Beyond Modeling Private Actions: Predicting Social Shares

Si Si [*]
University of Texas-Austin
ssi@cs.utexas.edu

Atish Das Sarma
eBay Research Labs
atish.dassarma@gmail.com

Elizabeth F. Churchill
eBay Research Labs
churchill@acm.org

Neel Sundaresan
eBay Research Labs
nsundaresan@ebay.com

ABSTRACT

We study the problem of predicting sharing behavior from e-commerce sites to friends on social networks via share widgets. The contextual variation in an action that is private (like rating a movie on Netflix), to one shared with friends online (like sharing an item on Facebook), to one that is completely public (like commenting on a Youtube video) introduces behavioral differences that pose interesting challenges. In this paper, we show that users' interests manifest in actions that spill across different types of channels such as sharing, browsing, and purchasing. This motivates leveraging all such signals available from the e-commerce platform. We show that carefully incorporating signals from these interactions significantly improves share prediction accuracy.

Categories and Subject Descriptors

J.4 [**Computer Applications**]: Social and Behavioral Sciences

Keywords

Recommender systems; e-commerce; user behavior analysis

1. INTRODUCTION

Social media platforms provide an outlet for expressing personal preferences or generic opinions with friends through features like "share'. Increasingly, these "social signals" include events, products, and services. Shares pertaining to commerce are particularly relevant from a monetization standpoint. Share widgets are being rolled out on platforms like Amazon and eBay in the hope that they increase sales through network marketing. There however remains little understanding on how users actually engage with these share widgets on such marketplaces. There has been a vast body of work towards predicting private preferences (e.g. ratings). However, predicting users' preferences for sharing items from e-commerce sites to external social networks remains challenging even at the broad category level (i.e. likelihood of user sharing an item in fashion vs. sports memorabilia). Designing good share prediction algorithms, have practical implications in prioritizing content for user feeds, marketing strategies, and general recommendations.

Recommendation algorithms do not always distinguish between actions that often elicit vastly different behaviors: for e.g. (a) actions that are private (like rating a movie on Netflix), (b) those that are shared with online friends (like sharing an item on Facebook), (c) actions that are completely public (like commenting on a Youtube video). The focus of this paper is on predicting the second type of action, although originating from an e-commerce platform. We demonstrate that combining interests through multiple distinct, yet related, types of actions: for e.g. browsing through item pages, sharing through widgets, or engaging in purchases. yields improved prediction algorithms.

Dataset. In this paper, we focus on the objective of predicting users' sharing behavior across the categories of items on data from eBay. The dataset is collected from a large sample of users and actions between 2012 and 2013. We consider 950,000 users, with around 2.5 million shares, 20 million transactions, and over 66 million page-views. Further, each item involved falls in a category tree structure, with 39 top-level categories (such as fashion, home and garden, books) and one of around 15000 leaf-categories. To illustrate the rough statistics and variation across shares, pageviews, and purchases, we show distributional plots in Figure 1.

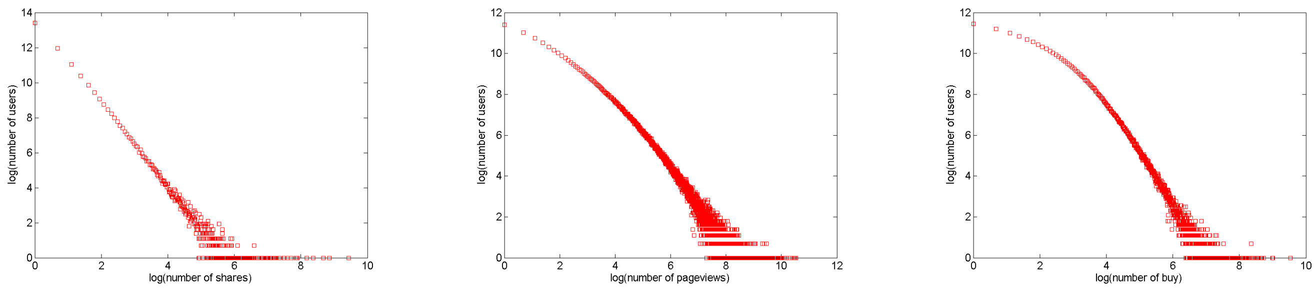
2. OUR PROPOSED APPROACH

The problem is formulated as follows: given a share matrix $S \in R^{m\times n}$ with missing entries, where each S_{ij} gives the number of times the user i shares items in category j and m and n being the number of users and categories respectively, the task is to predict all missing entries based on the existing known entries in S. We are also given two other matrices P and B showing users' pageview and purchase histories over categories. So P_{ij} and B_{ij} give the number of times user i has purchased or clicked on items in category j respectively.

To predict a user's share interest in some categories, one common way is to consider recommendation system based approaches. For recommendation system problems, it has been widely shown that matrix factorization based approaches perform better than traditional neighborhood based approaches [4]. For matrix factorization, we want to find two matrices $W_S \in R^{m\times k}$ and $H_S \in R^{n\times k}$, such that for the known entry, $\boldsymbol{w}_i^s \boldsymbol{h}_j^{sT}$ is close to S_{ij}, where $\boldsymbol{w}_i^s$ and $\boldsymbol{h}_j^s$ are the i-th row for W_S and j-th row for H_S respectively.

[*]Work done while interning at eBay Research Labs.

WWW'14 Companion, April 7–11, 2014, Seoul, Korea.
ACM 978-1-4503-2745-9/14/04.
http://dx.doi.org/10.1145/2567948.2577278.



(a) share (b) pageview (c) purchase

Figure 1: User count distributions for number of shares, pageviews and purchases (in log-log scale).

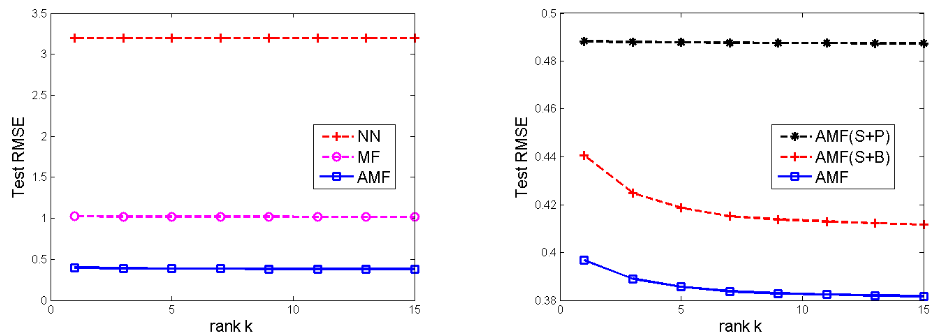


Figure 2: Varying rank k for AMF for share prediction. We see that AMF always performs better than MF and NN. Further, AMF is better than augmenting the share matrix with just one of purchases(AMF(S+B)) or pageviews(AMF(S+P)).

We want to solve:

$$\min_{\substack{W_S \in \mathbb{R}^{m\times k} \\ H_S \in \mathbb{R}^{n\times k}}} \sum_{(i,j)\in\Omega_S} (S_{ij} - \boldsymbol{w}_i^s \boldsymbol{h}_j^{sT})^2 + \lambda \left(\|W_S\|_F^2 + \|H_S\|_F^2 \right), \quad (1)$$

where Ω_S is the set of observed shares in S; λ is the regularization parameter for both W_S and H_S. The share matrix S is approximated by $W_S H_S^T$ and the number of shares the user i makes for j-th category is predicted as $\boldsymbol{w}_i^s \boldsymbol{h}_j^{sT}$. Similarly, we can solve the Problem (1) to approximate P by $W_P H_P^T$ and B by $W_B H_B^T$. Based on that, we can predict P_{ij} and B_{ij} accordingly. A simple way to use this information is to combine the interactions together as follows, named Augmented Matrix Factorization(AMF):

$$\bar{S}_{ij} = \alpha_s \boldsymbol{w}_i^s \boldsymbol{h}_j^{sT} + \alpha_p \boldsymbol{w}_i^p \boldsymbol{h}_j^{pT} + \alpha_b \boldsymbol{w}_i^b \boldsymbol{h}_j^{bT}, \quad (2)$$

where $\boldsymbol{w}_i^p$ and $\boldsymbol{h}_j^p$ are the i-th row for W_P and j-th row for H_P respectively; $\boldsymbol{w}_i^b$ and $\boldsymbol{h}_j^b$ are the i-th row for W_B and j-th row for H_B respectively; and α_s, α_p, α_b are weights learned from the share matrix S. These three scores show the likelihood of user i sharing, purchasing, and clicking on the items in j-th category. So the combination of these three shows combined interaction of user i.

In Figure 2, we show the benefit of using auxiliary information from pageviews and purchases to improve the share prediction. In this figure, we compare AMF with performing matrix factorization on S alone(MF) and the neighborhood based approach(NN), which predicts the number of shares based on the neighbors of the user. We can clearly see that both MF and AMF perform better than NN, showing that the matrix factorization based methods performs better than neighborhood-based approaches for share prediction. Furthermore, using just purchases (in AMF(S+B)) or just pageviews (in AMF(S+P)) is not sufficient to derive the full predictive accuracy. The figure also shows that AMF is robust to varying k.

3. RELATED WORK

There are studies that investigate effects of integrating social networks into e-commerce websites[2, 7]. The goal of providing users with the flexibility to share their interests is to increase consumption based on the belief that social influence is a strong promoter for the consumption. Studies [7, 1, 3] show that while social links are sparse, transactions between friends' friends in the network usually correlate with higher user satisfaction, and take this to indicate the positive impact of social media. Integration of recommender systems in e-commerce marketplaces to foster and leverage social interaction has shown some success [6, 5].

4. REFERENCES

[1] S. Guo, M. Wang, and J. Leskovec. The role of social networks in online shopping: information passing, price of trust, and consumer choice. In *EC*, pages 157–166, 2011.

[2] B. J. Jansen, K. Sobel, and G. Cook. Being networked and being engaged: the impact of social networking on ecommerce information behavior. In *iConference*, 2011.

[3] S.-W. Kim, C. Faloutsos, and J. Ha. Blogcast effect on information diffusion in a blogosphere. In *SIGIR*, pages 1149–1150, 2011.

[4] Y. Koren. Collaborative filtering with temporal dynamics. In *KDD*, pages 447–456, 2009.

[5] C. Lam. Snack: incorporating social network information in automated collaborative filtering. In *EC*, 2004.

[6] J. Leskovec, L. A. Adamic, and B. A. Huberman. The dynamics of viral marketing. *ACM Trans. Web*, 2007.

[7] G. Swamynathan, C. Wilson, B. Boe, K. Almeroth, and B. Y. Zhao. Do social networks improve e-commerce? a study on social marketplaces. In *WOSN*, pages 1–6, 2008.