How Diverse Is Your Audience? Exploring Consumer Diversity in Recommender Systems

Jacek Wasilewski
Insight Centre for Data Analytics
University College Dublin
Dublin, Ireland
jacek.wasilewski@insight-centre.org

Neil Hurley
Insight Centre for Data Analytics
University College Dublin
Dublin, Ireland
neil.hurley@insight-centre.org

ABSTRACT

On-line recommender systems have different challenges to overcome to provide content to users. One of these is the potential of isolating users from a diverse set of items by recommending very narrow content. In this paper we propose an item-centric view of a recommender system, looking at the exposure of items to groups of consumers, and how diverse those groups are, to identify if items are recommended to narrower groups of consumers. This is opposite to current practice where diversity of content is typically analysed. Preliminary results on the MovieLens 20M dataset show that recommender systems expose items to narrower groups of consumers, and these groups are less diverse.

KEYWORDS

recommender systems; diversity; consumer diversity; item-centric evaluation

1 INTRODUCTION

Recommender systems have become ubiquitous in the interfaces to product catalogues provided by on-line retailers. From the user's perspective, recommender algorithms are used to filter a large set of possible selections into a much smaller set of items that the user is likely to be interested in. On the other hand, from the business point of view, as important as users getting engaging recommendations is the utilisation of products in the catalogue.

Sales increase or redistribution across the whole catalogue of items might not be the only business goal to be addressed by a recommender system. In some sense, recommender systems are marketing tools that identify customers and target these customers with personalised items. Questions arise: are we exposing items to users that showed an interest before? Are we promoting items to reach new groups of customers? How diverse are these groups? From market development perspective, recommender system should help us in achieving all of these business goals. To measure and control for this, we need a picture of how items are exposed to different groups of people, and if the exposure is diverse.

In this paper we tackle the problem of *item exposure* to understand who consumes items and if potential consumers are reached by recommendations. We measure diversity of the people getting recommendations for an item, using approaches coming from ecology, such as *species diversity* of a habitat. This is different to the *content diversity* of recommendations that has been typically considered in the context of diversity in recommender systems. The main goal of this paper is to find the answer to the following question: do recommender systems expose items to the same, wider or narrower groups of consumers, and how diverse are these groups.

2 CONSUMER DIVERSITY

Recommender systems have to deal with the long tail of items that are rarely recommended. This includes niche items that are rarely liked, but also items that have not penetrated the market. To identify and promote these items, we argue it is not enough to ask *how many* users have rated each item in the past, but also *which users* have rated the items, which define its *item exposure*.

An item's user profile, $\mathcal{U}_i$, contains the set of users who rated the item in the past. A diversity measure over these users gives insight into the extent to which item has been exposed to a wide range of different user types. Similarly, the set of users to whom the item is recommended, $\mathcal{R}_i$, can be analysed to reveal the extent to which recommendations extend the exposure of an item. If an item is recommended to diverse consumers, it is possible that the item can reach a wider potential market.

As it is commonplace for marketeers to model their customer-base through customer segmentation, we find it useful to measure the diversity in terms of the spread across different consumer segments. Given a partition $\mathcal{P}_c$ of $\mathcal{U}$ into k consumer segments, $\mathcal{U} = C_1 \cup C_2 \cup \ldots \cup C_k$, where C_j is the j^{th} consumer segment, we define *consumer diversity* of a set of consumers S, as functions of $(p_1, \ldots, p_k)$, where $p_j = \frac{|S \cap C_j|}{|S|}$ is the proportion of the set S that belong to consumer segment C_j.

A similar problem is considered in ecology, where a habitat can be quantified in terms of *species diversity* [5, 6], which measures diversity in terms of the *proportionality abundance* of each species in a sample. It assigns a high diversity value when the sample is evenly spread across the different species. In biodiversity, different measures like species richness, Shannon entropy, Simpson concentration, can be generalised through the *Hill number* [5], or *diversity of order q* defined as:

$$^qD \triangleq \left(\sum_{j=1}^{k} p_j^q \right)^{1/(1-q)}$$

and $^1D = \lim_{q \to 1} {}^qD = \exp(H(p))$. In biodiversity these are called *true diversities* or *effective number of species* [6]. With $q = 0$ we obtain richness, $q = 1$ true diversity of Shannon entropy, and for $q = 2$ inverse Simpson index. Entropy increases as both richness and evenness increase, where Simpson index measures dominance and is less sensitive to richness. In our context, each consumer segment corresponds to a "species". Then, with the help of the true diversity we can evaluate the diversity of a habitat—that is, an item in our case. We can use true diversity to compare the exposure

Project funded by Science Foundation Ireland under Grant No. SFI/12/RC/2289.
RecSys 2017 Poster Proceedings, August 27-31, Como, Italy.

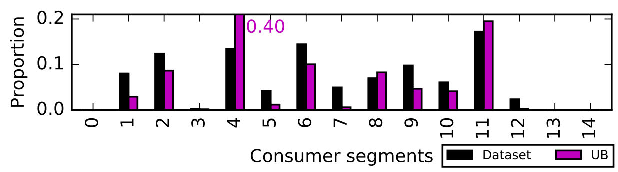

Figure 1: Distribution over segments of *The Matrix* movie (pop: 51,334), in the dataset and recommendation.

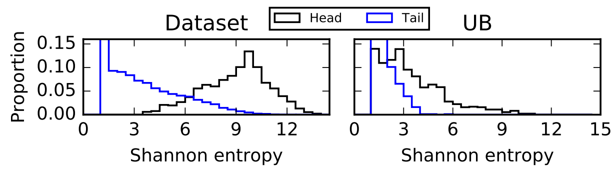

Figure 2: Histograms of Shannon entropy for dataset and recommendations. Items are distinguished by their popularity: head and tail items.

	Dataset	UB	IB	MF
Richness ($q = 0$)	12.99	5.11	3.73	6.43
Shannon ($q = 1$)	8.41	2.84	2.99	3.53
Simpson ($q = 2$)	6.90	2.34	1.95	2.84

Table 1: Average values of true diversity indices for all items in the dataset and recommendations (UB, IB, MF).

of different items to the consumer segments or to compare the exposure of a single item under different conditions.

3 ANALYSIS OF CONSUMER DIVERSITY

We investigate consumer diversity on the MovieLens 20M dataset [4]. For that, a partition into consumer segments is required. We create behavioural segments based on past interactions. X-means [7] clustering algorithm is used to define segments—$k = 15$ clusters have been created based on interactions. Results of such clustering depends on the initialisation parameters, which is a limitation, but it still enables comparison of diversity. We analyse recommendations (of $N = 20$ items) generated by collaborative filtering algorithms available in the RankSys framework (http://ranksys.org): user- (UB) and item-based (IB) kNN, and matrix factorisation (MF).

We wonder if recommender systems might suffer not only from narrowing content served to users, but also items being exposed to narrow audiences. To illustrate that, we take a movie (*The Matrix*) for which we show distribution of consumers over segments—Figure 1. It can be seen that one segment (no. 4) is over-represented almost 4 times in recommendations. We measured its true diversities: richness, Shannon and Simpson indices. Richness decreased from 15 to 13 which means 2 segments are not reached, Shannon and Simpson indices also dropped, respectively, from to 9.80 to 6.14, and 8.85 to 4.43, which means that recommendations are generally less diverse in terms of consumers to which this movie has reached. True diversities are also easy to interpret—they tell the effective number of species, the number of equally abundant species that produce same diversity. In our case, recommendations are 1.5 times less diverse on Shannon index, and 2 times on Simpson index.

Table 1 contains values of considered true diversity indices, averaged over all items. Richness shows that on average items are consumed by users of 13 out of 15 segments, but only recommended to 3-6 segments. If concentration is taken into account, Shannon and Simpson indices drop, indicating items being 2-3 times less diverse. Paired t-test show significance of the differences ($p < 0.001$).

As item's popularity can affect collaborative filtering methods, we wonder if lower diversity is due to low item popularity. To examine this, we split the items into the head of most popular items (80% of interactions), and the rest in the tail—histogram of Shannon diversity in Figure 2. On dataset, head items have higher diversity, while tail tends to obtain lower values. Recommendations do not follow these—both groups of items have distribution of diversity skewed towards 0. This suggests that even popular items, receiving more interactions, are isolated from wide and diverse consumers.

4 RELATED WORK

Diversity is commonly studied in the context of items that are recommended to users, which might help mitigating the problem of users being exposed to narrower spectrum of item types. A number of frameworks have been proposed to measure and increase diversity, such as *Intra-List Diversity* [10].

Sales diversity [1, 3] is a notion of diversity which attempts to capture how items perform, e.g. how evenly they are consumed. It tackles the *long tail* problem, where most popular items drive the recommendations. *Aggregate Diversity* [1], *Gini index* [3], *Shannon entropy* [9] are some of the measures of sales performance over items. In [2] an item-centric evaluation is conducted to detect pathologies hindering novel recommendations. These method, however, analyse impacts on items globally, not individually, and also without considering different groups of consumers.

In information retrieval, a concept of profile diversity [8] has been proposed, where a profile contains information about the user's community. Then queries should retrieve documents that different communities find useful. However, the framework does not analyse consumers reached by these documents.

5 CONCLUSIONS

In this paper we identified and explored the problem of consumer diversity, which measures how diverse each item is in terms of consumer segments. Our analysis shows that popular recommendation techniques expose items to much narrower and less diverse consumers. Although the overall quality of recommendations might be good, items are hidden from certain groups of people who expressed an interest in them in the past.

REFERENCES

[1] G. Adomavicius and Y. Kwon. 2012. Improving Aggregate Recommendation Diversity Using Ranking-Based Techniques. *IEEE TKDE* 24, 5 (2012).
[2] Ò. Celma and P. Herrera. 2008. A New Approach to Evaluating Novel Recommendations *(RecSys '08)*.
[3] D. Fleder and K. Hosanagar. 2009. Blockbuster Culture's Next Rise or Fall: The Impact of Recommender Systems on Sales Diversity. *Manage. Sci.* 55, 5 (2009).
[4] F. M. Harper and J. A. Konstan. 2015. The MovieLens Datasets: History and Context. *ACM Trans. Interact. Intell. Syst.* 5, 4 (2015).
[5] M. O. Hill. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology* 54, 2 (1973).
[6] L. Jost. 2006. Entropy and diversity. *Oikos* 113, 2 (2006).
[7] D. Pelleg and A. W. Moore. 2000. X-means: Extending K-means with Efficient Estimation of the Number of Clusters *(ICML '00)*.
[8] M. Servajean, E. Pacitti, S. Amer-Yahia, and P. Neveu. 2013. Profile Diversity in Search and Recommendation *(WWW '13 Companion)*.
[9] Z. Szlavik, W.J. Kowalczyk, and M.C. Schut. 2011. Diversity measurement of recommender systems under different user choice models *(ICWSM'11)*.
[10] C. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen. 2005. Improving Recommendation Lists Through Topic Diversification *(WWW '05)*.