighbourhoods. In: 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009), pp. 796–797. ACM Press, New York (2009)
7. O'Donovan, J., Smyth, B.: Trust in Recommender Systems. In: 10th International Conference on Intelligent User Interfaces (IUI 2005), pp. 167–174. ACM Press, New York (2005)
8. Wang, J., de Vries, A. P., Reinders, M. J. T.: Unified relevance models for rating prediction in collaborative filtering. ACM Trans. Inf. Syst. 26(3), 1–42 (2008)