User-Item Reciprocity in Recommender Systems: Incentivizing the Crowd

Alan Said[1], Martha Larson[1], Domonkos Tikk[2], Paolo Cremonesi[3], Alexandros Karatzoglou[4], Frank Hopfgartner[5], Roberto Turrin[6], and Joost Geurts[7]

[1] Delft University of Technology - alansaid@acm.org, m.a.larson@tudelft.nl
[2] Gravity R&D - domonkos.tikk@gravityrd.com
[3] Politecnico di Milano - paolo.cremonesi@polimi.it
[4] Telefonica Research - alexk@tid.es
[5] Technische Universität Berlin - frank.hopfgartner@dai-labor.de
[6] Moviri - roberto.turrin@moviri.com
[7] JCP-Connect - joost.geurts@jcp-connect.com

Abstract. Data consumption has changed significantly in the last 10 years. The digital revolution and the Internet has brought an abundance of information to users. Recommender systems are a popular means of finding content that is both relevant and personalized. However, today's users require better recommender systems, able of producing continuous data feeds keeping up with their instantaneous and mobile needs. The CrowdRec project addresses this demand by providing *context-aware*, *resource-combining*, *socially-informed*, *interactive* and *scalable* recommendations. The key insight of CrowdRec is that, in order to achieve the dense, high-quality, timely information required for such systems, it is necessary to move from passive user data collection, to more active techniques fostering user engagement. For this purpose, CrowdRec activates the crowd, soliciting input and feedback from the wider community

1 Introduction

The new generation of recommender systems find recommendations for their users in particular situations and at certain moments in time. As a result, the amount of information needed on an item increases dramatically since such systems deal not in single items, but rather in pairs, (i.e., recommend a *book* for a user to read at a specific *point in time*). The result can potentially exacerbate the data sparsity problem. To overcome this problem, rich and reliable sources of information on the items available for recommendation are necessary. This information needs to include not only views and ratings, but also contextual information on the user's situation, the device in use, etc. In particular, recommender systems focusing on user experience can exploit users' comments and reviews, the context of users and items, and other interaction data [5]. Conventional recommender systems are passive, i.e. they wait for users to start interacting with the system. However, if a recommender system could actively

Fig. 1. A recommender system actively interacting with the crowd. The arrows show reciprocal relationships between *a user interested in an item* and analogously *an item in need of user attention*. This type of relationships motivate user's to interact with items and actively contribute data that can be used to improve recommendation quality.

"request" more information for certain items, or certain item-context pairs (e.g. movie to watch during a flight), the sparse data problem could be addressed directly, and recommendation quality substantially improved. This is what the CrowdRec[8] project is focusing on.

Outsourcing micro-tasks to many users, i.e. crowdsourcing, is an unmined information resource; by actively collecting information from users on items, a richer and denser dataset is collected and can be used to improve recommendation quality. Crowdsourcing does however come with a drawback, i.e. users need to be remunerated for their contributions. We propose that recommender systems utilize their own users to contribute data. Instead of the traditional financial incentives common in crowdsourcing, we intend to motivate user contributions by matching users with the specific items that they find interesting enough to comment on, review or interact with at specific points in time. The result is a recommender system with *crowd activation*, as illustrated in Fig. 1.

The key factor in ensuring high quality of data is *user-item reciprocity*, i.e. if the recommended item is of interest to the user but does not invite interaction (e.g. tagging, reviewing, etc.), little or no data will be added by the user. However, if the user is likely to interact with the item (item-user reciprocity), more data will be generated. Consider this example: a new restaurant opens in a location that is off the beaten track. In the regular case, it would take considerable time for the restaurant to be discovered by the crowd. The system recommends this restaurant to a person familiar with that area, this person is also likely to write a review. Additionally, the person often frequents restaurants and is motivated by the fact that this particular review can make an important contribution to popularize the neighborhood she is living in. The result is a richer description of the restaurant in questions which results in better recommendations for everyone.

Analogously, in the case of user-generated video portals, conventional recommender systems tend to down-weight videos that receive little attention in the first few days after uploading [2]. With the proposed methodology niche content may also find audiences easier, thus potentially contributing to the diversity of recommendations.

[8] http://crowdrec.eu

2 Incentivizing Active Participation from the Crowd

Incentivization in the context of crowdsourcing is the act of motivating contributions from the crowd, and at the same time caring about the quality of these contributions. According to Antin and Shaw [1] and Kaufmann et al. [4], crowd member incentives can be divided into two basic categories; 1) *intrinsic motivation*, motivation arising from internal factors, e.g. enjoyment, identification with a community and need for social contact, and 2) *extrinsic motivation*, arising from external factors such as awards and external obligation. The CrowdRec vision of crowd activation for recommendation exploits both motivation types, with a specific focus on intrinsic motivation. This is attained by pairing users with items they likely are interested in, and in parallel targeting items which will benefit from user attention. The result being satisfied users engaging and contributing for the good of the community. The power of incentives in crowdsourcing has been demonstrated by real work applications, e.g. Podcastle (podcast transcripts) and Songle[9] (music annotation). Their creators report that the quality of contributed information exceeds that of commercial crowdsourcing platforms [3]. Our vision extends these approaches by pairing users with content, and vice versa. We anticipate reciprocal recommendations to actively match users to items not only based on interest, but also on need and likelihood of interaction.

3 User-Item Reciprocity

The vision of crowd activated recommender systems presented here aims to create a symbiosis between users and items. To realize this symbiosis, two factors must be taken into account: 1) users who are recommended to items should be interested in those items and have the potential to enrich them, and 2) it is necessary to create mechanisms enticing users to provide data on recommended items. The first consideration could be perceived as a traditional recommendation scope, the second will however require extending existing recommender system techniques in order to reach the goal. One direction for this extension is to build on existing work in reciprocal recommendation, creating matches between both the target item, and the target user. This technique, commonly used in dating recommendation, generates pairs of users with mutual preference [6]. In order to establish this symbiotic relationship, the concept of reciprocity in recommendation must be extended to scenarios where various constraints, e.g. duration of availability, novelty, interestingness, intrinsically limit the potential users an item can be recommended to. The ways in which reciprocal preference modeling can improve recommendation performance must be thoroughly analyzed before it can be understood. This includes the relation between the recommendations displayed and the response rate, the collection of critical amounts of feedback to better characterize media content, cutting the duration of cold-start for new

[9] http://en.podcastle.jp/ and http://songle.jp/

users and items, and optimization not only for users but also technical factors and business objectives.

4 Conclusions and Outlook

We presented a vision combining crowdsourcing and recommender systems. Our insight is that recommender systems can utilize their own user base as a crowd that can contribute the rich information needed to address the sparse data problem faced by recommender system. By using reciprocal recommendation to identify items that are suited to users, and additionally users that are suited to items, we propose that it is possible to incentivize users to contribute information on items. The resulting symbiotic user-item relationship will generate richer, high-quality information, resulting in better recommendations. We emphasized real-time and context-aware recommendation as contributing to the sparse data problem. Other factors could be important as well, e.g. in video recommendation it is interesting to recommend not only whole videos, but also time-points within specific videos. It is possible that rich information at the time-point level is only possible if recruited users are also interested in specific videos, making them interested in interacting with and tagging videos in their entire length. By activating the crowd, we can move beyond the problem of data sparsity to the problem of addressing low quality data. In addition to contributing, the crowd can also validate information that is used as a basis for recommendations.

Acknowledgement

This work received funding from the European Union's Seventh Frame work Programme (FP7/2007-2013) under CrowdRec Grant Agreement no. 610594.

References

1. Judd Antin and Aaron Shaw. Social desirability bias and self-reports of motivation: a study of amazon mechanical turk in the us and india. In *CHI*. ACM, 2012.
2. Flavio Figueiredo, Jussara M. Almeida, Marcos Andre Goncçales, and Fabrício Benevenuto. On the dynamics of social media popularity: A YouTube case study. *arXiv*, 1402.1777v1, 2014.
3. Masataka Goto, Jun Ogata, Kazuyoshi Yoshii, Hiromasa Fujihara, Matthias Mauch, and Tomoyasu Nakano. Podcastle and songle: Crowdsourcing-based web services for retrieval and browsing of speech and music content. In *CrowdSearch*, 2012.
4. Nicolas Kaufmann, Thimo Schulze, and Daniel Veit. More than fun and money. worker motivation in crowdsourcing-a study on mechanical turk. In *AMCIS*, volume 11, pages 1–11, 2011.
5. Joseph A Konstan and John Riedl. Recommender systems: from algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(1-2):101–123, 2012.
6. Luiz Pizzato, Tomek Rej, Thomas Chung, Irena Koprinska, and Judy Kay. Recon: a reciprocal recommender for online dating. In *RecSys*, pages 207–214. ACM, 2010.