Evolving Recommender Systems: Modeling, Analysis and Experiments*

Hongru Zhu
Shanghai Jiao Tong University
auberginepp@sjtu.edu.cn

Luoyi Fu
Shanghai Jiao Tong University
yiluofu@sjtu.edu.cn

Xinbing Wang
Shanghai Jiao Tong University
xwang8@sjtu.edu.cn

ABSTRACT

In many realistic recommender systems (RSs), user ratings on items and social connections among users usually interact in the context of an embedded evolving process where new users and items constantly arrive over time. However how to mathematically model such evolving RSs, along with the corresponding quantitative characterizations, remains unexplored.

Motivated by this, we take the initiative to propose a novel evolving RS model, which, as validated by our empirical results, can well capture some basic features of RSs. Particularly, two types of results are presented in this paper. (i) Our model is primarily based on a weighted bipartite graph structure composed of users and items. With edge weights abstracted from user ratings on their purchased items, an evolving process is proposed to highlight the establishment of social connections for new users and the characterization of their purchases based on rating-driven preferential attachment. (ii) By analytical derivations, along with further empirical validation on eight different real RS datasets, we disclose several RS properties. Specifically, in addition to power-law degree distributions of both users and items, we find that the distribution of total rating scores given by each user or received by each item, is bounded by power-law distributions determined by the rating scale in the RS. And interestingly, the beginning of the distribution of total rating scores, instead of being pure power-law, resembles the global rating probability mass function in certain situations.

Our model discloses how users' rating behaviors, together with RS parameters like rating scales, can alter item markets by affecting item popularity, and can provide insights into the design and analysis of RSs for RS operators and developers.

1 INTRODUCTION

Recommender systems (RSs) [1], which are traditionally designed to assist users to find relevant information and make better choices in face of chaos of overloaded information, offer customized information access for specific domains. In an effective RS where users are provided with relevant and useful items, they are more willing to make purchases or revisit the website, which in turn increases the site traffic and revenue. Recently RSs tend to introduce "social" features in their design. For instance, Amazon allows users to post their purchases on social media like Facebook or Twitter. Netflix shows users the films that their friends have recently watched. Yelp makes use of opinions from users' friends to personalize restaurant rankings. These features bring apparent benefits to websites revenue since friends tend to affect each other on purchases and tastes for product likes movies, and additionally people enjoy interacting with friends online. Considering that friends probably have common interests and are more likely to trust each other, it is almost certain that users tend to follow recommendations from their friends [2, 3]. Therefore it is essential to explore the structure of the underlying social network, which contributes to provide more trusted recommendations, to improve RSs.

Regarding this, there have been some initial efforts directed toward performance improvement of RSs through exploitation of social characteristics. For example, Su et al. [4] study friend recommendation in Twitter, and show that recommendations of popular users are more likely to be accepted than recommendations of "average" users. With both local and global metrics introduced, Wong et. al. [5] analyze the efficiency of social RSs, aiming at simultaneously maximizing individual's benefit and the efficiency of network in information dissemination. Other types of work include investigating the limit of RSs that rely heavily on popularity [6] and more precise recommendation from the perspective of user interests [7]. However, all those works fail to consider a key factor, i.e., the network evolution, a common phenomenon in a flurry of social networks. For example, in purchasing websites such as Amazon, new items are constantly increasing along with new customers arriving and making purchases. The same phenomenon holds in Youtube, where new videos keep being uploaded and new users are registering.

Inspired by the introduced "social" features and the evolving nature of RSs, we alternatively view RSs as evolving networks composed of groups of users and items, which can better capture "social" and evolving features simultaneously. Yet one difficulty to approach the problem from this perspective is to characterize the interaction both within the user group and between users and items. Moreover, with the constant arrival of new users and items these two types of interaction will gradually alter the item market, and the user and item popularity, or rather the total rating scores given or received, will also evolve in the rating-driven evolving process, normally resulting in a virtuous cycle:

- A newly arrived user (item), with a high probability, will show preference on popular item (user) whose total rating is high.
- Preference, with a high probability, will result in a new rating to popular item (user), adding up to its total rating and strengthening its popularity.
- As the RS keeps evolving, popular agents become more popular while outmoded agents gradually become neglected.

Under such circumstance, the distribution of degrees and the distribution of total rating scores given or received manifest their crucial roles in RSs and can be the key factor influencing newly added entities. Therefore it is essential to a fundamental question in this context, and the one we intend to tackle in this paper is the following: What affects the user and item popularity in an evolving RS with social characteristics, and how? For instance, suppose a community of users rate a set of movies on a movie recommender

*Produces the permission block, and copyright information

system. If we change some RS parameters like setting the rating scale from a scale of one to five stars to a scale of one to ten stars, will the movie popularity greatly differs in these two settings?

To understand how popularity is affected in the evolving process, we start by introducing an rating-driven model to capture the evolving and "social" features of RSs. As Figure 1 illustrates, the evolving process in our model incorporates two symmetrical aspects, i.e., the arrival of new users and items. At each time step, a new user arrives, selects an existing user (someone who possibly shares common interests with him and has a high total rating) as the "prototype" and establishes connection with him. Then the new user chooses among the items purchased by the "prototype" with a probability proportional to the rating that the "prototype" gives to each item. The new user will further rate those selected items according to his own judgement and give ratings from a discrete rating set $\{1, 2, \ldots, R\}$. A symmetrical process also occurs to items.

Based on the proposed model, we derive some general results for our evolving RS model. First we prove that when an RS evolves in the above prescribed way, the degrees of users and items are power-law distributed, and the exponents relate to the ratio of the number of items and users as well as the least number of ratings each has given or received. Then we proceed to conclude that under any assumption of users' rating habits the distribution of total ratings given or received, can be upper and lower bounded by power-law distributions, which, to a great extent, are affected by the rating scale in the RS. Additionally, we also look into the beginning of item vertex weight distribution where total ratings are small and prove that it is far from being a pure power-law and resembles the global rating probability mass function in certain situations.

Our main contributions are summarized as follows:

- Modeling: We develop an evolving RS model from the weighted bipartite graph structure which well captures basic features of evolving RSs. Our proposed evolving process highlights the establishment of social connections for new users and the characterization of their purchases based on rating-driven preferential attachment. We also exhibit the alternate influence of user and item popularity in the evolving process.
- Analysis: We analyze our evolving RS model and show that the degree distribution and the distribution of total rating scores are either power-law or bounded by power-laws determined by RS parameters like the rating scale. We also look into the beginning of the distribution of total rating scores, which resembles the global rating probability mass function in certain situations.
- Experiments: We empirically validate our model through experiments on eight real RS datasets, from which we demonstrate that our model fits well to real data and well captures features of realistic RSs. Our validation further indicate that these findings could provide insights for RS operators and developers into the RS design and analysis from both user and item perspective.

The rest of this paper is organized as follows. In Section 2 we provide the relevant literature. We define our evolving RS model in Section 3 and mathematically analyze its properties in Section 4. The theoretical results are further testified through experiments in Section 5. We provide concluding remarks in Section 6. The proofs are either presented in line or available at Appendix.

2 RELATED WORKS

We note that there is no prior work, other than ours, that focuses on the analysis of evolving RSs. However, there is prior work on non-evolutional RS models and the modelling of evolving networks.

Prior studies in non-evolutional RS models have covered serval aspects on RS improvements and design. Some are devoted to RS performance analysis from algorithmic perspectives [8–10]. Others are directed toward RS performance evaluation from network perspectives [5, 11]. Most recently, several studies also exploit social features in RS design. Vahabi et al. [12] devise a novel RS exploiting the anticipated social-network information diffusion. Bressan et al. [6] introduce a RS model based on popularity and the power of users in influencing others. Zeng et al. [13] clarify the existence of core users carrying most of information for recommendation in RSs. Liu et al. [7] propose a unified RS to integrate user interests and evolving sequential preference with temporal interval assessment.

There is also a flurry of prior work in evolving network sciences. A series of studies [14–17] have illustrated the structure of evolving social networks with the arrival and departure of users [18, 19] and temporary dynamics of interest [20]. Some work features the evolving process in the network model. Chung et al. [21] introduce the assumption of either vertex-arrival or edge-arrival at each time step. Ghoshal et al. [22] propose a model that is able to clarify the role of individual elementary mechanisms. Lattanzi et al. [23] introduce the affiliation network model, where preferential attachment and edge copy are emphasised in the evolving process.

3 MODEL OF EVOLVING RS

In this section, we first present our RS model, assumptions and the evolving algorithm. Then we provide a situation to understand the intuition behind the evolving model and discuss some essential aspects that reflect key characteristics of our model.

3.1 Mathematical Modeling

We begin by introducing our weighted bipartite graph structure that we refer to as $B(U, I)$ for evolving RSs and providing definitions of key concepts to understand how this structure is related to RSs. Specifically the two groups of vertices in such a weighted bipartite graph correspond to the groups of users and items in RSs respectively and the edge weights between vertices are the rating scores that users give to items. We further define the vertex weight of a vertex to be the sum of edge weights on all the edges that are connected to that vertex in our weighted bipartite graph. Therefore the vertex weight is naturally a representation of the total rating score a user gives or an item receives in RSs. This structure is designed to suit and manifest the properties of ordinary social networks and the evolving process in the bipartite graph.

Before we proceed to exhibit the detailed evolving process in the weighted bipartite graph, we introduce our assumptions on how users rate items in our model, namely the assignment of edge weights. Generally we assume that all rating scores are sampled from a mixture distribution yet we have explored two distinct frameworks for our assumptions. The first framework is simple and ratings are assumed to be affected by either the new user or the new item itself. We also have a second framework and it is more complex in that ratings are assumed to be affected by the joint influence of the user and the selected item.

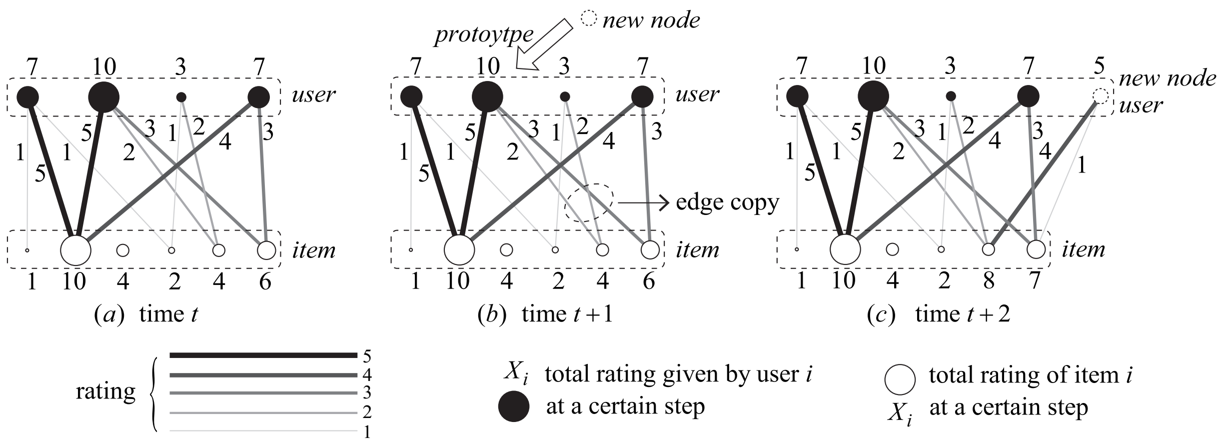


Figure 1: An illustration of how the recommender system evolves in terms of users, items, and ratings.

The common assumptions in the two frameworks are:

(1) There are K basic user types and each real user u reflects a mixture of those K types with weight vector v_u when rating items.

(2) There are L basic item levels and each real item i reflects a mixture of those L levels with weight vector t_i when being rated.

(3) For all users, their weight vectors are distributed according to some parameterized distribution F (possibly Gaussian) with a parameter vector θ and the probability density function (pdf) is f_θ.

(4) For all items, their weight vectors are distributed according to some parameterized distribution G (possibly Gaussian) with a parameter vector γ and the pdf is g_γ.

(5) The global expectation of a rating score is Er.

In the first framework we also assume:

(1) For any basic user type k there is a unique corresponding rating probability mass function (pmf) $h_k(r)$, $r = 1, 2, ..., R$. Symmetrically for any basic item level l, there is a unique corresponding rating pmf $h^l(r)$, $r = 1, 2, ..., R$.

(2) When a newly added user u gives his/her ratings on selected items, his/her rating pmf can be presented as:

$$H_u(r) = \sum_{k=1}^{K} v_u(k) h_k(r).$$

Symmetrically when a newly added item i are being rated by several users, its rating pmf can be presented as:

$$H^i(r) = \sum_{l=1}^{L} t_i(l) h^l(r).$$

Yet in the second framework we assume:

(1) Given any pairs of the basic user type k and the basic item level l, there is a unique corresponding rating probability mass function (pmf) $h_k^l(r)$, $r = 1, 2, ..., R$.

(2) When a user u gives his/her rating on the item i, his/her rating pmf can be presented as:

$$H_u^i(r) = \sum_{k=1}^{K} \sum_{l=1}^{L} v_u(k) t_i(l) h_k^l(r).$$

The detailed evolving process in our weighted bipartite graph $B(U, I)$ is shown in Table 1.

Table 1: Rating-Driven Evolving Process in $B(U, I)$

Fix two integers c_u,c_i > 0, and let $\beta \in (0, 1)$. Fix an integer R as the highest rating.
Process:
At time 0: Weighted bipartite graph $B(U, I)$ is a simple graph with at least $c_u * c_i$ edges, where each vertex in U has at least $c_u$edges and each vertex in I has at least c_i edges. The edge weights are sampled from the assumed mixture distribution in the discrete rating set $\{1, 2, . . . , R\}$ with expectation Er.
At time $t > 0$: **(Evolution of U)** With probability β: **(Arrival)** A new vertex u is added to U. **(Preferentially chosen Prototype)** A vertex $u' \in U$ is chosen as $prototype$ with a probability proportional to its vertex weight, i.e., the weights on all edges connected to it. **(Edge-copying)** c_u edges are "$copied$" from u'; that is, c_u neighbors of u', denoted by $n_1, n_2, \ldots, n_{c_u}$ are chosen with probability proportional to the weight of the edges in between (without replacement), and the edges $(u, n_1), (u, n_2), \ldots, (u, n_{c_u})$are added to the graph with weights sampled from the assumed mixture distribution in the discrete rating set $\{1, 2, . . . , R\}$. **(Evolution of I)** With probability $1 - \beta$, a new vertex $i$is added to I following a symmetrical process, adding c_i edges to i.

3.2 An Intuitive Understanding

To understand the intuition behind the evolving process, let us consider, for example, a movie RS. In this case the bipartite graph consists of the user set U and the movie set I. When a new user arrives, he is likely to watch movies that receive high ratings from one of his friends or one popular RS user, who may have written many reviews and have a trusted taste for movies. That friend can be viewed as "*prototype*" and influence the new user on his selection of movies to watch and write reviews on. Similarly when a new movie is added, before it begins advertising on potential viewers, it may seek in the existing movie set another popular movie similar in theme or cast and treat it as "*prototype*". The new movie is likely to build connections with those who have watched the "*prototype*" movie and given positive reviews on it. These two processes alternate and evolve as the RS gradually grows.

Here we explore some essential aspects of our model. The degree distribution is a main property for evolving RSs. The evolutionary forms of the model are showed by edges and vertices, which are used to analyze network characteristics. In conventional unweighted models proposed in prior literature, evolutionary forms are mainly related to degree distributions yet in our weighted bipartite graph, the vertex weight distribution, namely the distribution of total rating scores is another characteristic that relates to evolving features of RSs. The vertex weight distribution can affect the evolutionary form and provide insight into how ratings influence new purchases and how previous purchases influence new ratings in turn. In the evolving algorithm, we emphasize the establishment of social connections between users and thus embed social characteristics in our analysis of evolving RSs.

4 ANALYSIS OF DEGREE AND VERTEX WEIGHT DISTRIBUTIONS

We present here theoretical results regarding degree and vertex weight distributions in $B(U, I)$. Theorem 4.2 shows power-law degree distributions. Theorems 4.3 and 4.4 show power-law bounds for vertex weight distributions. Proposition 4.7 and Theorem 4.8 analyze the beginning of vertex weight distributions.

4.1 Degree Distribution of Evolving RSs

We start with the degree distribution in our RS model $B(U, I)$ and introduce Lemma 4.1 before we state Theorem 4.2 regarding degree distributions.

Lemma 4.1 ([24]). *If a sequence a_t satisfies the recursive formula $a_{t+1} = (1 - b_t/t)a_t + c_t$ for $t \geq t_0$, where $\lim_{t\to\infty} b_t = b > 0$ and $\lim_{t\to\infty} c_t \geq c$ exists. Then $\lim_{t\to\infty} a_t/t$ exists and equals $c/(1 + b)$.*

With an approach similar to that in [23], We derive Theorem 4.2.

Theorem 4.2. *For the weighted bipartite graph $B(U, I)$ generated after n steps, when $n \to \infty$, the ensemble average of the degree sequence of vertices in U (resp. I) follows a power-law distribution with exponents $\alpha = -2 - \frac{c_u\beta}{c_i(1-\beta)}$ $\left(\alpha = -2 - \frac{c_i(1-\beta)}{c_u\beta}\right)$.*

Proof. Provided in Appendix. A in technical report [30]. □

4.2 Vertex Weight Distribution of Evolving RSs

Degree distributions reflect popularity from the perspective of purchase times. Here we investigate another reflection of popularity, i.e., the vertex weight, namely the total ratings given or received. we give the upper and lower bounds of vertex weight distributions in Theorems 4.3 and 4.4 separately and provide illustrations on techniques used to prove these two theorems in Figure 2.

4.2.1 *Upper Bound of Vertex Weight Distribution.*

Theorem 4.3. *For the weighted bipartite graph $B(U, I)$ generated after n steps, when $n \to \infty$, the ensemble average of the vertex weight sequence of vertices in U (resp. I) has an upper bound which follows a power-law distribution with exponent $\alpha = -\frac{1+C}{R}$, where C is a constant and equals $Er\left(1 + \frac{c_u\beta}{c_i(1-\beta)}\right)$ $\left(Er\left(1 + \frac{c_i(1-\beta)}{c_u\beta}\right)\right)$ for any vertex with vertex weight greater than $c_uR + 2R$ $(c_iR + 2R)$.*

Proof. The proof of Theorem 4.3 includes 3 sequential parts:
I. derivation of the recurrence relation of vertex weight sequence,
II. derivation of the intermediate upper bound,
III. derivation of the power-law upper bound.

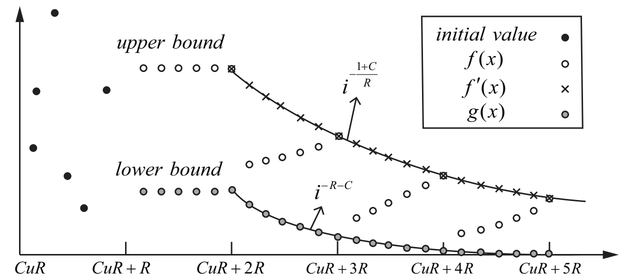


Figure 2: An illustration of proving upper and lower bounds of the vertex weight distribution.

PART I. *Derivation of the Recurrence Relation*

Let V_t^k be the expected number of vertices in U with vertex weight k at time t. When $k > c_uR + R$ we have

$$\begin{aligned} V_t^k =& V_{t-1}^k - E[\#\text{ of vertices in } U \text{ with vertex weight } k \\ & \text{at time } t-1 \text{ and increase at time } t] \\ & + E[\#\text{ of vertices in } U \text{ with vertex weight} < k \\ & \text{at time } t-1 \text{ and increase to } k \text{ at time } t]. \end{aligned}$$

The vertex weight of vertices in U can increase if and only if a new vertex is added to I. Thus we have

$$\begin{aligned} V_t^k =& V_{t-1}^k - (1-\beta)c_i\frac{kV_{t-1}^k}{W_{t-1} + W_{B_0}} \\ & + (1-\beta)\sum_{j=k-R}^{k-1} E[\#\text{ of vertices in } U \text{ with vertex weight} \\ & j \text{ at time } t-1 \text{ and increase to } k \text{ at time } t| \\ & \text{a new vertex is added to } I]. \end{aligned}$$

Similar to the proof of Theorem 4.2 the probability that an edge is selected in a single selection is proportional to its weight. Thus,

$$\begin{aligned} & E[\#\text{ of vertices with vertex weight } j \text{ at time } t-1 \text{ that is chosen} \\ & \text{as end point}|\text{ a new vertex is added to } I] = \frac{jc_iV_{t-1}^j}{W_{t-1} + W_{B_0}}. \end{aligned}$$

Since we have

$$\begin{aligned} & E[\#\text{ of vertices in } U \text{ with vertex weight } j \text{ at time } t-1 \\ & \quad \text{and increase to } k \text{ at time } t|\text{ a vertex is added to } I] \\ =& E[\#\text{ of vertices in } V_{t-1}^j \text{ that is chosen as end points}| \\ & \quad \text{a new vertex is added to } I] \\ & \times E[Pr[\text{the edge weight is assigned to be } k-j]], \end{aligned}$$

We call $H(r) = E[Pr[\text{the edge weight is assigned to be } r]]$ to be the expected rating pmf for new users. Using our assumptions in Section 3, we have the following results under the two different frameworks. In the first framework, we have:

$$\begin{aligned} H(r) &= E[Pr[\text{the edge weight is assigned to be } r]] \\ &= \int_\gamma g_\gamma H_\gamma(r)d\gamma \\ &= \int_\gamma g_\gamma \sum_{l=1}^{L} t_\gamma(l)h^l(r)d\gamma. \end{aligned}$$

Or in the second framework, we have:

$$
\begin{aligned}
&E[Pr[\text{the edge weight is assigned to be } r]]\\
&=H(r) = \int_\theta \int_\gamma f_\theta g_\gamma H_{\theta,\gamma}(r) d\gamma d\theta\\
&= \int_\theta \int_\gamma f_\theta g_\gamma \sum_{k=1}^{K} \sum_{l=1}^{L} v_\theta(k) t_\gamma(l) h_k^l(r) d\gamma d\theta.
\end{aligned}
$$

Since we know that

$$\sum_{r=1}^{R} h_k^l(r) = 1, \sum_{r=1}^{R} h^l(r) = 1,$$

we always have

$$\sum_{r=1}^{R} H(r) = 1.$$

we can derive

$$V_t^k = V_{t-1}^k \left(1 - \frac{(1-\beta)c_i k}{W_{t-1} + W_{B_0}}\right) + (1-\beta) \sum_{j=k-R}^{k-1} \frac{c_i j V_{t-1}^j H(k-j)}{W_{t-1} + W_{B_0}}.$$

Let $X_k = \lim_{t\to\infty} V_t^k / t$. Again, using Lemma 4.1, we get the following recurrence relation

$$X_k = \frac{(1-\beta)\frac{c_i}{Er(c_u\beta + c_i(1-\beta))}}{1 + (1-\beta)\frac{c_i k}{Er(c_u\beta + c_i(1-\beta))}} \sum_{j=k-R}^{k-1} H(k-j) j X_j,$$

namely,

$$X_k = \frac{1}{k+C} \sum_{j=k-R}^{k-1} H(k-j) j X_j.$$

PART II. *Derivation of the Intermediate Upper Bound*

This high order homogeneous recurrence relation is not detachable as in the proof of Theorem 4.2. Without loss of generality, we assume an arbitrary set of initial values which are not all zeros, as shown in Figure 2 and suppose f is one intermediate upper bound. We have the following 2 cases:

When $c_uR + R + 1 \le k \le c_uR + 2R$, let

$$f(k) = \max_{cuR+R+1 \le j \le cuR+2R} X_j.$$

When $k > c_uR + 2R$, $H(r)$ sums to 1 and we have

$$
\begin{aligned}
X_k &= \frac{1}{k+C} \left(H(R)(k-R)X_{k-R} + \ldots + H(1)(k-1)X_{k-1}\right)\\
&\le \frac{1}{k+C} \times \max_{k-R\le j\le k-1} f(j)\\
&\quad \times \left(H(R)(k-R) + H(R-1)(k-R+1) + \ldots + H(1)(k-1)\right)\\
&\le \frac{k-1}{k+C} \max_{k-R\le j\le k-1} f(j),
\end{aligned}
$$

namely,

$$f(k) = \frac{k-1}{k+C} \max_{k-R\le j\le k-1} f(j). \tag{1}$$

Before we use induction to find the upper bound, we first deal with the initial case when k is in the range $[c_uR + 2R + 1, c_uR + 3R]$.

Recall that when $c_uR + R + 1 \le k \le c_uR + 2R$, $f(k)$ is fixed and equals $f(c_uR + 2R)$. Thus for $c_uR + 2R + 1 \le k \le c_uR + 3R$, since $\frac{k-1}{k+C} < 1$, using Equation (1) we have

$$\max_{c_uR+2R+1\le j\le c_uR+3R} f(j) < f(c_uR + 2R),$$

which is demonstrated in Figure 2.

Moreover, since $\frac{k-1}{k+C}$ is strictly increasing, we also have

$$\max_{c_uR+2R+1\le j\le c_uR+3R} f(j) = f(c_uR + 3R).$$

Ultimately, we get the following two equations for the initial case when $c_uR + 2R + 1 \le k \le c_uR + 3R$:

$$f(k) = \frac{k-1}{k+C} f(c_uR + 2R), \quad \max_{c_uR+2R+1\le j\le c_uR+3R} f(j) = f(c_uR + 3R).$$

By induction, suppose

if $n = K$, when $c_uR + nR + 1 \le k \le c_uR + (n+1)R$, we have

$$f(k) = \frac{k-1}{k+C} f(c_uR + nR), \text{ and}$$

$$\max_{c_uR+nR+1\le j\le c_uR+(n+1)R} f(j) = f(c_uR + (n+1)R).$$

Next we show the conditions above also hold when $n = K + 1$.

We know from Equation (1) that

$$f(k) = \frac{k-1}{k+C} \max_{k-R\le j\le k-1} f(j) < \max_{k-R\le j\le k-1} f(j).$$

Thus we can derive the following inequality

$$\max_{c_uR+(K+1)R+1\le j\le c_uR+(K+2)R} f(j) < f(c_uR + (K+1)R).$$

Hence we have

$$
\begin{aligned}
f(k) &= \frac{k-1}{k+C} f(c_uR + (K+1)R)\\
&= \frac{k-1}{k+C} f(c_uR + nR)\\
&< f(c_uR + nR).
\end{aligned}
$$

Since $\frac{k-1}{k+C}$ is increasing as k increases, $f(k)$ is increasing in the given range for k and

$$\max_{c_uR+nR+1\le j\le c_uR+(n+1)R} f(j) = f(c_uR + (n+1)R).$$

Now we can conclude that for any positive integer $n \ge 2$, when $c_uR + nR + 1 \le k \le c_uR + (n+1)R$, we have

$$f(k) = \frac{k-1}{k+C} f(c_uR + nR).$$

PART III. *Derivation of the Power-law Upper Bound*

We can derive the following relation

$$f(k) = \frac{k-1}{k+C} f(k-R),$$

$$\text{for every } k = c_uR + 3R, c_uR + 4R, \ldots$$

Recursively, the above equation results in

$$
\begin{aligned}
f(k) &= \frac{k-1}{k+C} \cdot \frac{k-R-1}{k-R+C} \cdot \frac{k-2R-1}{k-2R+C} \cdot \ldots \cdot \frac{c_uR+2R-1}{c_uR+2R+C}\\
&\quad \cdot f(c_uR + 2R)\\
&= \frac{\frac{k-1}{R}}{\frac{k+C}{R}} \cdot \frac{\frac{k-R-1}{R}}{\frac{k-R+C}{R}} \cdot \frac{\frac{k-2R-1}{R}}{\frac{k-2R+C}{R}} \cdot \ldots \cdot \frac{\frac{c_uR+2R-1}{R}}{\frac{c_uR+2R+C}{R}} f(c_uR + 2R)\\
&= \frac{\Gamma(\frac{k-1}{R}+1)}{\Gamma(\frac{k+C}{R}+1)} \frac{\Gamma\left(\frac{c_uR+2R+C}{R}+1\right)}{\Gamma\left(\frac{c_uR+2R-1}{R}+1\right)} f(c_uR + 2R)\\
&\sim \left(\frac{k}{R}\right)^{-\frac{1+C}{R}},
\end{aligned}
$$

for every $k = c_uR + 3R, c_uR + 4R, \ldots$.

As illustrated in Figure 2, we want to find a new upper bound f'. To do this, we first define the initial case

$$f'(c_u R + 2R) = f(c_u R + 2R),$$

and then instead of $k = c_u R + 3R, c_u R + 4R, \ldots$, suppose for any positive integer $k > c_u R + 2R$,

$$f'(k) = \frac{\Gamma(\frac{k-1}{R} + 1)}{\Gamma(\frac{k+C}{R} + 1)} \frac{\Gamma(\frac{c_u R + R + C}{R} + 1)}{\Gamma(\frac{c_u R + R - 1}{R} + 1)} f'(c_u R + 2R).$$

$f(k)$ is increasing in interval $[c_u R + nR + 1, c_u R + (n+1)R]$ for any positive integer n while $f'(k)$ is always decreasing and $f(c_u R + nR) = f'(c_u R + nR)$. As shown in Figure 2, $f'(k)$ is also an upper bound of X_k and it follows a power-law distribution with exponent $\alpha = -\frac{1+C}{R}$.

Thus we have proved that the upper bound of the ensemble average of the vertex weight distribution in set U follows a power-law distribution with exponent $\alpha = -\frac{1+C}{R}$. Symmetrically, we can prove a similar result for the vertices in I. □

4.2.2 Lower Bound of Vertex Weight Distribution.

Theorem 4.4. *For the weighted bipartite graph $B(U, I)$ generated after n steps, when $n \to \infty$, the ensemble average of the vertex weight sequence of vertices in U (resp. I) has a lower bound which follows a power-law distribution with exponent $\alpha = -R - C$, where C is a constant and equals $Er(1 + \frac{c_u \beta}{c_i(1-\beta)}) \left(Er\left(1 + \frac{c_i(1-\beta)}{c_u \beta}\right)\right)$ for any vertex weight greater than $c_u R + 2R$ ($c_i R + 2R$).*

Proof. Provided in Appendix. B in technical report [30]. □

Theorem 4.2 is consistent with Theorems 4.3 and 4.4 when the degree distribution is viewed as a special case of the vertex weight distribution where all edge weights are assigned to 1 in the weighted graph.

When $R = 1$ it is easy to verify that the exponents of both upper and lower bound power-law distributions are $-1 - C$, which is exactly the same as the exponent of the degree distribution we derive in Theorem 4.2. And note that the following inequality always holds

$$-R - C \leq -\frac{1+C}{R}.$$

4.2.3 Beginning of Vertex Weight Distribution.

In Sections 4.2.1 and 4.2.2, we show the vertex weight distribution is bounded by power-laws when the vertex weight k is greater than a certain value. Here we proceed to explore the beginning of vertex weight distribution when k is relatively small. Defined as the above, V_t^k is the expected number of vertices in U with vertex weight k at time t. Again let $X_k = \lim_{t\to\infty} V_t^k / t$. We define new random variable S to be the vertex weight of a newly added vertex in U with a pmf $s(k)$. Symmetrically we define random variable S' to be the vertex weight of a newly added vertex in I with a pmf $s'(k)$. Similar to the proof of Theorems 4.3 and 4.4, we define $H(r)$ to be the expected rating pmf for new users and $H(r)'$ to be the expected rating pmf for new items.

We introduce the following 3 propositions before we provide Theorem 4.8. The first two proposition give the recurrence relation and an upper bound of X_k. The third one describes the distributions of random variables S and S'.

Proposition 4.5. *For the ensemble average of the vertex weight distribution in set U (resp. I), when $c_u < k < c_u R$ ($c_i < k < c_i R$), the recurrence relation of X_k is*

$$X_k = \frac{1}{k+C} \sum_{j=k-R}^{k-1} H(k-j) j X_j + \frac{C\beta s(k)}{k+C}$$

$$\left(X_k = \frac{1}{k+C} \sum_{j=k-R}^{k-1} H'(k-j) j X_j + \frac{C(1-\beta) s'(k)}{k+C} \right),$$

where $C = Er\left(1 + \frac{c_u \beta}{c_i(1-\beta)}\right) \left(Er\left(1 + \frac{c_i(1-\beta)}{c_u \beta}\right)\right)$.

Proof. Provided in Appendix. C in technical report [30]. □

Proposition 4.6. *For $c_u \leq k \leq c_u R$ (resp. in I, $c_i \leq k \leq c_i R$), X_k have an upper bound $l(k) \leq 1$.*

Proof. Provided in Appendix. D in technical report [30]. □

Proposition 4.7. *When c_u (resp. c_i) is large enough, S (S') follows a unimodal probability distribution.*

Proof. Provided in Appendix. E in technical report [30]. □

Now we introduce Theorem 4.8 regarding the difference of vertex weight sequence and the difference of the nonhomogeneous term in the recurrence relation.

Theorem 4.8. *Let $p(k)$ be the nonhomogeneous term in recurrence relation in set U (resp. I),*

$$p(k) = \frac{C\beta s(k)}{k+C} \quad \left(p(k) = \frac{C(1-\beta) s'(k)}{k+C} \right)$$

. For $c_u < k < c_u R$ ($c_i < k < c_i R$), asymptotically

$$|\Delta X_k - \Delta p(k)| \leq H_u + \frac{1}{c_u + C}$$

$$\left(|\Delta X_k - \Delta p(k)| \leq H_i + \frac{1}{c_i + C} \right)$$

where

$$C = Er\left(1 + \frac{c_u \beta}{c_i(1-\beta)}\right) \left(Er\left(1 + \frac{c_i(1-\beta)}{c_u \beta}\right)\right),$$

and

$$H_u = \max\{H(1), H(R)\} + \sum_{r=1}^{R-1} |H(r+1) - H(r)|,$$

$$H_i = \max\left\{H'(1), H'(R)\right\} + \sum_{r=1}^{R-1} |H'(r+1) - H'(r)|.$$

Proof. Provided in Appendix. F. □

Here we discuss more about function $p()$ in set I, which is closely related to the beginning of the vertex weight distribution of items. For $p(k) = \frac{C(1-\beta)s'(k)}{k+C}$, it is a product of $s'(k)$ and another power function $q(k) = \frac{C(1-\beta)}{k+C}$ with exponent equal to -1. We know that for the derivative of a power function $q(k)$ with exponent -1 is

$$\frac{dq(k)}{dk} < 0, \frac{dq(k)}{dk} \to 0, k \to \infty.$$

When c_i is large enough, $s'(k)$ is a unimodal pmf. Since the absolute derivative of $q(k)$ is rather small and leads to steady $q(k)$, with a high probability, the product $p(k)$ increases before some value

Table 2: Datasets statistics and fitting parameters

Dataset	Rating scale	Number of users	Number of items	Number of ratings	α (item degree distribution)	Parameter C
MovieLens	[1,5]	700	10,000	100,000	-2.397	4.191
AmazonMovie	[1,5]	1,884,911	198,797	4,607,047	-2.018	3.054
AmazonCD	[1,5]	1,460,632	470,130	3,749,004	-2.184	3.552
YahooMusic	[1,100]	1,948,882	98,211	11,557,943	-2.015	51.256
Audioscrobbler	[1,5]	146,946	1,493,930	24,296,858	-2.003	3.009
AmazonBook	[1,5]	5,518,811	2,078,816	22,507,155	-2.287	3.861
BookCrossing	[1,10]	278,858	271,379	1,149,781	-2.335	7.343
AmazonElectronics	[1,5]	3,431,122	464,673	7,824,482	-2.019	3.057

and then decreasing to 0 as k increases. By Proposition 4.7 and Theorem 4.8, we can roughly depict the vertex weight distribution of items for $c_i \le k \le c_i R$ and it is probable that $p(k)$ and X_k are both unimodal function and the vertex weight distribution will demonstrate a peak at the beginning of the sequence.

When c_i is small, for instance in the extreme case when c_i equals 1, pmf $s'(k)$ is the same as expected rating pmf for new items $H'(r)$. Thus the shape of $p(k)$ and X_k will largely depend on the shape of $H'(r)$. $p(k)$ will fluctuate when $H'(r)$ has fluctuations since a power function tends to be steady and has small absolute derivatives when $k > 1$. Therefore the shape of $p(k)$ and X_k, or rather the vertex weight distribution will resemble $H'(r)$.

Symmetrically we have similar results for the analysis in set U.

5 EXPERIMENTS

In this section, We validate our theoretical results by conducting experiments on different realistic RS datasets and show the generalizable and explanatory ability of our model. We use the term "item" to represent any products in the RSs like music, book, moives and etc.

5.1 Experiment Settings

As shown in Figure 2, we use 8 publicly available RS datasets: the MovieLens movie recommendation datasets [25], the Audioscrobbler music artist rating datasets [26], the Yahoo music artist rating datasets [27], the BookCrossing book recommendation datasets [28], and the Amazon product recommendation datasets for movies, CDs, electronics and books [29]. These datasets vary in size, types of products and ratio of user and item numbers, testing the applicability of our model in different situations.

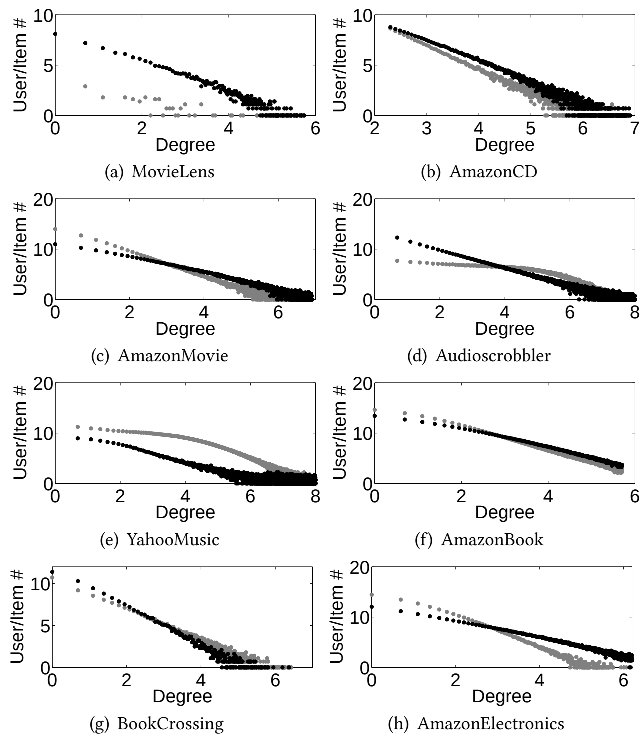
We make subtle modifications to some datasets before the experiments, including the adjustments of rating scales. For instance, in Audioscrobbler dataset, we map implicit rating information, namely the users' play count for each music artist to explicit integer ratings within range $[1, 5]$. Specifically if users listen only once to an artist, then the rating is 1. When the play count is in range [2,4],[5,9] or [10,19], the rating is set to 2,3,4 separately. A rating of 5 will be given if and only if users listen to the artist more than 19 times.

5.2 Performance Evaluations

In this section we first give the result of degree distributions of users and items in different datasets. Then we focus on the vertex weight distribution of items specifically for some datasets.

5.2.1 Degree Distributions of Items and Users.

Figure 3 shows degree distributions of users and items in 8 datasets in logarithm, which are demonstrated to be power-laws. The results fit to the symmetrical feature of degree distributions of users and items in our bipartite graph model.



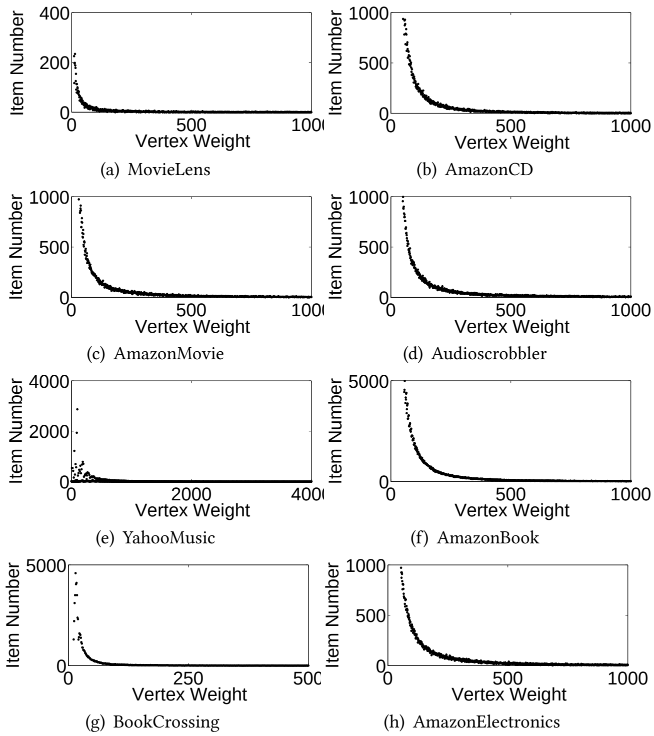
(a) MovieLens (b) AmazonCD (c) AmazonMovie (d) Audioscrobbler (e) YahooMusic (f) AmazonBook (g) BookCrossing (h) AmazonElectronics

Figure 3: Degree distribution of users (grey dot) and items (black dot) in logarithm

Clearly, item degree distributions have an unambiguous fit of power-laws in some datasets while in others like the Yahoo dataset the distribution seems divergent with high degrees. This phenomenon might result from a bias in favor of popular items that receives much attention during data collection. Thus the number of items with higher degrees fails to decay as assumed in some datasets. Here we fit item degree distributions in 8 datasets and avoid use the divergent part with high degrees. As shown in Table 2, the exponents of item degree distributions are ciphered out and they appear to be lower than -2, which is consistent with the conclusion in Theorem 4.2 that the exponents should be

$$\alpha = -2 - \frac{c_i(1-\beta)}{c_u \beta}. \tag{2}$$

Moreover, using Equation (3) we also calculate the parameter C used in deriving the exponents of the upper and lower bounds of item vertex weight distributions in Theorems 4.3 and 4.4 and provide them in Table 2 for later use in the exhibition of experiment



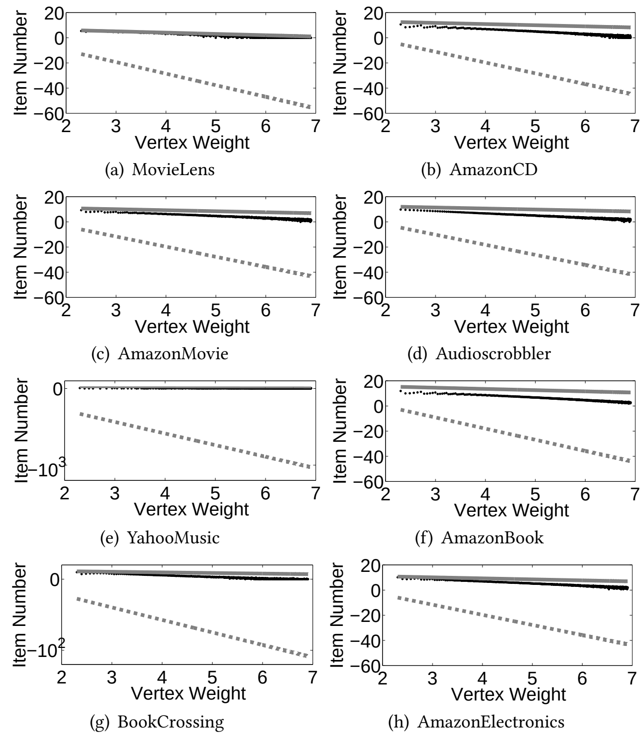
(a) MovieLens (b) AmazonCD

(c) AmazonMovie (d) Audioscrobbler

(e) YahooMusic (f) AmazonBook

(g) BookCrossing (h) AmazonElectronics

Figure 4: Vertex weight distribution of items



(a) MovieLens (b) AmazonCD

(c) AmazonMovie (d) Audioscrobbler

(e) YahooMusic (f) AmazonBook

(g) BookCrossing (h) AmazonElectronics

Figure 5: Upper (solid grey) and lower bound (dash grey) of vertex weight distribution (black) of items in logarithm

results for vertex weight distributions.

$$C = Er\left(1 + \frac{c_i(1-\beta)}{c_u\beta}\right). \tag{3}$$

There are some other interesting findings for degree distributions. Figure 3 shows a noticeable common feature for 2 music artist recommendation datasets, i.e., the Audioscrobbler dataset and the Yahoo dataset. User degree distributions in these two datasets, i.e. the grey plots in Figure 3(d) and 3(e) seem to be piecewise. When degrees are small, the fitted power-law distributions have small absolute exponents or rather small absolute slopes in logarithm graphs while the absolute exponents are larger with large degrees. Since we use asymptotic analysis in our proofs, the results are not necessarily true for very small degrees but is more trusted for large degrees.

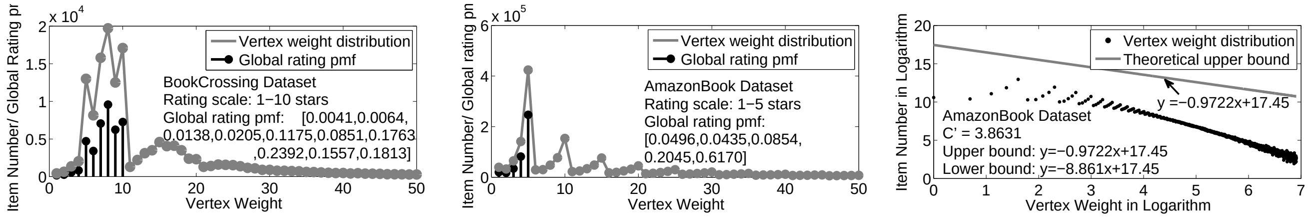
One explanation consistent with our evolving RS model is that users for music artist RSs can be roughly separated into 2 groups, i.e. senior users and junior users. Senior users are those who have used the RS for long and given many ratings, and junior users, on the contrary, do not often use the RS or have recently arrived. These two types of users behave quite differently in behavioral habits. Senior users might be more socially interactive while junior users are not much influenced by others. In the evolving process, senior users with large vertex weights have higher probabilities to be selected as "prototype" and socially interact with others by sharing their preference list. As time increases, some junior users become senior and their behavioral habits are altered and start to build up social features as their vertex weights grow larger. These features provide us with insights into how the evolving process and the social interactions can influence user degree distributions.

5.2.2 *Vertex Weight Distribution of Items.*

We plot in Figure 4 the vertex weight distributions of items in 8 different datasets. To further verify our model, we also plot in Figure 5 vertex weight distributions and their theoretical upper and lower bounds in logarithm. The upper and low bounds are calculated using the parameter C from Table 2.

As shown in Figure 5, item vertex weight distributions in 8 datasets are well bounded, implying that the distribution either follows a single power-law or a sum of power-laws with exponents also bounded by those of upper and lower bounds. We know from Theorems 4.3 and 4.4 that the theoretical bounds are affected by rating scales or rather the highest possible rating score R. The item vertex weight distribution is closer to the upper bound, providing an insight into how rating scales may influence item vertex weight distributions by affecting the exponents of its upper bound.

We subsequently look into the beginning of item vertex weight distributions. In Figure 4, no plot exhibits a peak at the beginning of the item vertex weight sequence, indicating a small c_i for selected datasets according to Proposition 4.7 and Theorem 4.8. With small c_i, it is more likely that item vertex weight distributions have similar fluctuations to those in expected rating pmf for items $H'(r)$. In Figure 6 we show item vertex weight distributions in book RSs from BookCrossing and AmazonBook. Figure 6(a) and 6(b) show that the beginning of item vertex weight distributions resemble $H'(r)$, which suggests that the average users' rating habit can influence the beginning of item vertex weight distributions. In an evolutional view, small c_i leads to a large number of items having small vertex weights. New Items are added to RS with small vertex weights and



(a) BookCrossing (b) AmazonBook (c) AmazonBook in logarithm

Figure 6: Vertex weight distribution for book recommendation datasets.

are not competitive with other popular items. New users choose items with probabilities proportional to edge weights, indicating that popular items with large vertex weights are more probable to be selected by new users and become more popular.

Surprisingly we find that in the vertex weight distributions in AmazonBook dataset, as shown in Figure 6(b), 5 different power-laws seem to coexist and the plot in logarithm in Figure 6(c) also supports the idea. With a detailed examination on the dataset, we find that users prefer to give a rating score of 5. In 6(b) we present the rating probabilities for different scores. Recall the recurrence relation we use in proving Theorems 2 and 3 is

$$X_k = \frac{1}{k+C} \sum_{j=k-R}^{k-1} H'(k-j) j X_j.$$

If $H'(5)$ is much larger, the recurrence relation yields to

$$X_k \approx \frac{H'(R)(k-R)X_{k-R}}{k+C},$$

indicating that X_k is dependent mostly on X_{k-5} and this results in 5 different power-law curves starting from 5 distinct initial values.

6 CONCLUSION

In this paper, we propose a novel evolving RS model that highlights the establishment of social connections for new users and the characterization of their purchases based on rating-driven preferential attachment. The superiority of our model lies in three aspects: good capture of realistic RSs, mathematically tractability and novelty in the context of evolving RSs with social characteristics. Subsequently we analyze our model and prove some basic properties including power-law degree distributions and the vertex weight distributions bounded by power-laws. We also investigate the beginning of the item vertex weight distribution, which resembles expected rating pmf for new items. Our model is finally verified through experiments on 8 real RS datasets, from which we validate our theoretical results and demonstrate that our evolving model can well capture realistic RSs.

REFERENCES

[1] P. Resnick and H. Varian. "Recommender Systems," in *Communications of the ACM*, Vol. 40, No. 3, pp. 56–58, Mar., 1997.

[2] D. Kempe, J. M. Kleinberg and É Tardos, "Maximizing the Spread of Influence through a Social Network," in *Proc. of KDD*, 2003, pp. 137–146.

[3] W. Chen, A. Collins, R. Cummings, T. Ke, Z. Liu, D. Rincon, X. Sun, W. Wei, Y. Wang and Y. Yuan, "Influence Maximization in Social Networks When Negative Opinions May Emerge and Propagate," in *Prof. of SDM*, Mesa, Arizona, USA, April, 2011.

[4] J. Su, A. Sharma and S. Goel. "The Effect of Recommendations on Network Structure," in *Proc. of WWW. 2016.*

[5] F. M. F. Wong, Z. Liu and M. Chiang, "On the Efficiency of Social Recommender Networks," in *IEEE/ACM Trans. on Networking*, Vol. 24, No. 4, pp. 2512–2524, Aug., 2016.

[6] M. Bressan, S. Leucci, A. Panconesi, P. Raghavan and E. Terolli, "The limits of popularity-based recommendations, and the role of social ties," in *Prof. of ACM KDD*, San Francisco, CA, USA, Aug., 2016.

[7] Y. Liu, C. Liu, B. Liu, M. Qu and H. Xiong, "Unified point-of-interest recommendation with temporal interval assessment," in *Proc. of ACM KDD*, San Francisco, CA, USA, Aug., 2016.

[8] B. J. Mirza, B. J. Keller and N. Ramakrishnan, "Studying Recommendation Algorithms by Graph Analysis," in *J. of Intelligent Information Systems* (2003) Vol. 20, No. 2, pp. 131–160, 2003.

[9] S. Chojnacki and M. A. Klopotek. "Random Graphs for Performance Evaluation of Recommender Systems," in *Journal of Control and Cybernetics*, Vol. 40, No. 2, 2011,

[10] Zhao, J., Zhang, H. and Y. Lian, "Analysis and Design of Personalized Recommender System Based on Collaborative Filtering," in *J. Internet of Things*, Vol. 312 of the Series Communications in Computer and Information Science, pp. 473-480, 2012.

[11] D. Lamprecht, M. Strohmaier and D. Helic, "Improving Reachability and Navigability in Recommender Systems," in *arXiv preprint arXiv:1507.08120*, 2015.

[12] H. Vahabi, I. Koutsopoulos, F. Gullo and M. Halkidi, "Difrec: A social-diffusion-aware recommender system," in *Proc. of CIKM*, pp. 1481-1490, Melbourne, VIC, Australia, Oct., 2015.

[13] W. Zeng, A. Zeng, H. Liu, M. Shang and T. Zhou, "Uncovering the Information Core in Recommender Systems," in Scientific Reports, 4:6140, Aug., 2014.

[14] M. Atzmueller, A. Ernst, F. Krebs, C. Scholz and G. Stumme, "On the Evolution of Social Groups During Coffee Breaks," in *Proc. of ACM WWW*, pp. 631–636, 2014.

[15] N. Z. Gong, W. Xu, L. Huang, P. Mittal, E. Stefanov, V. Sekar and D. Song, "Evolution of Social-Attribute Networks: Measurements, Modeling, and Implications using Google+," in *Proc. of ACM IMC*, pp. 131-144, 2012.

[16] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. T. Sukhavasi, C. Patel and S. Geirhofer, "Network Densification: The Dominant Theme for Wireless Evolution into 5G," in *IEEE Communications Magazine*, Vol. 52, No. 2, pp. 82–89, 2014.

[17] Y. Wu, N. Pitipornvivat, J. Zhao, S. Yang, G. Huang and H. Qu, "egoSlider: Visual Analysis of Egocentric Network Evolution," in *IEEE Trans. on Visualization and Computer Graphics*, Vol. 22, No. 1, pp. 260– 269, 2015.

[18] S. Wu, A. D. Sarma, A. Fabrikant, S. Lattanzi and A. Tomkins, "Arrival and Departure Dynamics in Social Networks," in *Proc. of ACM WSDM*, pp. 233–242, 2013.

[19] T. Zhang, P. Cui, C. Faloutsos, W. Zhu and S. Yang, "Come-and-Go Patterns of Group Evolution: A Dynamic Model", in *Proc. of ACM KDD*, 2016.

[20] B. Yin, Y. Yang and W. Liu, "Exploring Social Activeness and Dynamic Interest in Community-based Recommender System," in *Proc. of ACM WWW*, pp. 771–776, 2014.

[21] F. Chung and L. Lu, "Complex Graphs and Networks," in *Providence: American Mathematical Society*, Vol. 107, 2006.

[22] [14] G. Ghoshal, L. Chi and A. Barab'asi, "Uncovering the Role of Elementary Processes in Network Evolution," in *Scientific Reports*, 2013.

[23] S. Lattanzi and D. Sivakumar, "Affiliation Networks," in *Proc. ACM STOC*, pp. 427-434, 2009.

[24] W. Aiello, F. Chung, L. Lu, "Random Evolution in Massive Graphs", in *42nd IEEE symposium on Foundations of Computer Science*, FOCS 2001.

[25] MoiveLens Dataset: http://grouplens.org/datasets/movielens/

[26] Audioscrobbler Music Artist Dataset: http://www-etud.iro.umontreal.ca/~bergstrj/audioscrobbler_data.html

[27] Yahoo Music Artist Rating Dataset: http://webscope.sandbox.yahoo.com/

[28] Book Crossing Dataset: http://www2.informatik.uni-freiburg.de/ cziegler/BX/

[29] Amazon Product Recommendation Dataset: http://jmcauley.ucsd.edu/data/amazon/links.html

[30] Technical report: https://github.com/hongruz/Evolving-RSs/blob/master/Evolving-RSs-Technical-Report.pdf