Affinity-Driven Prediction and Ranking of Products in Online Product Review Sites

Hui Li
herolee@pmail.ntu.edu.sg

Sourav S Bhowmick
assourav@ntu.edu.sg

Aixin Sun
axsun@ntu.edu.sg

School of Computer Engineering
Nanyang Technological University
Singapore

ABSTRACT

Large online product review websites (*e.g., Epinions, Blippr*) have recently come into being containing information related to various types of products. Typically, each product in these sites is associated with a group of members who have provided ratings and comments on it. These people form a *product community*. A potential member can join a product community by giving a new rating to the product. We refer to this phenomenon of a product community's ability to "attract" new members as *product affinity*.

The knowledge of a ranked list of products based on product affinity is of much importance to be utilized for implementing policies, marketing research, online advertisement, and other applications. In this paper, we identify and analyze an array of features that exert effect on product affinity and propose a novel model, called *AffRank*, that utilizes these features to predict the future rank of products according to their affinities. Evaluated on a real-world dataset, we demonstrate the effectiveness and superior prediction quality of *AffRank* compared to baseline methods. Our experiments show that features such as *affinity rank history, affinity evolution distance*, and *average rating* are the most important factors affecting future rank of products.

Categories and Subject Descriptors

H.3.3 [**Information Search and Retrieval**]: Information Search and Retrieval—*Information Filtering*; J.4 [**Computer Applications**]: Social and Behavior Sciences

General Terms

Algorithms, Experimentation

Keywords

Social networks, community evolution, product community, product review, product ranks, ratings
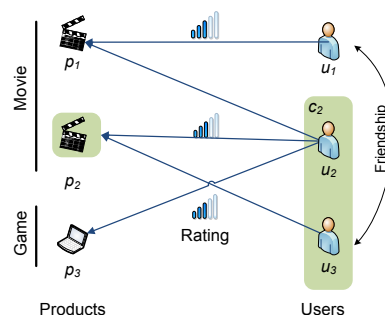
Figure 1: Product review website structure.

1. INTRODUCTION

Large online product review websites (*Epinions*[1], *Blippr*[2]) have recently come into being containing information related to many categories of products. Within these websites, individual users are allowed to publish their comments or give ratings on different products. Besides, they can set up friendships by linking to each other. Figure 1 depicts the structure of such a product review site. Observe that for a particular product (*i.e.,* p_2), there is a group of people who have given ratings and published their comments on it (*i.e.,* u_2, u_3). These people form a community (*i.e.,* c_2). In other words, each product in a product review site is associated with a community [6]. In the sequel, we refer to such a community as *product community*. Clearly, these communities are a potential gold mine for all kinds of marketing and business analysts as users' comments and ratings toward a particular product may affect other consumers' purchasing behavior [6].

A new user (member) can join a product community by giving a new rating to the product. Hence, *the growth of a product community's size can be implicitly measured by the number of new ratings*[3] *it receives during a particular time slot.* We refer to this phenomenon of a product community to "attract" new users as *product affinity*. Specifically, it is measured by the number of *new* ratings a product receives from *new* members during a particular time period. In fact, existing product review websites often display a ranked list of products that received the most number of *new* ratings on a daily, weekly, or monthly basis. We refer to this ranked product list at a particular time slot as *affinity rank*. For

---

[1] http://www.epinions.com

[2] http://www.blippr.com

[3] A rating that is updated by an existing member is not considered as new.

Table 1: Statistics of *Blipper* dataset.

Dataset	Number of products	Number of categories	Number of ratings	Number of users	Number of user-user links
Blippr	75	5	8,032	2,219	10,480

example, consider Figure 2. It depicts two affinity ranks of top-5 movies extracted from the *Blippr* website during weeks $t-2$ and $t-1$, respectively. Observe that *Ninja Assassin* and *Sherlock Holmes* received the most number of new ratings.

The affinity ranks of products in the past weeks/months highlight the popularity of products among new users in the (recent) past. Although such historical information is important for several applications, prediction of future affinity ranks of products is even more important to marketing and business strategists. For example, reconsider Figure 2. Suppose in week $t-1$ a company intends to put advertisements on five movie products during week t. Then, it makes sense to predict 5 most popular products at time t so that optimum benefit can be achieved. That is, it is desirable to predict the affinity ranks of products in the near future. In this context, instead of just listing most popular products, ranking them based on their affinity ranks makes more sense as a company may allocate different shares of their advertisement budget depending on the popularity of the products. Note that such top-k products may vary considerably at two different time points. For instance, consider the top-5 products during weeks $t-2$ and $t-1$ in Figure 2. Observe that *Sherlock Holmes* moved from rank 5 to the top rank in successive weeks. Further, *Invictus* first appeared in the top-5 list in week $t-1$ whereas *Crazy Heart* failed to remain in the top-5 list in this week.

Recent research on ranking products in product review websites have primarily focused on evaluating a product by the strength of connections among its users [6] or by the features related to a particular product item (*i.e.,* price, released time etc.) [3]. However, these techniques are not designed to predict the ranks of products based on their affinities. In this paper, we propose a novel quantitative model called *AffRank* that utilizes historical and evolutionary affinity information as well as other features to predict the future ranks of products according to their affinities. We propose three features, namely *affinity rank history*, *average ratings*, and *affinity evolution distance*, related to the product communities that may exert significant effect on affinity. In particular, *affinity rank history* represents the historical affinity ranks of products over time; *average ratings* measures the average of ratings received by a product at a given time point; and *affinity evolution distance* measures the distance between affinity evolution of a pair of different products. *To the best of our knowledge, these features have not been studied systematically in the literature.*

2. FEATURES FOR AFFINITY PREDICTION

In this section we describe the features that are used in our ranking model. All the values for these features are normalized into the interval [0,1] using Min-Max Normalization. We begin by introducing the real-world dataset we have used for our study. Table 1 describes the *Blippr* dataset was crawled using *Blippr* API[4] till August, 2009. It includes user ratings toward 75 different products.

Figure 2: Product affinity and affinity ranks.
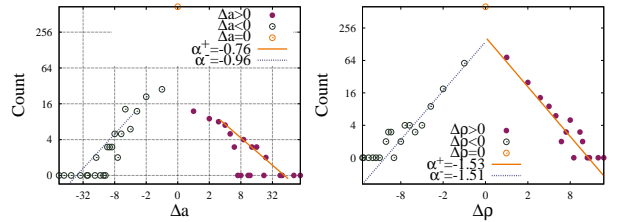

(a) Δa (b) $\Delta \rho$.

Figure 3: Distributions of Δa and $\Delta \rho$.

Affinity Rank History. Recall that each product p_i is associated with an affinity rank at time t, denoted by $\rho_t(p_i)$. Since the affinity rank of a product depends on the number of new users (product affinity), here we characterize the evolutionary behaviors of product affinity and affinity rank. We first investigate how the affinity changes between consecutive time points in the history. We denote the number of new users for product p_i at time slot t as $a_i^t = |\mathbf{C}_i^t| - |\mathbf{C}_i^{t-1}|$. Figure 3(a) reports the distribution of $\Delta a_i^t = a_i^t - a_i^{t-1}$ over all products p_i and time t. It indicates that the product affinity for all the products discussed in this paper is more likely to *increase spikily* and *drop down smoothly*. As affinity is the number of new users associated to a product within a time interval, this phenomenon reflects the speed of growth of product community size. It suggests that the speed of growth tends to increase to a peak in a short time and diminishes slowly subsequently. This phenomenon is generally applicable for most products, although we do acknowledge that in some specific categories (e.g., car batteries) this may not be true.

We now investigate the change of affinity rank between consecutive time slots in the history for each product. In particular, the change of affinity rank for a product between times $t-1$ and t is measured as follows.

$$\Delta \rho_t(p_i) \quad = \quad \rho_t(p_i) - \rho_{t-1}(p_i). \tag{1}$$

We calculate the count for each $\Delta \rho_t(p_i)$ value over all products and time slots. The distribution of $\Delta \rho$ is reported in Figure 3(b). Clearly, it follows a long-tail distribution. If we fit the curve in Figure 3(b) using power law models at both the left and right sides of $\Delta \rho = 0$ separately, the exponents equal to -1.51 for the negative changes where $\rho_{t-1}(p_i) > \rho_t(p_i)$ and -1.53 for the positive ones. The exponents are so close that the distributions of positive $\Delta \rho$ and negative $\Delta \rho$ are almost symmetrical. Besides, such a
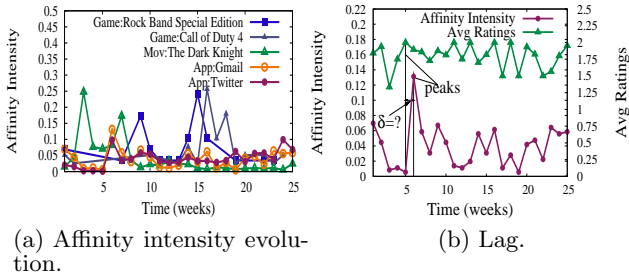
(a) Affinity intensity evolution.

(b) Lag.

Figure 4: Affinity intensity evolution and Lag between the average rating evolution and AIC.

phenomenon also indicates that the change of rank is most probable to be within a small range. Note that the aforementioned phenomenon also exists for each specific category of products. Thus, the rank of a product at any time slot is highly related to its previous rank. In the next section, we shall exploit the symmetric property of $\Delta\rho$ instead of asymmetric Δa in our *AffRank* model.

Affinity Evolution Distance. We now analyze and compare the evolutionary nature of different product affinities. We begin by introducing the notion of *affinity intensity*. Let $|\mathbf{U}_i^t|$ be the number of new users towards product p_i at time slot t. Then the *affinity intensity* of p_i at t is defined as: $\alpha_i(t) = \frac{|\mathbf{U}_i^t|}{\sum_{\tau=1}^{T} |\mathbf{U}_i^\tau|}$. Note that T is the number of time slots over which the affinity is normalized.

Figure 4(a) reports the evolution of affinity intensity values of five different products over time. Note that the label in front of a product name indicates the category of the product. The x-axis denotes the number of weeks since the product first appeared in the website, while y-axis represents the affinity intensity towards the product over different weeks. In the sequel, we refer to such curve as *Affinity Intensity Curve* (AIC). Observe that out of the five AIC in Figure 4, the AIC of *Twitter* and *Gmail* look *similar*, the two products of `game` also exhibit similar AIC while that of *The Dark Knight* is quite different from the rest. In other words, *these curves show that products in the same category tend to have similar affinity evolution patterns.*

We compute the distance between different curves using the DTW (Dynamic Time Warping) distance with Sakoe-Chiba band [7] which adds a window constraints w to the warping path found by DTW algorithm. Our results show that the average distances between products in the same category are always smaller than those from different categories for all values of w [5].

We compute *affinity evolution distance* for a product p_i at time t (denoted by ϕ_i^t) using the notion of DTW distance. Since we intend to measure how similar an AIC of a product is compared to the product in the same category with most number of new users, we quantify ϕ_i^t by measuring the DTW distance between the AIC of a product p_i and the AIC of product p_j whose affinity rank is the highest in the same category. The formal algorithm is given in [5].

Average Rating. Let $\mathbf{R}_i^t$ be the bag of ratings that product p_i received during time slot t. Then the *average rating* of p_i during t, denoted by $\overline{\mathbf{R}}_i^t$, is defined as: $\overline{\mathbf{R}}_i^t = \frac{1}{|\mathbf{R}_i^t|} \sum_{r \in \mathbf{R}_i^t} r$.

Figure 4(b) depicts the AIC of *Gmail* in *Blippr* dataset as well as its average rating evolution. Observe that the evo-
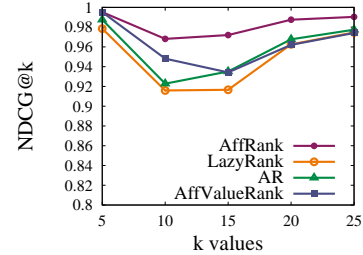

Figure 5: Comparison with different models.

lutions of affinity intensity and average rating follow similar trend except that there is a certain delay δ between the peaks of the two curves. In general, most of the δ values fall in the interval $[0, 3]$ [5]. Hence, we incorporate the average users' ratings for $\delta \in [0, 3]$ days as a feature in our affinity rank prediction model. If we denote the upper bound of δ as ℓ, then this feature can be computed as $\overline{\mathbf{R}}_i^{t-\delta-1}$ over all $\delta \in [0, \ell]$.

3. AFFINITY RANKS PREDICTION

As the affinity rank of a product is highly related to its ranks in the near past, it can be modeled by an ARX (AutoRegressive model with exogenous inputs) model [2]. Specifically, the ARX model of orders g and h is given as follows.

$$y_t = \varepsilon_t + \sum_{i=1}^{g} \varphi_i y_{t-i} + \sum_{j=1}^{s} \sum_{i=1}^{h} w_{i,j} b_{t-i,j} \qquad (2)$$

In this equation, y is the time series data (*i.e.,* product ranks in various different time slots), s is the number of exogenous input features; $\varphi_1, \ldots, \varphi_g$ and $w_{1,1}, \ldots, w_{h,s}$ are the parameters to be estimated from the training data. Both g and h are the orders in the model to be manually determined before model estimation. In our context, g is the *order of product rank* which determines the number of previous product ranks to be considered in the modeling; h is the *order of feature* determining the number of past time slots from which the values of the corresponding features to be involved in the model estimation. The variable $b_{t-i,j}$ is the value for feature j at time $t-i$, and ε_t is white noise. The estimation of the ARX model is efficient as it solves linear regression equations in analytic form. Also, the solution is unique and always satisfies the global minimum of the loss function [2].

In the ARX model, the order of product rank g and the order of feature h need to be manually determined. A common way of selecting g and h is to fix the value of g (or h) and try a range of values for h (or g); for each pair of g and h values, evaluate the accuracy of the model estimated using a measure called FPE [1]. A smaller FPE value means a more accurate model. Based on the FPE measures, we set $g = 5$ and $h = 1$ for our experimental study. Also, we use the latest 10 weeks data to estimate the ARX model. The algorithm to predict the affinity ranks using the learned ARX model is given in [5].

4. EXPERIMENTS SUMMARY

In this section, we report experimental results on *Blippr* dataset (Table 1). To evaluate the accuracy of predicted product rank against the ground-truth rank at a given time slot, we adopted Normalized Discounted Cumulative Gain (NDCG) measure: $NDCG@k = \frac{1}{Z} \sum_{i=1}^{k} \frac{rel_i}{\log_2(i+1)}$ [4]. We vary

Table 2: Top 10 popular products in weeks 12 & 13 (2009) and predicted ranks. (Blippr)

Rank	Week 12 (*LazyRank*)	Week 13	Predicted rank in week 13		
			AffRank	*AR*	*AffValueRank*
1	Twitter	Twitter	Twitter	Twitter	Twitter
2	Gmail	Gmail	Gmail	Gmail	Gmail
3	Mashable	Mashable	Mashable	Mashable	Mashable
4	Google	Tweetie for iPhone	Google Earth	Google Earth	Google
5	Google Earth	Google Earth	Google	Google	Google Earth
6	OK Computer	Google	Tweetie for iPhone	Google Reader	The Dark Knight
7	Google Reader	The Dark Knight	The Dark Knight	The Dark Knight	Tweetie for iPhone
8	Dropbox	The Shawshank Redemption	The Shawshank Redemption	Tweetie for iPhone	The Shawshank Redemption
9	WordPress	Google Reader	Google Reader	WordPress	Google Reader
10	In Rainbows	Watchmen	Watchmen	OK Computer	OK Computer

k from 5 to 25 at the step of 5 to evaluate the accuracy of the rank involving top-k ranked products. We define the relevance of a product p_i to be the inverse of its ground-truth rank: $\frac{1}{\rho_t(p_i)}$. Note that the value of NDCG heavily depends on the definition of relevance (i.e., rel_i). However, for a given relevance definition, NDCG well reflects the accuracies of different ranking models.

We compare the performance of the proposed *AffRank* with three other methods.

LazyRank. This model predicts product rank at time slot t to be the same as the rank obtained in the last time slot $t − 1$.

AR. *AR* model refers to the AutoRegressive model without taking exogenous input features (see Equation 2). In another word, the product rank y_t is predicted solely by the ranks in the past g time slots, $y_{t−g}$ to $y_{t−1}$, for a given product rank order g.

AffValueRank. With *AffValueRank*, instead of predicting the affinity rank, the exact affinity value is predicted using the ARX model. That is, $y_{t−i}$ in Equation 2 is set to be the affinity $a_i^{t−i}$. Using the learned model, the affinity values a_i^t is predicted; the products are then ranked accordingly.

As reported in Figure 5, the proposed *AffRank* model outperform all three baseline models for every k value. Overall, *AffValueRank* is the second best performing model followed by *AR*. The *LazyRank* model performs the worst. Considering the features involved in the four models, the experimental results show that, (i) product affinity rank can be better predicted using the past few product ranks than the single last rank (i.e., $AR > LazyRank$), (ii) the extra features besides the product rank lead to better prediction (i.e., *AffValueRank > AR*), and (iii), product affinity ranks can be better predicted than the affinity values (i.e., *AffRank > AffValueRank*) probably due to the smoother distribution of product affinity ranks than affinity values.

Table 2 shows an example of predicted affinity rank in *Blippr*. The second and third columns show the top-10 ranked products in weeks 12 and 13 of year 2009, respectively. The remaining 3 columns list the predicted ranks using different models. Obviously, our model *AffRank* accurately predicts the top 10 products in week 13 although the exact ranks for products *Google Earth*, *Google* and *Tweetie for iPhone* are not accurate. Besides, our model also detects that *Google Earth* will acquire more affinity than *Google* in week 13. Compared to week 12, there are four products

newly listed in top-10 (i.e., *Tweetie for iPhone*, *The Dark Knight*, *The Shawshank Redemption*, and *Watchmen*). Our *AffRank* predicted all four accurately, *AffValueRank* missed one of them, *AR* missed two of them, and the *LazyRank* missed all four.

5. CONCLUSIONS

In this paper, we proposed a predictive model called *AffRank*, that utilizes an array of features to predict the future rank of products according to their affinities in product review websites. Informally, product affinity refers to a product community's ability to "attract" new members and is measured by the number of new ratings during a specific time slot. We formulate the product affinity prediction problem as an autoregressive model with exogenous inputs (features). We have identified three features, namely affinity rank history, affinity evolution distance, and average rating, for predicting product affinity. Specifically, we discovered several interesting findings related to these features which we exploit in our model. Firstly, the affinity of a product for most products in *Blipper* tends to increase spikily and decrease smoothly. Secondly, the average DTW distances between products in the same category are always smaller than those between products from different categories. Thirdly, as the average rating increases the affinity intensity increases accordingly within 3 days. Our experimental study demonstrate the effectiveness and superior prediction quality of *AffRank* compared to three baseline methods.

6. REFERENCES

[1] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Aut. Contr.*, 19:716–723, 1974.

[2] P. J. Brockwell and R. A. Davis. *Introduction to Time Series and Forecasting*. Springer, March 2002.

[3] Q. Feng, K. Hwang, and Y. Dai. Rainbow product ranking for upgrading e-commerce. *IEEE Internet Computing*, 13(5):72–80, 2009.

[4] K. Järvelin and J. Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM Trans. Inf. Syst.*, 20(4):422–446, 2002.

[5] H. Li, S. S. Bhowmick, and A. Sun. Affinity-Driven Prediction and Ranking of Products in Online Product Review Sites. *Tech. Report*, Available at *www.cais.ntu.edu.sg/∼assourav/TechReports/PAffinity-TR.pdf*.

[6] Y. Matsuo and H. Yamamoto. Community gravity: measuring bidirectional effects by trust and rating on online social networks. In *WWW*, 2009.

[7] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. on Acoustics, Speech, and Signal Proc.*, 26(1):43–49, 1978.