Probabilistic Collaborative Filtering with Negative Cross Entropy

Alejandro Bellogín[1,3], Javier Parapar[2], Pablo Castells[3]

alejandro.bellogin@uam.es, javierparapar@udc.es, pablo.castells@uam.es

[1] Information Access, Centrum Wiskunde & Informatica

[2] Information Retrieval Lab, University of A Coruña

[3] Information Retrieval Group, Universidad Autónoma de Madrid

Introduction

- **Neighbourhood identification** in memory-based CF algorithms is based on selecting those users who are most similar to the active user according to a certain similarity metric.
- **Hypothesis:** neighbour-based CF techniques may be improved by using Relevance Models (RM) from Information Retrieval (IR) to identify such neighbourhoods (and also to weight them).
- **Experiments:** a **relevance-based language model** has been introduced into a neighbour-based CF algorithm which outperforms other standard techniques in terms of **ranking precision**. Further improvements are achieved when we use a **complete probabilistic representation** of the problem.

Relevance Modelling for Recommendation

We decompose the rating prediction task $\hat{r}(u,i) = C \sum_{v \in N_k(u)} \mathrm{sim}(u,v) r(v,i)$ (*) as:

1. We compute a relevance model for each user in order to capture how relevant any other user would be as a potential neighbour.
   Probability of a neighbour v under the relevance model R_u for a given user u:

$$p(v|R_u) = \sum_{i \in PRS(u)} p(i) p(v|i) \prod_{j \in I(u)} p(j|i) \qquad (1)$$

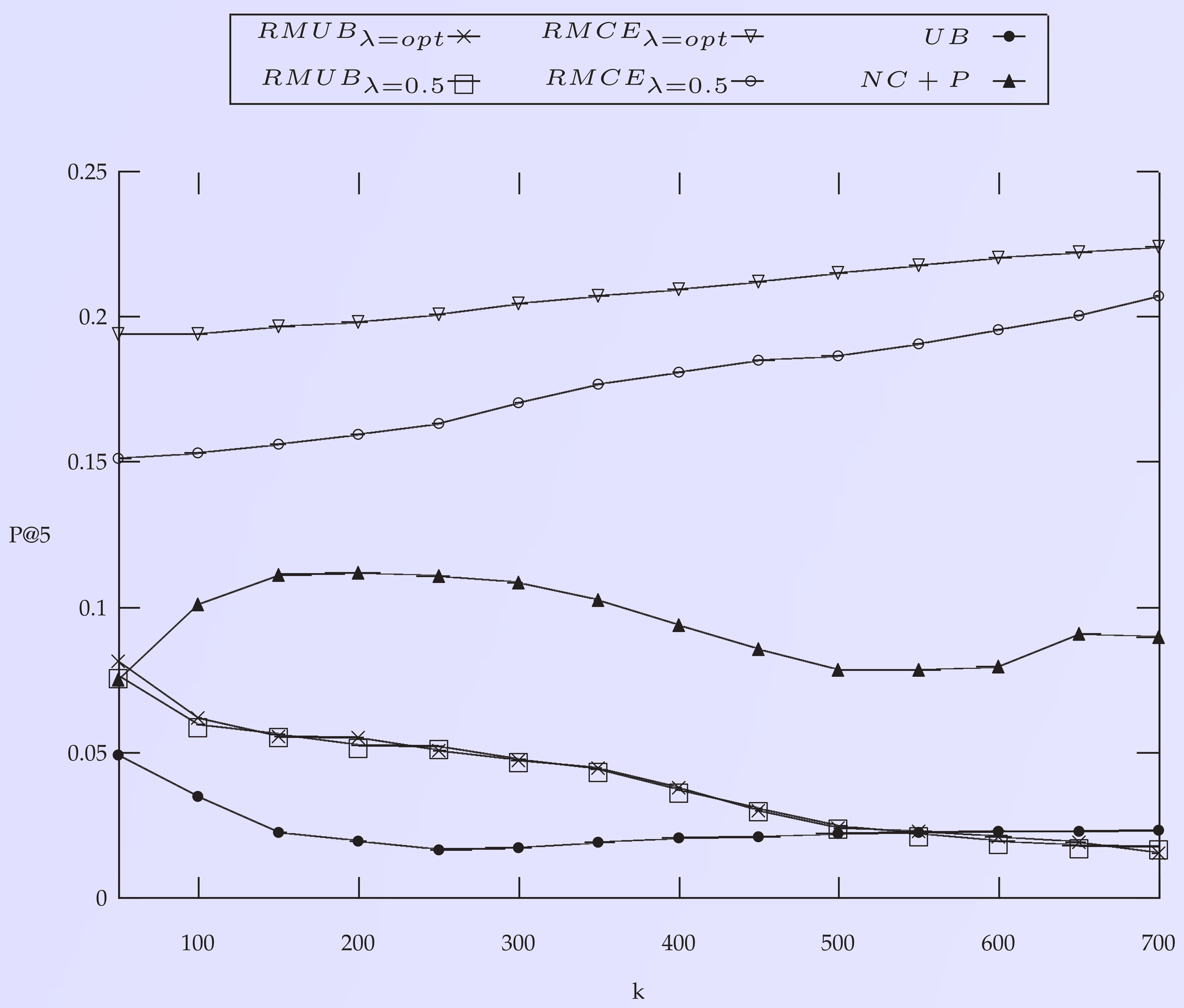
   where $p(i)$ is the probability of the item i in the collection, $p(v|i)$ is the probability of the neighbour v given the item i, and $p(i|j)$ is the conditional probability of item i given another item j. $I(u)$ corresponds to the set of items rated by user u, and $PRS(u) \subset I(u)$ is the subset of items rated by u above some specific threshold.
2. Then we replace the rating prediction by weighted average in CF with the negative cross entropy (from IR) to incorporate the information learnt from the RM:

$$\hat{r}(u,i) = H(p(\cdot|R_u); p(\cdot|i)|N_k(u)) = \sum_{v \in N_k(u)} p(v|R_u) \log p(v|i) \qquad (2)$$

Notation: $\boxed{\textit{RMUB}\text{: Eqs. (1) + (*)}}$ and $\boxed{\textit{RMCE}\text{: Eqs. (1) + (2)}}$

Experiments and Results

Evaluation methodology: *TestItems* [1] (for each user a ranking is generated by predicting a score for every item in the test set).
Baselines: *UB* (user-based CF with Pearson's correlation as similarity measure), **NC+P** (Normalised Cut (NC) with Pearson similarity [2]) (better performing than a matrix factorization algorithm), **UIR** (relevance model for log-based CF [6]), **URM** (rating-based probability estimations [7]).



Evolution of the performance of the compared methods in terms of $P@5$ when varying k on the MovieLens 100K collection.

MovieLens 100K

Method	P@5	nDCG@5	nDCG@10	P@50	cvg
UB	0.049^{cd}	0.041^{cd}	0.047^{cd}	0.056^{ce}	100%
NC+P	0.111^{acde}	0.097^{acde}	0.095^{acde}	0.058^{ce}	83%
UIR	0.004	0.002	0.002	0.002	100%
URM	0.005	0.003	0.018	0.054^{ce}	100%
RMUB	0.081^{acd}	0.064^{acd}	0.062^{acd}	0.050^{c}	60%
RMCE	$\mathbf{0.224}^{abcde}$	$\mathbf{0.204}^{abcde}$	$\mathbf{0.204}^{abcde}$	$\mathbf{0.138}^{abcde}$	100%

MovieLens 1M

Method	P@5	nDCG@5	nDCG@10	P@50	cvg
UB	0.035^{cd}	0.031^{cd}	0.031^{cd}	0.039^{cde}	100%
NC+P	0.037^{acd}	0.033^{acd}	0.036^{acd}	0.048^{acde}	99%
UIR	0.001	0.001	0.001	0.001	100%
URM	0.001	0.001	0.006	0.034^{c}	100%
RMUB	0.075^{abcd}	0.061^{abcd}	0.057^{abcd}	0.038^{c}	41.4%
RMCE	$\mathbf{0.187}^{abcde}$	$\mathbf{0.176}^{abcde}$	$\mathbf{0.168}^{abcde}$	$\mathbf{0.108}^{abcde}$	100%

Summary of comparative effectiveness

Conclusions

- *RMUB* outperforms other state-of-the-art approaches but is not optimal.
- The complete probabilistic model (*RMCE*) achieves even larger improvements.
- Improvement in performance are consistent across different datasets

We have also produced different mappings the involved variables [5], nevertheless, more research is still needed on this point.

References

[1] Bellogín, A., Castells, P., and Cantador, I. Precision-oriented evaluation of recommender systems: an algorithmic comparison. In *RecSys* (2011)

[2] Bellogín, A., and Parapar, J. Using graph partitioning techniques for neighbour selection in user-based collaborative filtering. In *RecSys* (2012)

[3] Lavrenko, V., and Croft, W. B. Relevance based language models. In *SIGIR* (2001)

[4] McLaughlin, M. R., and Herlocker, J. L. A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In *SIGIR* (2004)

[5] Parapar, J., Bellogín, A., Castells, P., and Barreiro, A. Relevance-based language modelling for recommender systems. *IPM. 49*, 4 (2013)

[6] Wang, J., de Vries, A., and Reinders, M. A user-item relevance model for log-based collaborative filtering. In *ECIR* (2006)

[7] Wang, J., de Vries, A. P., and Reinders, M. J. T. Unified relevance models for rating prediction in collaborative filtering. *ACM TOIS 26* (2008)

RecSys 2013, 7th ACM Conference on Recommender Systems. October 12–16, 2013. Hong Kong, China.