This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

Evolving Bipartite Model Reveals the Bounded Weights in Mobile Social Networks

Jiaqi Liu, Cheng Deng, Luoyi Fu, Huan Long, Xiaoying Gan, Xinbing Wang, Guihai Chen and Jun Xu

Abstract—Many realistic mobile social networks can be characterized by evolving bipartite graphs, in which dynamically added elements are divided into two entities and connected by links between these two entities, such as users and items in recommendation networks, authors and scientific topics in scholarly networks, male and female in dating social networks, etc. However, given the fact that connections between two entities are often weighted, how to mathematically model such weighted evolving bipartite relationships, along with quantitative characterizations, remains unexplored.

Motivated by this, we develop a novel Evolving Bipartite Model (EBM), which, based on empirically validated power-law distribution on multiple realistic mobile social networks, discloses that the distribution of total weights of incoming and outgoing edges in networks is determined by the weighting scale and bounded by certain ceilings and floors. Based on these theoretical results, for evolving bipartite networks whose degree follows power-law distribution, their overall weights of vertices can be predicted by EBM. To illustrate, in recommendation networks, the evaluation of items, i.e., total rating scores, can be estimated through the given bounds; in scholarly networks, the total numbers of publications under specific topics can be anticipated within a certain range; in dating social networks, the favorability of male/female can be roughly measured. Finally, we perform extensive experiments on 10 realistic datasets and a synthetic network with varying weights, i.e., rating scales, to further evaluate the performance of EBM, and experimental results demonstrate that given weighting scales, both the upper bound and the lower bound of total weights of vertices in mobile social networks can be properly predicted by the EBM.

Index Terms—Evolving bipartite model, mobile social network, theoretical bound.

1 Introduction

In mobile social networks, various sensors collect a large amount of real-time sensing data like geographical location, browsing history, user comment and rating, etc. Such a rich set of sensing data enables research [1]–[4] on modeling relationships among individuals, groups, organizations, and even entire societies. In the above relationships, there exists an important one - bipartite relationship that characterizes connections between two types of entities. For instance, as users often have the personal preferences to different items, a bipartite relationship, i.e., user preference, naturally forms. In general, user preference can be estimated by the obtained sensing data. For example, we can infer that a user located in China has a higher preference on tea. In addition, since sensing data is sampled and collected at a certain time, the preference is time-stamped. Regarding it, evolving bipartite graph is a useful tool to mathematically model such time-stamped bipartite relationships. There have existed some efforts that are directed towards analysis of evolving bipartite graphs [3], [5], [6], and several works extend the bipartite graphs to tripartite ones and even K-partite ones [7], [8]. However, few of them combine the edge weight into model, and as a result cannot characterize users' preferential level to item. While in some other related research, even though some weighted bipartite graphs [9]–[11] have been developed, most of them are static ones that are not applicable to capture evolving nature of the networks.

Motivated by this, in this work our objective is to design a model that can mathematically characterize the weighted and evolving bipartite relationships integrated from various sensing data, and analyze the properties of the model with theoretical guarantee. The main challenges of this work are summarized as follows: i) Given the fact that user preference (weight) collected and integrated from sensing data is usually complicated with multiple causes and variety, it is difficult to find a specific characterization for the weight; ii) The uncertain characterization of weight illustrated above, as well as the evolving nature of weight, makes it difficult to provide theoretical analysis on the properties of the model.

To bridge this gap, in this paper we propose a novel Evolving Bipartite Model (EBM) to incorporate the weights of edges into the whole framework. This model is founded on observed power-law distribution in bipartite social networks, and by both theoretical and experimental validation, EBM reveals that the total weights of vertices, e.g., the total rating scores items receive in recommendation networks, are limited by certain upper and lower bounds. These bounds are highly significant due to the fact that knowing the total weights of vertices helps a lot, e.g, estimate the product popularity in recommendation systems, the total number of publications under specific topics and the favorability of male/female in dating social websites. In addition, the benefit to obtain the total weights through theoretical bounds rather than directly collecting data lies in two aspects: on one hand, in some cases data is missing or incomplete; on the other hand, estimating the limits of total weights can only be achieved by theoretical bounds rather than collecting data. For example, in recommendation networks where we consider users, items and ratings as node type 1, node type 2 and weighted edges, the theoretical bound of total ratings received by an item reflects its potential performance, which contributes to product evaluation, item recommendation, marketing strategy adjustment, etc. There have existed some state-of-the-art models for bipartite networks [12]–[16]. However, to our best knowledge, EBM is the first one that reveals the rounded weights in evolving bipartite networks. To begin with, employing 10 realistic datasets, we observed both vertices' degree and total weight follow the power-law distribution in bipartite mobile social networks. Based on these observations, we propose EBM - a weighted evolving bipartite model that is inspired by the intuition of influence of the elderly, i.e., the edge formation of new vertices is often influenced by the existing vertices, especially their weight distribution, since new vertices often tend to imitate the connections of existing ones. Further, to explicitly clarify mathematical notations in EBM, we characterize EBM with notations in the context of recommendation networks and analyze it theoretically. By theoretical analyzing results and empirical validation in real networks, we reveal the total weight of a vertex in evolving bipartite networks, e.g., total rating scores an item receives, is limited by the obtained bounds.

We summarize the contributions of EBM:

- The first contribution of EBM is to originally incorporate the edge weight in model construction of evolving bipartite graph. By translating rating scores into weights of edges, EBM properly considers the weight as one important component of the model. To illustrate, when a new user/item arrives at the network, EBM supposes the weights distribution of the new arrival user/item will be exposed to the impact from weights distributions of existing users/items. Further, EBM also portrays the bound of items' total rating scores in recommendation networks, which can be leveraged to predict bounds of total weights of vertices in other social networks.
- The second contribution is to provide mathematical proofs to consolidate the model's reasonability. The theoretical depth of mathematical analysis lies in three aspects: i) We present the exact expression of weight distribution with a recurrence relation; ii) Since the recurrence relation is difficult to calculate directly, we offer its upper and lower bounds. Especially, the two bounds are the same when all the weights are set to 1 (unweighted bipartite graph), which closes up the gap between unweighted and weighted bipartite graphs; iii) Through theoretical analysis on the beginning of weight distribution, we figure out the reason of discontinuity in the distribution sequence of item weight with theoretical support.
- EBM's third contribution relies on the extensive experimental validation. In the experiments, we launch exhaustive investigations on 10 real networks that belong to 3 types: recommendation networks, scholarly networks and social networks, with varying weighting, i.e., rating scales, to further evaluate the performance of EBM. The results validate the effectiveness of the model assumption and demonstrate that knowing weighting scales, both the upper and the lower bounds of total weights of the vertices can be predicted by EBM.

The paper is organized as follows. Section 2 presents the literature review. In Section 3, we introduce our datasets and present our observation results. We illustrate our design of EBM and characterize it mathematically in Section 4. Then, we theoretically analyze EBM in Section 5. In Section 6, extensive experiments are conducted to validate the EBM. Finally, we conclude the paper in Section 7.

2 Related Work

2.1 Bipartite relationships in mobile social networks

Many existing literature works have discovered the bipartite relationships in mobile social networks, and studied their characterization, modeling and application correspondingly. Li et al. [17] study the social features through mobile phone records that can be modeled as bipartite complex networks, with one set of nodes regarded as the mobile phone users and the other one as the calls that they perform. Luo et al. [18] propose a friend discovery scheme based on character attributes, where users and their attributes like gender, age and interest naturally form the bipartite social relationships. Barbera et al. [19] collect the data of probes sent by different devices, and uncover the social relationships by modeling the bipartite social relationships between the set of devices and the set of the network SSIDs. The above works study bipartite relationships in mobile social networks from a general view. In the case of recommendation in mobile social networks, there are also a furry of related ones. De Spindler et al. [20] conduct the recommendation based on the shared information among co-located mobile devices, where users perform as one type of nodes and their locations related to a specific event are the other one. Schifanella et al. [21] define the user-item bipartite networks, based on which propose an epidemic protocol designed to distribute and exchange user ratings among mobile devices in ad-hoc networks. Del Prete et al. [22] propose diffeRS, a decentralized recommend system to reduce the complexity of Collaborative Filtering based approach in Mobile Social Networks. With the support of literature review above, we can conclude that bipartite social relationships widely exist in mobile social networks. Some researchers have made effort towards modeling bipartite social relationships, which are presented in details as below. In the following, we introduce the models that characterize bipartite relationships in mobile social networks from two aspects: network evolution and weighted relationship.

2.2 Models that characterize network evolution

There is a flurry of existing work that contribute to evolving network sciences. A series of studies [23]–[26] have illustrated the network structure evolves with the arrival and departure of users [27] and the temporary dynamics of interest [28]. Some work features the evolving process with the network model. Lattanzi et al. [3] introduce the model of affiliation networks, where preferential attachment and edge copying are used in the proposed evolving process. Liu et al. [6] give Crosslayer 2-hop Path algorithm that leverages the bipartite evolving network to study friend recommendation. Koskinen et al. [29] give a class of models and a Bayesian inference scheme that extends previous one-mode networks to two-mode ones. Hernandez et al. [13] propose an evolving model for the lodging-service network in a tourism destination, where links are determined by a random and preferential attachment rule. Preferential attachment is adopted by many models as one of their basic rules. For example, the aforementioned models [3], [6], [13] introduce this rule for edge coping. Further, there are also weighted versions of preferential attachment. Dai et al. [30] and Athreya et al. [31] assume that each new node arrives

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

TABLE 1
Performance comparison between state-of-the-art models and ours

<table>
<tr><td>Property</td><td>Network evolution</td><td>Weighted relationship</td></tr>
<tr><td rowspan="2">Algorithm</td><td>[3], [6], [13], [29]</td><td>[9]–[11], [16], [33]</td></tr>
<tr><td colspan="2">[34], Our Model</td></tr>
</table>

at the network and attaches to an existing node with the probability proportional to its sum of weights.

2.3 Models that characterize weighted relationship

A multitude of researchers have made effort to mathematically characterize weighted bipartite relationships. As one of the most important item recommendation algorithms, collaborative filtering [32] records the ratings by a matrix, in which each entry $a_{i,j}$ represents the rating of user i on item j. Besides the matrix, some work introduces the weighted bipartite graph to make the representation. Liu et al. [9] introduce a weighted bipartite user-item network to model the recommender systems, where the weight of an edge is the rating that the user gives the item. Similarly, Pan et al. [33] model the recommender systems by the weighted bipartite network, and then analyze the degree distribution in a numerical way. Hu et al. [10] and Zhang et al. [11] also use the weighted bipartite network to characterize user-item ratings and propose some new recommendation algorithms. In addition, network embedding method has been used to model weighted relationships in bipartite social networks. Gao et al. [16] develop a representation learning method for bipartite networks, and learn the weight by regarding it as one of edge features.

The aforementioned work differs from ours as follows: i) they only pay attention to either network evolution or weighted relationship while we study both the properties; ii) most of them make the evaluation numerically while we provide the theoretical bound. Last but not least, the research most related to ours may be [34]. It presents a model of weighted bipartite evolving social networks and explores the strength distributions of user/item. However, we note this model cannot be adopted to characterize user-item ratings since it assumes the weight decreases with time. A performance comparison between state-of-the-art models and ours is provided in Table 1.

3 Observation

This section launches our observation on 10 real datasets. We first introduce datasets and then illustrate our discovery, based on which EBM is founded.

3.1 Dataset

We observe 10 real datasets, including 6 recommendation datasets [35]–[37], 3 academic datasets [38] and 1 social network dataset [39]. The dataset statistics are presented in Table 2. The practical meaning of [node type 1, node type 2, weight] for recommendation networks, scholarly networks and social networks are [user, item, rating], [author, research field, number of publications that the author publishes in the research field] and [male, female, favorability from male to female]. Weight scale of each dataset is obtained in an implicit way: mapping the real ratings to integers. For instance, in Audioscrobbler dataset, we map the users' play counts for musical artist to explicit integers in $[1, 5]$. Specifically, when the play count is 1, $[2, 4]$, $[5, 9]$, $[10, 19]$, and no less than 20, the weight is mapped to 1 to 5 respectively. Besides, the table presents the number of nodes of type 1/2, the sum of all the edge weights, the number of edges, the average degree of node of type 1/2 and density that is calculated as # Edges/(# Node 1 + # Node 2).

3.2 Power-law Distributed Rating Scores

Then we will show the degree and the total weight, i.e., the number and the total weight of edges connected to the user/item, are power-law distributed. In this observation, we will figure out i) Whether the data is power-law distributed? ii) If so, how to calculate the power-law exponent α. The computation of power-law distribution uses the method provided in [40]. This work demonstrates x obeys a power law if

$$p(x) \propto x^{-\alpha}, \quad \text{for } x > x_{min},$$

where α denotes the exponent of the distribution and x_{min} denotes the lower bound to the power-law behavior. This work provides the estimation of α using maximum likelihood and that of x_{min} using Kolmogorov-Smirnov statistic. Further, it introduces p-value to quantity the goodness of fit, and the authors assert the power-law is ruled out if $p \leq 0.1$. With this method, we make the observations for degree and total weight respectively.

The results of degree distribution are given in Figure 1. From the figure we observe that the user degree and the item degree in all the 10 datasets are power-law distributed, i.e., $p > 0.1$. In particular, all these 10 datasets have a large p that is greater than 0.3, which we call strongly power-law distributed. Besides, the results of total weight distribution are given in Figure 2. As the figure shows, all the p are greater than 0.1 and thus there also exhibits the power-law distribution for total weight. The above results further provide insights for us to build EBM.

Interestingly, by observing Figures 1 and 2, we discover that the beginning of the distribution sequence shows great discontinuity. This phenomenon can be well reproduced by EBM both in theory and practice. Here, in advance, we say the beginning of distribution sequences in evolving bipartite networks, rather than being pure power-law, resembles the expected mass function in certain situations.

4 Model of EBM

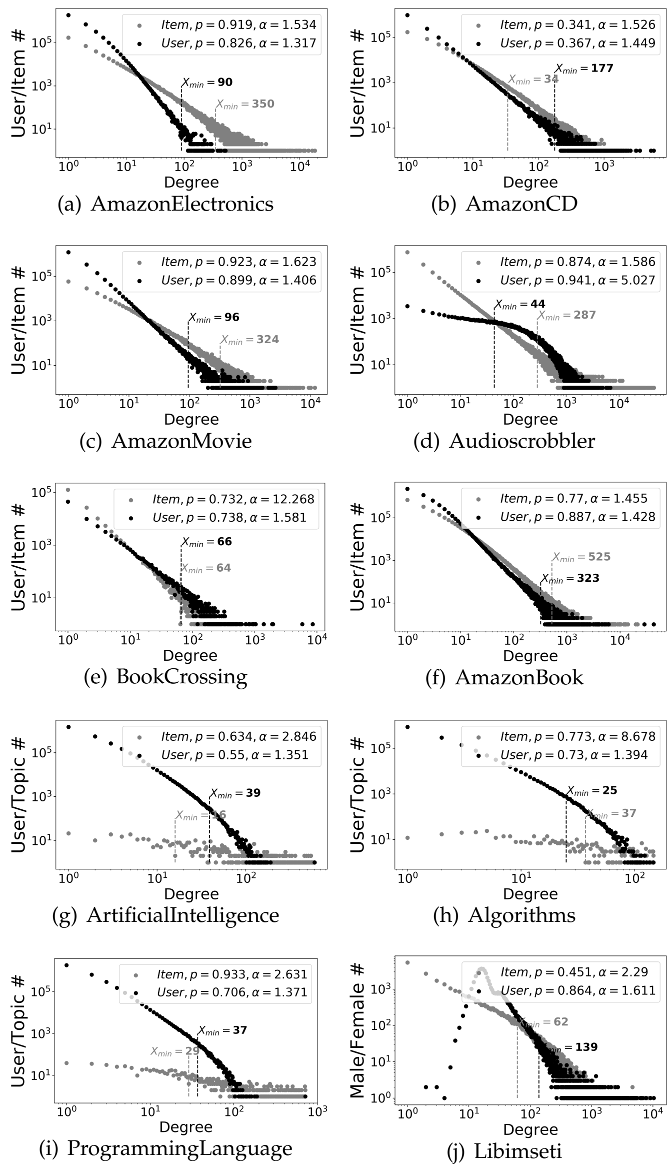
Based on the above observations, we propose evolving bipartite model (EBM). In this section, we begin with explaining the intuition behind this model. Then, we introduce EBM with its model structure, new definitions, two basic assumptions and evolving algorithm.

4.1 Intuition of EBM

The core of EBM is that we assume the behavior of new vertices are exposed to the influence of elderly, i.e., existing vertices. Here we firstly use recommendation networks as an instance, in which users and items are two sets of vertices in bipartite graph, and users' rating scores are weights of

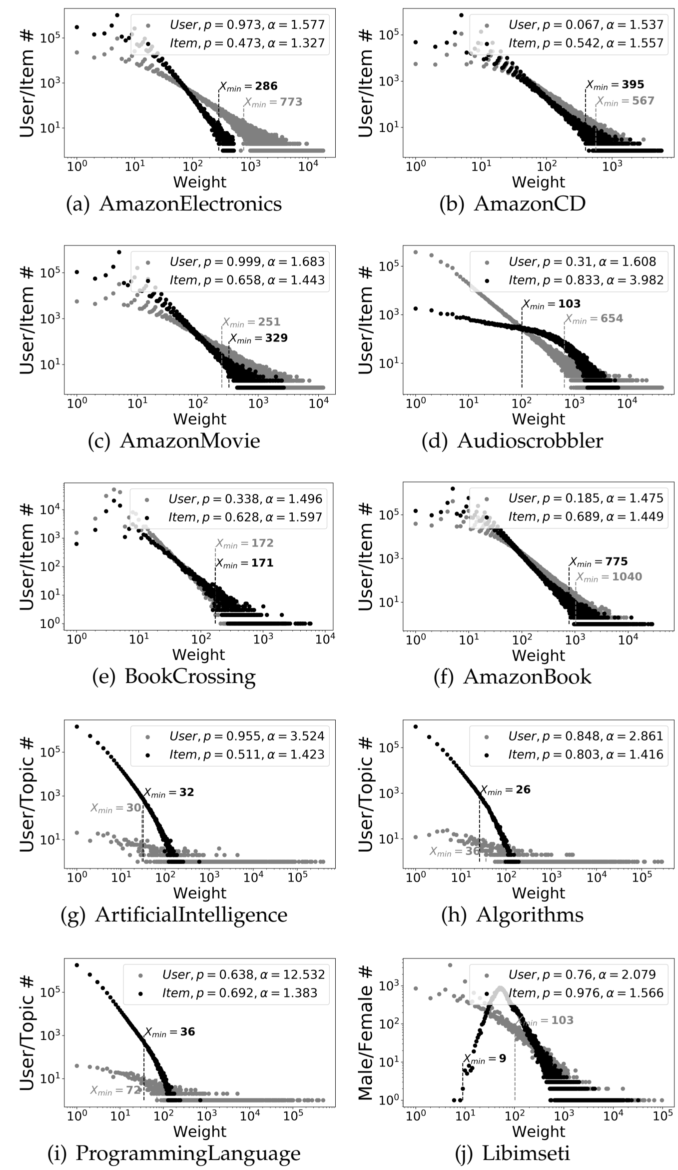
TABLE 2
Dataset statistics

Network type	Dataset	Weight scale	# Node 1	# Node 2	Total weights	# Edges	Avg. degree of node 1	Avg. degree of node 2	Density
Recommendation networks	AmazonMovie	[1,5]	2,088,620	200,940	19,289,148	4,607,047	2.21	22.93	2.01
	AmazonCD	[1,5]	1,578,596	486,360	16,505,353	3,749,004	2.37	7.71	1.82
	Audioscrobbler	[1,5]	146,946	1,493,930	59,353,961	24,296,858	165.35	16.26	14.81
	AmazonBook	[1,5]	8,026,324	2,330,066	96,685,308	22,507,155	2.80	9.66	2.17
	BookCrossing	[1,5]	77,803	185,846	1,749,539	433,669	5.57	2.33	1.64
	Amazon Electronics	[1,5]	4,201,696	476,001	31,394,452	7,824,482	1.86	16.44	1.67
Scholarly networks	Artificial Intelligence	[1,5]	3,276,869	1,130	7,469,417	7,073,161	2.16	6,259.43	2.16
	Algorithms	[1,5]	1,686,210	1,092	4,162,485	3,895,052	2.31	3,566.89	2.31
	Programming Language	[1,5]	3,957,282	1,691	7,891,791	7,499,261	1.89	4,434.81	1.89
Social networks	Libimseti	[1,5]	60,145	38,433	9,670,628	3,232,064	53.73	84.09	32.78



(a) AmazonElectronics (b) AmazonCD (c) AmazonMovie (d) Audioscrobbler (e) BookCrossing (f) AmazonBook (g) ArtificialIntelligence (h) Algorithms (i) ProgrammingLanguage (j) Libimseti

Fig. 1. Degree distribution of users (grey dot) and items (black dot).



(a) AmazonElectronics (b) AmazonCD (c) AmazonMovie (d) Audioscrobbler (e) BookCrossing (f) AmazonBook (g) ArtificialIntelligence (h) Algorithms (i) ProgrammingLanguage (j) Libimseti

Fig. 2. Total weight distribution of users (grey dot) and items (black dot).

edges in the graph. In the perspective of users, a new user is likely to learn from rating scores given by an influential elderly user, who is thus selected as a prototype and impacts the new user on launching rating scores, i.e., weights. In the perspective of items, a new item that yearns for success is likely to imitate an outstanding item, i.e., the item with high total rating score. Therefore, since this outstanding elderly item is chosen to be prototype, the rating score of the new item is, to some extent, allied with the score of its prototype. In the example of scholarly networks, authors and topics are represented as vertices while the numbers of publications are represented as edge weights. In the case of a new author, he/she may learn from the elderly authors who have large amount of publications. In the case of a topic, it is likely to be a branch new of an existed topic in which massive papers have been published. As for social networks, user and topics can be represented as two sets of vertices while the number of messages that a user posted to a topic is represented as the weight. A new user often pays more attention to the hot topics while a newly generated topic prefers to attract those influential users. The above examples show the universal existence of influence of the elderly. Based on this intuition, with the constant arrival of new vertices and the weights of edges, a virtuous cycle will be normally resulted:

- A newly arrived vertex often has preference on popular vertices, i.e., the elements with high total weights.
- Preference brings new weights into popular vertices, which further strengthens their popularity.
- As the bipartite graph evolves, popular vertices become more popular while outmoded ones gradually become neglected, which contributes to final power-law distribution in the network.

Based on these intuitions, we build EBM. In the rest parts, EBM is illustrated in details and examined by both theoretical and experimental analysis, which, in reverse, validates that the intuition is consistent with network evolving nature.

4.2 Mathematical Characterization

In order to simplify mathematical symbols and endow our model with more concrete meaning, we characterize EBM under a specific case, i.e., recommendation networks. Here users and items can be extended to vertices in other bipartite social networks which yielded to power-law distribution, and rating scores can be extended to edge weights. For example, users and items can be extended to authors and research fields in scholarly networks, as well as users and topics in social networks. Rating scores can be extended to number of publications that the author publishes in the research field in scholarly networks, and number of messages that a user posted to a topic in social networks.

Model Structure: The network structure is characterized by a weighted bipartite graph $B(U, I)$. Vertices in U represent users and vertices in I represent items in recommendation networks. Intuitively, an edge between user u and item i indicates that user u makes a purchase and gives a rating on item i. The edge has weight $w_{(u,i)}$, which denotes the rating score given to item i by user u.

Definition of Vertex Weight: Given an arbitrary vertex v in $B(U, I)$, let $N(v)$ be the set of vertices connected to v. The vertex weight of v is the sum of edge weights on all edges connected to vertex v, namely,

$$W(v) = \sum_{t \in N(v)} w_{(v,t)}.$$

The vertex weight is the total rating score a user gives or an item receives in recommendation networks.

Basic Assumptions: In realistic mobile social networks, the value of weight is jointly determined by the user and the item. For example, in recommendation systems, for a certain item, some kind users prefer to rate a high score while some critical users prefer to rate a low score. Similarly, for a certain user, items with high quality often receive high scores while those with low quality often receive low scores. In order to model the influence of users' and items' own character, we divide users to K types and items to L types. For the case that users (items) follow a single type, we set $K = 1$ ($L = 1$). For the case that users (items) follow a mixture of several basic types, we set $K > 1$ ($L > 1$). For example, in recommendation systems, a user is often not purely kind or purely critical, but a mixture of the two types with a weight vector like 70% kind and 30% critical, Then, the mathematical modeling of this mixture distribution is given as below.

(1) Rating scores are sampled from a mixture distribution.
(2) There are K user types and the rating preference of user u is a mixture of the K types with weight vector v_u.
(3) There are L item levels and the rating preference of item i is a mixture of the L levels with weight vector t_i.
(4) The distribution of users' weight vectors is denoted by F (possibly Gaussian) with a parameter vector θ, and the probability density function (pdf) is denoted by f_θ.
(5) The distribution of items' weight vectors is denoted by G (possibly Gaussian) with a parameter vector γ, and the pdf is denoted by g_γ.
(6) The expected value of rating scores is denoted by Er.

In some cases, the researchers only focus on the influence coming from users or items. Then, the following framework (the first framework) can be adopted.

(1) For any user type k, there is a unique rating probability mass function (pmf) $h_k(r)$, $r = 1, 2, ..., R$. Symmetrically for any item level l, there is a unique 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:
$$H_u(r) = \sum_{k=1}^{K} v_u(k) h_k(r).$$
The rating pmf of a newly added item i is
$$H^i(r) = \sum_{l=1}^{L} t_i(l) h^l(r).$$

In the other cases, the researchers focus on the influence coming from both users and items. Then, the following framework (the second framework) can be adopted.

(1) Given any pairs of the user type k and the item level l, there is a unique rating pmf $h_k^l(r)$, $r = 1, 2, ..., R$.
(2) When user u gives his/her rating on item i, the rating pmf can be presented as:
$$H_u^i(r) = \sum_{k=1}^{K} \sum_{l=1}^{L} v_u(k) t_i(l) h_k^l(r).$$

We note $H_u(r)$, $H^i(r)$ and $H_u^i(r)$ are discrete with respect to user type k, item level l and rating r, but continuous with respect to θ and γ - parameters of distributions of v_u and t_i.

We give experimental validation on this mixture distribution model. Results in Section 6.1 demonstrate the model

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

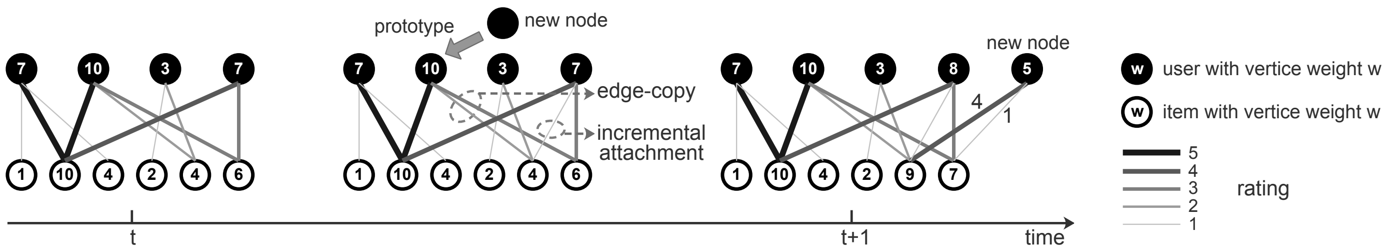


Fig. 3. An illustration on how the weighted bipartite network evolves.

can well characterize weight distribution in realistic mobile social networks.

In fact, weight vector v_u characterizes the user u's own preference on rating, which is jointly determined by

1) Internal Effects: emotional and psychological factors like subjective norms, social identity and group norms;
2) External Effects: social networks' structural parameters like clustering, centrality, and assortative mixing.

For example, assume there are two user types: mild character and demanding character, with rating pmfs $h_1(r)$ and $h_2(r)$ respectively. $h_1(r)$ increases ($h_2(r)$ decreases) with r since a kind (demanding) user prefers to rate a high (low) score. Internal effects reflect a user's own social character. A person with a mild character often prefers to rate a high score, and therefore the vector v_u provides a high weight to $h_1(r)$ and a low one to $h_2(r)$. As for external effects like clustering, the users within a same cluster often have some common preferences, and therefore their vectors are similar to some degree. Note that the social pattern on rating can be characterized by weight vector v_u and user type pmf $h_k(r)$, which well capture social aspects in mobile social networks.

In addition, we note that our model may not perfectly characterizes the social influence in some scenarios. However, a complete reproduction of all the social features is too difficult, if at all possible, and we believe it is beneficial to make some simplifications towards a tractable model and a meaningful look into the formation of weighted bipartite relationships. In this way, the proposed model well captures social influence in real-word mobile social networks.

Evolving Algorithm: As Figure 3 illustrates, the evolving process in EBM incorporates two symmetrical aspects, i.e., the arrival of new users and items. Here we use the arrival of a new user as an instance. At each time step, a new user arrives, selects an existing user (someone who possibly shares common interests with him/her and has a high total rating score) as the prototype. Then, the new user chooses among the items purchased by the prototype with a probability proportional to the ratings 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 newly added items. Algorithm 1 presents the details of the evolving process in $B(U, I)$. In addition, some edges will be created between the existing users and item, referred as incremental attachment.

Remarks: EBM assumes the weights are positive integers. While in real recommendation networks or other social networks, there exist decimal rating scores like 1.5 stars or decimal values of weights. However, this does not hurt since we can always map decimal or implicit ratings into explicit positive integer ratings based on pre-established rules as in subsection 3.1. For convenience, we present Table 3 to list all the notations that will be used later.

Algorithm 1: Rating-Driven Evolution in $B(U, I)$

Fix the integers c, c_u, $c_i > 0$, and let $\beta \in (0, 1)$.
Fix an integer R as the highest rating score.
Initialization at $t = 0$:
Give a weighted bipartite graph $B(U, I)$ with at least $c_u c_i$ edges, where each $u \in U$ has at least c_u edges and each $i \in I$ has at least c_i edges. The edge weights are sampled from the rating set $\{1, 2, \ldots, R\}$ following the mixture distribution with expectation Er.
At time $t > 0$:
begin
  (Evolution of U) With probability β:
  begin
    (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, namely the sum of edge weights of all edges connected to it.
    (Edge-copy) c_u edges are copied from u', i.e, c_u neighbors of u', denoted by $n_1, n_2, ..., n_{c_u}$ are chosen with a probability proportional to the weight of edges (without replacement). Edges $(u, n_1), (u, n_2), ..., (u, n_{c_u})$ are added to the graph with weights sampled from the set $\{1, 2, \ldots, R\}$ following the mixture distribution.
  end
  (Evolution of I) With probability $1 - \beta$, a new item i is added to I following a symmetrical process, adding c_i edges to i.
  (Incremental attachment) c new edges are added to $B(U, I)$, where the edge's two endpoints are selected from U and I with the probability proportional to the node's weight. Similarly, the weight is sampled following the mixture distribution.
end

4.3 Discussion

Rather than creating networks similar to certain real ones, the purpose of EBM is to figure out the inherent evolution mechanism of weighted bipartite networks, based on which deduce the theoretical bounds of user/item weight distribution. The model contributes to two aspects: (i) The study on inherent mechanism can help researchers better

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

TABLE 3
Notation and Definition

Notation	Definition
$B(U, I)$	Weighted bipartite graph.
$w_{(u,i)}$	Rating score given to item i by user u.
$W(v)$	Sum of weights on edges connected to node v.
W_t	Total edge weight of users in $B(U, I)$ at time T.
K	Number of user types.
v_u	Weight vector of user u.
f_θ	PDF of v_u with the parameter vector θ.
L	Number of item types.
t_i	Weight vector of item i.
g_γ	PDF of t_i with the parameter vector γ.
$h_k(r)$	Pmf of the user of type k giving the rating r.
$h^l(r)$	Pmf of the item of type l receiving the rating r.
$h_k^l(r)$	Pmf of the user of type k giving the rating r to the item of type l.
$H(r)$	Expected rating pmf for a new user.
Er	Global expectation of rating scores.
c_u	Number of edges attached to a new user.
c_i	Number of edges attached to a new item.
β	Probability of user arrival in each time slot.

understand the essence of weighted evolving bipartite relationships; (ii) The deduction on theoretical bounds is helpful in estimating user/item weight and reveal the performance limit of them.

5 Theoretical Analysis of EBM

We present the theoretical analysis of EBM in this section. Firstly, Theorem 1 demonstrates the degree in EBM is power-law distributed. Then, Theorems 2 and 3 give power-law bounds for weight distribution, and finally, Proposition 3 and Theorem 4 analyze the beginning of it.

5.1 Analysis of Degree Distribution

We start with the degree distribution in EBM and give Lemma 1 first.

Lemma 1. *[41] 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 approaches similar to [3], we derive Theorem 1.

Theorem 1. *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 exponent $\alpha = -2 - \frac{c_u\beta}{c_i(1-\beta)+c}$ $\left(\alpha = -2 - \frac{c_i(1-\beta)}{c_u\beta+c}\right)$.*

Proof. Please refer to the Supplemental Material. □

5.2 Analysis of Weight Distribution

In the remaining part we first give the upper and lower bounds of vertex weight distributions in Theorems 2 and 3, which is further illustrated in Figure 4. Then we discuss the beginning of vertex weight distributions in Theorem 4.

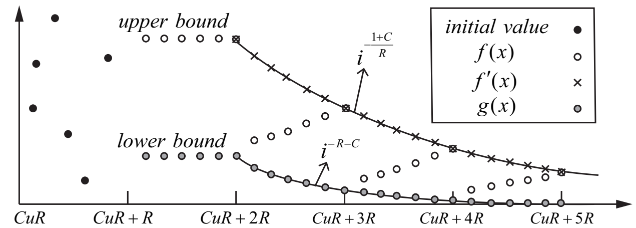


Fig. 4. An illustration on upper and lower bounds of the vertex weight distribution.

5.2.1 Upper Bound of Weight Distribution

Theorem 2. *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 $Er\left(1 + \frac{c_u\beta}{c_i(1-\beta)+c}\right)$ $\left(Er\left(1 + \frac{c_i(1-\beta)}{c_u\beta+c}\right)\right)$ for any vertex with vertex weight greater than $c_uR + 2R$ ($c_iR + 2R$).*

Proof. The proof of Theorem 2 includes 3 sequential parts:
I. Derivation of the recurrence relation.
II. Derivation of the intermediate upper bound.
III. Derivation of the power-law upper bound.

$$PART \text{I}. Derivation of the Recurrence Relation$$

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

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

The vertex weight of vertices in U can increase if a new vertex is added to I, or incremental attachment happens to it. Thus we have

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

$W_{t-1} + W_{B_0}$ denotes the total weight of users at time $t-1$, where W_{B_0} is the total weight at time 0 (the initial graph) and W_{t-1} is the total weight increased from time 0 to time $t-1$. Similar to the proof in Theorem 1, the probability that an edge is selected is proportional to its weight. Thus,

$$\begin{aligned} & E[\text{\# vertices with vertex weight } j \text{ at time } t-1 \text{ that are} \\ & \text{chosen as end point}] \\ = & \frac{jV_{t-1}^j}{W_{t-1} + W_{B_0}}. \end{aligned}$$

Since we have

$$\begin{aligned} & E[\text{\# vertices in } U \text{ whose vertex weight is } j \text{ at time } t-1 \\ & \text{and increases to } k \text{ at time } t] \\ = & E[\text{\# vertices in } V_{t-1}^j \text{ that is chosen as end points}] \\ & \times E\left[Pr[\text{the edge weight is assigned to be } k-j]\right], \end{aligned}$$

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

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

In the second framework, we have:

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

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

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

Let $X_k = \lim_{t\to\infty} V_t^k / t$. Again, using Lemma 1, we get the following recurrence relation

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

namely,

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

The value of this R-order homogeneous recurrence relation is hard to calculate since it is jointly determined by its R previous values, i.e., $X_{k-1}, X_{k-2}, ..., X_{k-R}$. Thus, instead of direct computation, we present its upper bound as below and its lower bound in Theorem 3 to make the estimation.

$PART \text{II}. Derivation of the Intermediate Upper Bound$

Without loss of generality, we assume the initial values 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

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

namely,

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

Before the induction, 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 $c_uR + 2R + 1 \le k \le c_uR + 3R$:

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

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

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

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

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

$PART \text{III}. Derivation of the Power-law Upper Bound$

We can derive the following relation

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

for every $k = c_uR+3R, c_uR+4R, \ldots$ Recursively, the above

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

equation results in

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

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

As illustrated in Figure 4, 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 $[c_u R+nR+1, c_u R+(n+1)R]$ for any positive integer n while $f'(k)$ is decreasing and $f(c_u R+nR) = f'(c_u R+nR)$. As shown in Figure 4, $f'(k)$ 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. □

5.2.2 Lower Bound of Vertex Weight Distribution

Theorem 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 a lower bound which follows a power-law distribution with exponent $\alpha = -R-C$, where C is a constant and equals $Er(1+\frac{c_u\beta}{c_i(1-\beta)+c})\left(Er\left(1+\frac{c_i(1-\beta)}{c_u\beta+c}\right)\right)$ for any vertex weight greater than $c_u R+2R$ ($c_i R+2R$).*

Proof. Please refer to the Supplemental Material. □

Theorem 1 is consistent with Theorems 2 and 3 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 graph.

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

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

5.2.3 Beginning of Vertex Weight Distribution

The above two sections have shown the vertex weight distribution is bounded by power-laws when vertex weight k is greater than a certain value. Here we proceed to explore the beginning of the 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 a new random variable S to be the vertex weight of a newly added vertex in U with a pmf $s(k)$. Symmetrically we define another 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 2 and 3, 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 providing Theorem 4. The first two propositions give the recurrence relation and an upper bound of X_k. The third one describes the distributions of random variables S and S'.

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

Proof. Please refer to the Supplemental Material. □

Proposition 2. *For $c_u \le k \le c_u R$ (resp. in I, $c_i \le k \le c_i R$), X_k have an upper bound $l(k) \le 1$.*

Proof. Please refer to the Supplemental Material. □

Proposition 3. *When c_u (resp. c_i) is large enough, S (S') follows a unimodal probability distribution.*

Proof. Please refer to the Supplemental Material. □

Now we introduce Theorem 4 that regards the difference of vertex weight sequence and the difference of the non-homogeneous term in the recurrence relation.

Theorem 4. *Let $p(k)$ be the non-homogeneous 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

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

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

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

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

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

Proof. Please refer to the Supplemental Material. □

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

Here we discuss more about function $p(\cdot)$ 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 and then decreases toward 0 as k increases. By Proposition 3 and Theorem 4, we can roughly depict the vertex weight distribution of items for $c_i \leq k \leq c_i R$ and it is very 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 the expected rating pmf for new items $H'(r)$. Thus the shapes 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 that in U.

5.3 Discussion on the Parameters

In this subsection, we will discuss how the parameters affect the derived upper and lower bounds. According to Theorem 1 and Theorem 3, we have

- Weight scale R: A larger R gives a greater upper bound and a smaller lower bound, which indicates the weight distribution is difficult to be estimated.
- The probability of user arrival in each time slot β: A larger β results in a greater C for user and a smaller C for item, and therefore generates the greater upper bound and lower bound for user as well as the smaller upper bound and lower bound for item.
- Number of edges attached to a new user c_u: A larger c_u results in a greater C for user and a smaller C for item, and consequently has the same effects as the parameter β.
- Number of edges attached to a new item c_i: A larger c_i results in a smaller C for user and a greater C for item, and therefore has the opposite effects as the parameter β.
- Parameter of incremental attachment c: A larger c results in a smaller C for both user and item, and thus generates the smaller upper bound and lower bound.

6 Experiments

The experiments are launched on the real mobile social bipartite networks presented in Table 4. The experiments include two parts: i) validation on EBM's ability to model real-world bipartite networks; ii) validation on EBM's other properties. This section presents experimental results on item, which also hold for user due to their symmetrical characteristic. In addition, the experiments are reproducible with the support of datasets and codes[1].

1. https://github.com/davendw49/ebm

6.1 Validation on EBM's ability to model real-world bipartite networks

6.1.1 Validation on model assumption

To begin with, we come to the justification on the need of mixture model assumption on recommendation ratings given in Section 4.2, i.e., recommendation rating pmf is a mixture of certain basic types. In the two proposed frameworks, when we fix the number of user type K, the rating pmf of item i is

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

When $L = 1$, we call the pmf a single one; when $L > 1$, we call it a mixture one. The recommendation rating shows the mixture characteristic in many real networks. To validate that, we set $L = 2$ and take the case where $L = 1$ as the baseline. Then, the rating pmfs of the mixture case and the baseline are

- Mixture case: $H^i(r) = t_i(1)h^1(r) + t_i(2)h^2(r)$.
- Baseline: $H^i(r) = t_i h(r)$.

We use Mean-Square Error (MSE) as the metric to measure the two models' ability on characterizing rating distribution. Let $p(x)$ denote the real pmf obtained from the recommendation dataset and $\hat{p}(x)$ denote the fitting one given by the model. The MSE between the real pmf and the fitting pmf is calculated as

$$MSE = E\left(\hat{p}(x) - p(x)\right)^2.$$

Obviously, a smaller MSE indicates a better fitting.

In the experiments, we assume $h^i(r)$ is a Gaussian distribution, i.e., $h^i(r) = N(\mu, \sigma^2)$, since it is widely accepted that the rating is Gaussian distributed. The fitting parameters σ and μ are obtained using the following methods. For the baseline, following the definition of Gaussian distribution, we set μ as the mean of the real data and σ^2 as the variance. For the mixture case, μ_1, σ_1, μ_2 and σ_2 for $h^1(r)$ and $h^2(r)$ are set empirically. The weight $t_i(1)$ is selected from 0.1 to 0.9, and $t_i(2)$ is determined since $t_i(1) + t_i(2) = 1$.

Figure 5 shows the results. In all the 10 bipartite mobile social datasets, MSE of the mixture case is smaller than that of the baseline, which indicates the mixture assumption can better characterize the rating distribution. Besides, we note the mixture case only shows a weak outperformance since the parameter setting for the baseline is almost optimal, but that for the mixture case is far from optimal. In specific, for the baseline, we set μ and σ as their optimal estimation. For the mixture case, there are too may parameters, i.e., $2L+nL$, which are difficult to find the optimal combination for them and thus we make the parameter setting empirically.

6.1.2 Weight distribution

In this part, we give the evaluation on weight distribution. To begin with, we will evaluate whether the weight is limited by the given bounds. The real weight distribution, the upper and lower bounds given by the proposed model are calculated as follows.

- Real weight distribution: the method in [40];
- Upper bound (EBM): $R + C$;
- Lower bound (EBM): $(1 + C)/R$,

where R is the weight scale that can be directly used, and $C = Er\left(1 + \frac{c_i(1-\beta)}{c_u\beta+c}\right)$ can be estimated from the data. In

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing

TABLE 4
Upper and lower bounds of item weight distribution

Dataset	C	R	-α	UpperBound (EBM)	LowerBound (EBM)	UpperBound (Baseline 1)	LowerBound (Baseline 1)	UpperBound (Baseline 2)	LowerBound (Baseline 2)
AmazonMovie	4.187	5	1.443	9.187	1.037	5.187	5.187	U	L
AmazonCD	4.403	5	1.557	9.403	1.081	5.403	5.403	U	L
Audioscrobbler	2.443	5	3.982	7.443	0.689	3.443	3.443	U	L
AmazonBook	4.296	5	1.449	9.296	1.059	5.296	5.296	U	L
BookCrossing	4.035	5	1.597	9.035	1.007	5.035	5.035	U	L
AmazonElectronics	4.012	5	1.327	9.012	1.002	5.012	5.012	U	L
ArtificialIntelligence	1.056	5	1.423	6.056	0.411	2.056	2.056	U	L
Algorithms	1.069	5	1.069	6.416	0.414	2.069	2.069	U	L
Programming Language	1.052	5	1.383	6.052	0.410	2.052	2.052	U	L
Libimseti	4.041	5	1.566	9.041	1.008	5.041	5.041	U	L

* $U > u, \forall u \in (0, \infty)$, $L < l, \forall l \in (0, \infty)$

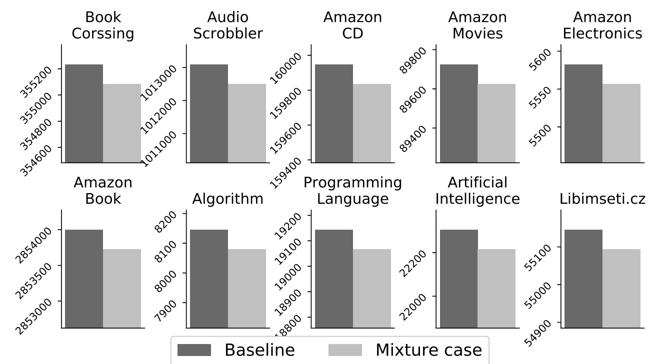


Fig. 5. MSE of rating pmf under the mixture case and the baseline (real-world networks).

addition to the proposed model, we include two baselines: Baseline 1: Bounds relying on network evolution only. In this baseline, the bound calculation relies only on network evolution but neglects weighted relationships. It indicates that the weighted bipartite network degenerates to a unweighted one. Without loss of generality, all the weights can be regarded as 1, i.e., $R = 1$. Then, we have

- Upper bound (network evolution only): $1 + C$;
- Lower bound (network evolution only): $1 + C$.

Baseline 2: Bounds relying on weighted relationships only. In this baseline, the bound calculation relies only on weighted relationships but neglects network evolution. It indicates that we only have the value of R and in this case the weight can follow an arbitrary power-law distribution, I.e.,

- Upper bound (weighted relationships only): U, where $U > u, \forall u \in (0, \infty)$;
- Lower bound (weighted relationships only): L, where $L < l, \forall l \in (0, \infty)$.

The values of R and C in the 10 datasets, along with the bounds given by the proposed model and two baselines, are listed in Table 4. From the table we can see that the item weight distribution is well bounded by the proposed model. For the baseline 1 whose bounds relies on network evolution only, the upper bound equals to the lower bound due to the loss of weight information, which indicates that weight information is necessary in order to obtain the correct bounds. For the baseline 2 whose bounds relies on weighted relationships only, the upper bound can be arbitrary large while the lower bound can be arbitrary small, which is the correct bounds but not the effective one. Therefore, we conclude that both network evolution and weighted relationships are important in the derivation of weight bounds.

Subsequently, we look into the beginning of item weight distribution. No plot in Figure 2 exhibits a consecutive curve at the beginning, which indicates c_i is small for the selected datasets according to Proposition 3 and Theorem 4. With a small c_i, item weight distribution tends to have the similar fluctuations to those in expected rating pmf for items $H'(r)$. We show item weight distribution in book recommendation datasets, i.e., BookCrossing and AmazonBook, in Figure 6. Figures 6(a) and 6(b) show that the beginning of item weight distribution resembles $H'(r)$, suggesting the average users' rating habits has influence on the beginning of item weight distribution. In an evolutional view, small c_i leads to a large number of items having small vertex weights. New items are added with small vertex weights and are not competitive with other existing popular items. Users choose items with probabilities proportional to edge weights, indicating that popular items with large vertex weights are more likely to be selected by users and have higher weights in the future.

Moreover, we find that in Figure 6(b), 5 different power-law curves seem to coexist and the logarithm plot from the same dataset in Figure 6(c) also supports this idea. With a detailed examination on the dataset, we present the rating pmf for different scores in Figure 6 and find the users prefer to give a rating score of 5. Recall the recurrence relation in proofs of Theorems 2 and 3

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

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

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

which indicates 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.

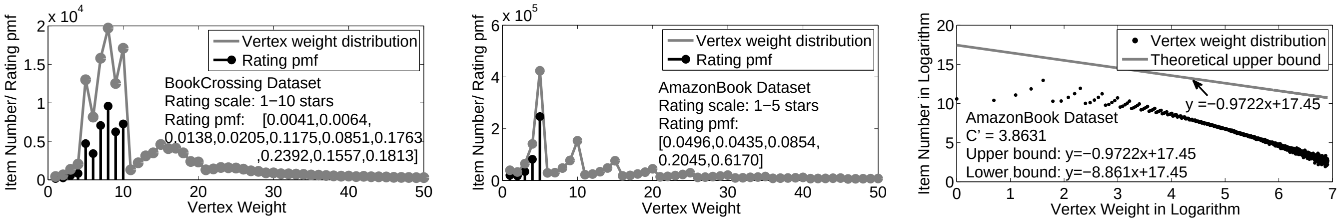
6.2 Validation on EBM's other properties

In addition to the experiments of EBM's performance on deriving upper and lower bounds, we provide some other experiments to show more features of the model.

6.2.1 Experiments on real-word networks

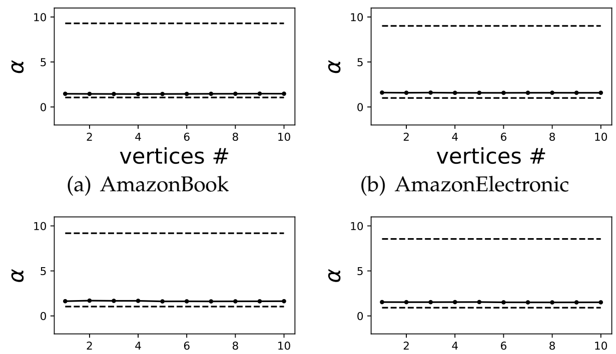
Firstly, we show how the weight distribution changes as new nodes are added, and whether the given bounds

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2020.3017630, IEEE Transactions on Mobile Computing



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

Fig. 6. The beginning of item weight distribution in book recommendation datasets (real-world networks).



(a) AmazonBook (b) AmazonElectronic (c) AmazonMovie (d) AmazonCD

Fig. 7. Item weight distribution with the vertices arrival (real-world networks).

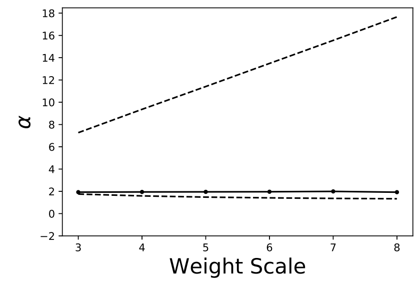


Fig. 8. Item weight distribution with the increase of weight scale (synthetic networks).

still hold in this process. The experiment is proceeded on the four real recommendation datasets. The corresponding statistics are presented in Table 2. The remaining two datasets are not used here since they do not include time information. In the datasets, each rating is related with a timestamp that records the time when the user rates the item, according to which we can obtain the time when an item is added, i.e., the time when an item receives its first rating. Using the same method as in Section 3, we assert the item weight degree are power-law distributed in the 4 real datasets with $p > 0.1$. Next, we calculate the exponents α of the power-law distributions following a similar way and show how they change as new nodes are added. The results are given in Figure 7. It shows that the weight distribution almost does not change when new vertices are added, and the weight distribution is well limited by the two bounds in the whole evolution.

6.2.2 *Experiments on synthetic networks*

Then, we show how the derived weight bounds change with the network structure. The experiments are conducted on the synthetic networks. In the evaluation, the parameters are set as $\beta = 0.6$, $c_u = 10$, $c_i = 20$ and $R \in \{3, 4, 5, 6, 7, 8\}$. The experimental results regarding the synthetic networks with specific characteristics are given in Figure 8. From the result we can observe that the item weight distribution, characterized as the power-law distribution with the exponent α, fluctuates with R, while the upper bound increases with R and the lower bound decreases with R. The intuition behind this observation lies in that a larger R allows a larger range of item weight and thus leads to a loose limit on its upper and lower bounds.

7 Conclusion and Future Work

In this paper, based on ten real datasets, we start with observing several interesting distributions in mobile social networks. Based on observations, we propose a novel evolving bipartite model called EBM that highlights the establishment of social connections for new vertices and the characterization of their behaviors based on weighting-driven preferential attachment. The superiority of our model lies in three aspects: good capture of realistic networks, mathematical tractability and novelty in predicting the bounds of final weights of connections. In the case study of recommendation networks, we also investigate the beginning of the item weight distribution, which resembles the expected rating pmf for new items.

However, there are also some constraints in the proposed model. For example, the model assumes that the user preferences, i.e., the weights, are independent of each other, which cannot exactly characterize the real case since a user's preference may be influenced by his, or her, friends. Therefore, it is a desirable future work to study the characterization of this relevance with weighted evolving bipartite graph.

References

[1] John Scott. *Social network analysis*. Sage, 2017.
[2] Elena Zheleva, Hossam Sharara, and Lise Getoor. Co-evolution of social and affiliation networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1007–1016. ACM, 2009.
[3] Silvio Lattanzi and D Sivakumar. Affiliation networks. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 427–434. ACM, 2009.
[4] Mark EJ Newman. Coauthorship networks and patterns of scientific collaboration. *Proceedings of the national academy of sciences*, 101(suppl 1):5200–5205, 2004.
[5] Nesserine Benchettara, Rushed Kanawati, and Celine Rouveirol. Supervised machine learning applied to link prediction in bipartite social networks. In *Advances in Social Networks Analysis and Mining (ASONAM), 2010 International Conference on*, pages 326–330. IEEE, 2010.

[6] Jiaqi Liu, Qi Lian, Luoyi Fu, and Xinbing Wang. Who to connect to? joint recommendations in cross-layer social networks. In *to appear in IEEE INFOCOM 2018*.

[7] Lingkun Kong, Fengyu Deng, Jiaqi Liu, Jinghao Zhao, Luoyi Fu, and Xinbing Wang. Msm: A multi-entity scholarly model for systematic understanding of evolving scholarly networks. ACM, 2017.

[8] Jiaqi Liu, Luoyi Fu, Yuhang Yao, Xinzhe Fu, Xinbing Wang, and Guihai Chen. Modeling, analysis and validation of evolving networks with hybrid interactions. *IEEE/ACM Transactions on Networking*, 27(1):126 – 142, 2019.

[9] Liu Jie, Mingsheng Shang, and Duanbing Chen. Personal recommendation based on weighted bipartite networks. In *International Conference on Fuzzy Systems & Knowledge Discovery*, 2009.

[10] Xiaohui Hu, Zichao Mai, Haolan Zhang, Xue Yun, Weixin Zhou, and Chen Xin. A hybrid recommendation model based on weighted bipartite graph and collaborative filtering. In *IEEE/WIC/ACM International Conference on Web Intelligence Workshops*, 2017.

[11] Xin Meng Zhang and Sheng Yi Jiang. Personalized recommendation algorithm based on weighted bipartite network. *Journal of Computer Applications*, 32(3):654–657, 2012.

[12] Sinan G Aksoy, Tamara G Kolda, and Ali Pinar. Measuring and modeling bipartite graphs with community structure. *Journal of Complex Networks*, 5(4):581–603, 2017.

[13] Juan M Hernández and Christian González-Martel. An evolving model for the lodging-service network in a tourism destination. *Physica A: Statistical Mechanics and its Applications*, 482:296–307, 2017.

[14] Isabella Gollini. A mixture model approach for clustering bipartite networks. In *Challenges in Social Network Research*, pages 79–91. Springer, 2020.

[15] Yaping Wang, Pengfei Jiao, Wenjun Wang, Chunyu Lu, Hongtao Liu, and Bo Wang. Bipartite network embedding via effective integration of explicit and implicit relations. In *International Conference on Database Systems for Advanced Applications*, pages 435–451. Springer, 2019.

[16] Ming Gao, Leihui Chen, Xiangnan He, and Aoying Zhou. Bine: Bipartite network embedding. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 715–724, 2018.

[17] Ming-Xia Li, Vasyl Palchykov, Zhi-Qiang Jiang, Kimmo Kaski, János Kertész, Salvatore Miccichè, Michele Tumminello, Wei-Xing Zhou, and Rosario N Mantegna. Statistically validated mobile communication networks: the evolution of motifs in european and chinese data. *New Journal of Physics*, 16(8):083038, 2014.

[18] Entao Luo, Qin Liu, and Guojun Wang. Hierarchical multi-authority and attribute-based encryption friend discovery scheme in mobile social networks. *IEEE Communications Letters*, 20(9):1772–1775, 2016.

[19] Marco V Barbera, Alessandro Epasto, Alessandro Mei, Vasile C Perta, and Julinda Stefa. Signals from the crowd: uncovering social relationships through smartphone probes. In *Proceedings of the 2013 conference on Internet measurement conference*, pages 265–276, 2013.

[20] Alexandre De Spindler, Moira C Norrie, and Michael Grossniklaus. Collaborative filtering based on opportunistic information sharing in mobile ad-hoc networks. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*, pages 408–416. Springer, 2007.

[21] Rossano Schifanella, André Panisson, Cristina Gena, and Giancarlo Ruffo. Mobhinter: epidemic collaborative filtering and self-organization in mobile ad-hoc networks. In *Proceedings of the 2008 ACM conference on Recommender systems*, pages 27–34, 2008.

[22] Lucia Del Prete and Licia Capra. differs: A mobile recommender service. In *2010 Eleventh International Conference on Mobile Data Management*, pages 21–26. IEEE, 2010.

[23] Martin Atzmueller, Andreas Ernst, Friedrich Krebs, Christoph Scholz, and Gerd Stumme. On the evolution of social groups during coffee breaks. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 631–636. ACM, 2014.

[24] Luoyi Fu, Jinbei Zhang, and Xinbing Wang. Evolution-cast: Temporal evolution in wireless social networks and its impact on capacity. *IEEE Transactions on Parallel and Distributed Systems*, 25(10):2583 – 2594, 2014.

[25] Naga Bhushan, Junyi Li, Durga Malladi, Rob Gilmore, Dean Brenner, Aleksandar Damnjanovic, Ravi Sukhavasi, Chirag Patel, and Stefan Geirhofer. Network densification: the dominant theme for wireless evolution into 5g. *IEEE Communications Magazine*, 52(2):82–89, 2014.

[26] Yanhong Wu, Naveen Pitipornvivat, Jian Zhao, Sixiao Yang, Guowei Huang, and Huamin Qu. egoslider: Visual analysis of egocentric network evolution. *IEEE transactions on visualization and computer graphics*, 22(1):260–269, 2016.

[27] Tianyang Zhang, Peng Cui, Christos Faloutsos, Yunfei Lu, Hao Ye, Wenwu Zhu, and Shiqiang Yang. Come-and-go patterns of group evolution: A dynamic model. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 1355–1364. ACM, 2016.

[28] Bin Yin, Yujiu Yang, and Wenhuang Liu. Exploring social activeness and dynamic interest in community-based recommender system. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 771–776. ACM, 2014.

[29] Johan Koskinen and Christofer Edling. Modelling the evolution of a bipartite network - peer referral in interlocking directorates. *Social Networks*, 34(3):309–322, 2012.

[30] Meifeng Dai, Qi Xie, and Lei Li. Dynamic weight evolution network with preferential attachment. *Physica A: Statistical Mechanics and its Applications*, 416:149–156, 2014.

[31] K. B. Athreya. Preferential attachment random graphs with general weight function. *Internet Mathematics*, 4(4):401–418, 2007.

[32] J. Ben Schafer, Frankowski Dan, Jon Herlocker, and Shilad Sen. Collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22(1):5–53, 2004.

[33] Xin Pan, Guishi Deng, and Jian Guo Liu. Weighted bipartite network and personalized recommendation. *Physics Procedia*, 3(5):1867–1876, 2010.

[34] Danping Zhang, Meifeng Dai, Lei Li, and Cheng Zhang. Distribution characteristics of weighted bipartite evolving networks. *Physica A Statistical Mechanics & Its Applications*, 428:340–350, 2015.

[35] Òscar Celma Herrada et al. *Music recommendation and discovery in the long tail*. Universitat Pompeu Fabra, 2009.

[36] Cai-Nicolas Ziegler, Sean M McNee, Joseph A Konstan, and Georg Lausen. Improving recommendation lists through topic diversification. In *Proceedings of the 14th international conference on World Wide Web*, pages 22–32. ACM, 2005.

[37] Julian McAuley and Alex Yang. Addressing complex and subjective product-related queries with customer reviews. In *Proceedings of the 25th International Conference on World Wide Web*, pages 625–635. International World Wide Web Conferences Steering Committee, 2016.

[38] Xinbing Wang and Luoyi Fu. Acenap: A general-purpose network analysis dataset. 2017.

[39] KONECT. Libimseti.cz network dataset, 2017.

[40] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM REVIEW*, 51(4):661–703, 2009.

[41] William Aiello, Fan Chung, and Linyuan Lu. Random evolution in massive graphs. In *Handbook of massive data sets*, pages 97–122. Springer, 2002.