B. Mitschang et al. (Hrsg.): BTW 2017 – Workshopband,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2017 203

Towards Complex User Feedback and Presentation Context in Recommender Systems

Ladislav Peska[1], Peter Vojtas[1]

Abstract: In this paper, we present our work in progress towards employing complex user feedback and its context in recommender systems. Our work is generally focused on small or medium-sized e-commerce portals. Due to the nature of such enterprises, explicit feedback is unavailable, but implicit feedback can be collected in both large amount and rich variety. However, some perceived values of implicit feedback may depend on the context of the page or user's device (further denoted as *presentation context*). In this paper, we present an extended model of presentation context, propose methods integrating it into the set of implicit feedback features and evaluate these on the dataset of real e-commerce users. The evaluation corroborated the importance of leveraging presentation context in recommender systems.

Keywords: Recommender systems, implicit feedback, presentation context, user preference, e-commerce.

1 Introduction

Recommender systems belongs to the class of automated content-processing tools, aiming to provide users with unknown, surprising, yet relevant objects without the necessity of explicitly query for them. The core of recommender systems are machine learning algorithms applied on the matrix of user to object preferences. The user preference is usually derived from explicit user rating (also referred as *explicit feedback*), but this is not possible e.g. for small or medium-sized e-commerce enterprises, where user ratings are extremely scarce. Instead, one can focus on collecting specific features of user behavior (*implicit feedback*) and estimate user preference from it [Cl01], [Pe14], [Pe16], [Yi14].

Our approach is to focus on a complex model of the implicit feedback and learn user preference via some machine learning method. Our working hypothesis is that by collecting more informative description of user behavior, we shall be able to better estimate user's preferences and thus provide him/her with better recommendations. Similar approach are not so common. We can mention e.g. Yang et al. [Ya12], or Claypool et al. [Cl01]. However, in both cases the domain and thus the set of collected feedback significantly differs from our approach.

Further, the key part of our approach is to relate implicit feedback features to the relevant context. We consider several *presentation context* features (e.g. related to page

[1] Charles University in Prague, Faculty of Mathematics and Physics, peska@ksi.mff.cuni.cz

complexity or user's device capability). Closest to our work is the approach by Yi et al. [Yi14], proposing several presentation context features (e.g. content type or article length) to be used while estimating user preference from *dwell time*. We differ from this approaches in both the usage of extended set of context (e.g. features related to the browser visible area), the method of incorporating *presentation context* and considered implicit feedback features.

This work follows to our previous paper [Pe16]. We extend it by presenting improved method for context incorporation and improved recommending algorithms.

2 Materials and Methods

In traditional recommender systems, user u rates some small sample $\boldsymbol{S}$ of all objects $\boldsymbol{O}$, which is commonly referred as user preference $r_{u,o}: o \in \boldsymbol{S} \subset \boldsymbol{O}$. The task of traditional recommender systems is to build suitable user model, capable to predict ratings $\hat{r}_{u,o'}$ of all objects $o' \in \boldsymbol{O}$. In domains without explicit feedback, user preference $\bar{r}_{u,o}$ can be learned from the set of implicit feedback $L(f_1, \dots, f_i) \to \bar{r}_{u,o}$. Then, traditional recommending algorithms can be used. There are in principle three possible approaches to construct the preference learning function L:

- We can suppose that the higher value of each feedback feature implies the higher user preference. Based on this hypothesis, the estimated rating $\bar{r}_{u,o}$ can be defined as the average of all feedback values. However, as the feedback features distribution varies greatly, the feedback values must be normalized in order to be comparable. In this study, we used standardization of features (denoted as *STD* in evaluation), and their empirical cumulative distribution (denoted as *CDF*). The estimated rating $\bar{r}_{u,o}$ can be then defined as the mean of *STD* or *CDF* values of all feedback features for the respective user and object.

$$\bar{r}_{u,o} = \sum_{l=1}^{i} CDF(f_{l,u,o})\, / i$$

- Another option is to consider some feedback feature as a golden standard $f_{k,u,o} \approx \bar{r}_{u,o}$. The obvious candidate in e-commerce are *purchases*. However, as they are quite sparse[2], we can hypothesize that also other visited objects were preferred to some extent. One way to derive such preference is to employ supervised machine learning aiming to learn the probability that the object was purchased, based on the other feedback feature values (*J48* decision tree was used in this study).

- A baseline option is to use binary *visits* (i.e. suppose that users equally prefer all visited objects). Such approach is also quite common in the literature (e.g. [Os13], [Re09]).

[2] Less than 0.4% of the visited objects were purchased in our dataset.

The implicit feedback features used in this paper are *view count*, *dwell time*, *travelled mouse cursor distance*, *cursor in-motion time*, *scrolled distance*, *time of scrolling*, *clicks count* and *purchases*. More details can be found in [Pe14] and [Pe16].

However, the perceived values of implicit feedback features might be highly biased by the way, how the object is presented to the user. Suppose for example that the content of a webpage fully fits into the browser visible area. Then no scrolling is necessary and thus we receive zero values of *scrolled distance* and *time of scrolling*. Similarly, if the page contains mostly text (e.g. news or tour domains), then the time spend by reading the page is mostly determined by the length of the text itself.

There can be more such factors related to the features of the object itself, or its presentation, which, altogether, can be denoted as the *presentation context*. We propose several presentation context variables, which should be generally observable on the webpages. These are *volume of text*, *links* and *images*, *page* and *browser dimensions*, *page visible area ratio* and *hand-held device* indicator. Furthermore, we propose two approaches to incorporate *presentation context* into the process of user rating $\bar{r}_{u,o}$ estimation.

- Extend the dataset of implicit feedback features by the presentation context features (denoted as *FB+C* in the evaluation). This approach leaves the context incorporation on the preference learning method.
- Use presentation context as a *baseline* value *predictor* and subtract these from the perceived feedback values. We can either derive an average baseline predictor based on all contextual features (*AVGBP*), or create a separate baseline predictor of each presentation context feature and use the Cartesian product of implicit feedback features and baseline predictors in the preference learning step (*CBP*).
- We further evaluate two baseline approaches: usage of all feedback features disregarding of any context features (*FB*) and usage of binary feedback based on visited objects (*Binary*).

After the computation of $\bar{r}_{u,o}$ ratings, these are supplied to the recommender system, which computes the final list of recommended objects. In this study, the *Vector Space Model* (*VSM*) algorithm was adopted [LGS11], with binarized content-based attributes serving as a document vector and cosine similarity over TF-IDF weights as objects' similarity measure. We further enhance the *VSM* algorithm by multiplying its results by the general objects popularity (in terms of total $\bar{r}_{u,o}$) and thus prioritize more popular objects (*popVSM*).

To sum-up, the presented approach works in three steps. In the first step, the implicit feedback and contextual features are collected and combined into a set of feedback features $\{f_1, \dots, f_i\}$. Those features are used in the second step – user preference learning, where either machine learning algorithms or simple estimators are employed to derive estimated ratings $\bar{r}_{u,o}$. The estimated ratings are forwarded to the recommending algorithm, which proposes the final list of top-k recommended objects to the user.

3 Evaluation and Discussion

We evaluated the proposed approach on a Czech travel agency dataset [Pe16], aiming to predict objects purchased by the user. We adopted the leave-one-out cross-validation applied on purchases and considered the problem as ranking (i.e. the object purchased by the user should appear on top of the recommended objects). Due to the evaluation protocol, the dataset was restricted only to users with purchased objects and more than one visited object, resulting into 405 purchases from 253 users. Table 1 contains the results w.r.t. the normalized discounted cumulative gain (nDCG).

Table 1. Evaluation results in terms of average nDCG. Baseline methods are depicted in grey italics, the best result in bold. Results outperformed by the best result according to the binomial significance test [Sa97] w.r.t recall@top-10 are marked with * ($p < 0.05$) or ** ($p<0.01$).

Processing method	**Feedback and Context composition**				
	Binary	*FB*	**FB+C**	**AVGBP**	**CBP**
STD + popVSM	*0.255**	*0.174***	0.197**	0.161**	0.158**
CDF + popVSM	*0.255**	*0.257**	0.253*	0.258*	0.257
J48 + popVSM	*0.255**	*0.256**	**0.274**	0.240**	0.247**
J48 + objects popularity	*0.180***	*0.205***	0.211**	0.168**	0.186**
J48 + VSM	*0.222**	*0.224**	0.233	0.225*	0.224*

3.1 Discussion and Conclusions

In this work in progress report, we presented our approach towards employing complex user feedback and its context in recommender systems. Our approach works in three steps: feedback and context collection and combination, preference learning and recommendation. The evaluation shown capability of such model to improve over both binary feedback baseline and usage of complex feedback without contextual information. Adding contextual features as further feedback features before the preference learning step (*FB+C*) generally achieves the best results, combining it with *J48*-based preference learning and *popVSM* recommender provided the overall best result so far. Using J48 preference learning outperformed both *STD* and *CDF* heuristics. While the results of *STD*-based methods was clearly inferior, the *CDF* results were quite close to the *J48* ones, so we can recommend such approach in situations, where purchases or similar feedback features are not available. Using *popVSM* recommender significantly outperformed individual usage of both of its components (*VSM* and *object popularity*).

There is substantial amount future of work to be done in both algorithm design and its evaluation. In particular, we would like to focus more on feedback and context incorporation, e.g., proposing some context-based feature weighting. The results should be further validated by using additional preference learning methods, recommending algorithms and evaluation datasets. Also, there is a space for improvement of evaluation scenario itself, e.g., evaluate recommendations for new sessions. Successful candidates from the offline evaluation should be also validated via on-line A/B testing.

Acknowledgments. The work on this paper was supported by the Czech grant P46. Supplementary materials can be obtained from: http://bit.ly/2g79VVO.

References

[Cl01] Claypool, M.; Le, P.; Wased, M.; Brown, D.: Implicit interest indicators. In: IUI '01. ACM, pp. 33–40, 2001.

[LGS11] Lops, P.; de Gemmis, M.; Semeraro, G.: Content-based Recommender Systems: State of the Art and Trends. In: Recommender Systems Handbook. Springer, 2011, pp. 73–105.

[Os13] Ostuni, V. C.; Noia, T. D.; Sciascio, E. D.; Mirizzi, R.: Top-N recommendations from implicit feedback leveraging linked open data. In: RecSys 2013. ACM, pp. 85–92, 2013.

[Pe14] Peska, L.: IPIget – The Component for Collecting Implicit User Preference Indicators. In: ITAT 2014, Ustav informatiky AV CR. Pp. 22–26, 2014, url: http://itat.ics.upjs.sk/workshops.pdf.

[Pe16] Peska, L.: Using the Context of User Feedback in Recommender Systems. In: MEMICS 2016. (to appear), EPTSC, 2016.

[Re09] Rendle, S.; Freudenthaler, C.; Gantner, Z.; Schmidt-Thieme, L.: BPR: Bayesian Personalized Ranking from Implicit Feedback. In: UAI 2009. AUAI Press, pp. 452–461, 2009.

[Sa97] Salzberg, S. L.: On Comparing Classifi ers: Pitfalls to Avoid and a Recommended Approach. Data Mining and Knowledge Discovery 1/3, pp. 317–328, 1997.

[Ya12] Yang, B.; Lee, S.; Park, S.; Lee, S.-g.: Exploiting Various Implicit Feedback for Collaborative Filtering. In: WWW 2012. ACM, pp. 639–640, 2012.

[Yi14] Yi, X.; Hong, L.; Zhong, E.; Liu, N. N.; Rajan, S.: Beyond Clicks: Dwell Time for Personalization. In: RecSys 2014. ACM, pp. 113–120, 2014.