This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

Probabilistic Modeling of Assimilate-Contrast Effects in Online Rating Systems

Hong Xie, Mingze Zhong, Xiaoyu Shi, Xiaoying Zhang, Jiang Zhong and Mingsheng Shang

Abstract—Online rating system serves as an indispensable building block for many web applications. Previous studies showed that due to assimilate-contrast effects, historical ratings could significantly distort users' ratings, leading to low accuracy of product quality estimation and recommendation. To understand assimilate-contrast effects, an "accurate" model is still missing as previous models do not capture important factors like rating recency, selection bias, etc. Furthermore, an analytical framework to characterize product estimation accuracy under assimilate-contrast effects is also missing. This paper aims to fill in this gap. We propose a probabilistic model to quantify the aforementioned important factors on assimilate-contrast effects. We apply stochastic approximation theory to show that when the rating bias satisfies mild contraction conditions, the aggregate rating converges under aggregate opinion heterogeneity. We also apply non-stationary Markov chain theory to show that when the strength of assimilate-contrast satisfies mild stable conditions, the aggregate rating converges under rating recency. We also derive an equation to characterize the converged aggregate ratings. These conditions reveal important insights on how the aforementioned factors influence the convergence and guide the online rating system operator to design appropriate rating aggregation rules and rating displaying strategies. We apply it to rating prediction tasks and product recommendation tasks. Experiment results on four public datasets show that our model can improve the rating prediction and recommendation accuracy over previous models significantly, under various metrics like RMSE, NDCG, etc. We also demonstrate the flexibility of our model by showing that it can be applied to enhance other rating behavior models.

Index Terms—Online rating system, assimilate-contrast effects, rating prediction, recommendation

1 INTRODUCTION

Online rating system serves as an indispensable building block for many web applications such as Amazon, TripAdvior and Yelp. It breaks the information asymmetry between users and sellers benefiting both sides. On the one hand, online ratings reveal the quality of products to users. Users assign ratings to products, which reflect their overall opinions or experiences on products. Assigned ratings are displayed to all users and the aggregation of online ratings (a.k.a. wisdom of the crowd) reveal product quality. This is helpful especially when a user encounters a product which he has never purchased before (or has no experience with) [1]. Thus, online rating system can enable better product purchase decisions, which in turn improves purchasing experiences of users [2], [3], [4]. On the other hand, online ratings reveal users' preferences, which is helpful for product recommendation. Users assign high ratings to preferred products. For example, a user who is fond of a certain brand may assign higher ratings to products of this brand. Online ratings can be utilized to infer users' personal preferences, which in turn improves the accuracy of product recommendation. Improving recommendation accuracy can improve the sales of sellers [5], [6], [7]. Formally, each user assigns ratings to a subset of products, i.e., we only have "partial information". This partial information makes the product quality estimation and recommendation non-trivial.

Unbiasedness of ratings is of vital importance to the accuracy of online rating based product quality estimation and product recommendation. A number of evidences showed that historical ratings distort user ratings, leading to biased or even erroneous ratings. Many survey studies revealed that users form opinions from historical ratings of a product [8], [9], [10], which makes their rating bias toward historical ratings. Several controlled experiment studies also showed that user ratings are bias toward historical ratings [11], [12], [13], [14]. More specifically, these experiment studies revealed that users tend to assign higher ratings when higher historical ratings are displayed to them. Via extensive data analysis, Zhang *et al.* [15] identified the assimilate-contrast phenomenon in online ratings, which is supported by the "Assimilate-Contrast" theory from psychology. This phenomenon shows that users either "assimilate" or "contrast" to historical ratings. In particular, users conform to historical ratings if a product is slightly over or under rated (assimilation), while users deviate from historical ratings if a product is significantly over or under rated (contrast). They proposed a model to quantify the assimilate-contrast effects and applied it debias online ratings. Experiments showed that their model can improve the rating prediction and recommendation accuracy.

However, both the micro aspect and macro aspect of the assimilate-contrast effects are not well understood. The micro aspect refers to user rating behavior under the assimilate-contrast effects. An accurate model on this micro aspect is still missing. Previous models [15], [16] on assimilate-contrast effects do not capture important factors in user rating behavior such as rating recency [8], [9], [10],

- *Mingsheng Shang is the corresponding author.*
- *Hong Xie, Mingze Zhong, Xiaoyu Shi and Mingsheng Shang are with Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences and Chongqing School, University of Chinese Academy of Sciences. Emails: {xiehong,xiaoyushi,msshang}@cigit.ac.cn*
- *Xiaoying Zhang is with Bytedance. Email: jingyuxy@gmail.com*
- *Jiang Zhong is with College of Computer Science, Chongqing University. E-mail: zhongjiang@cqu.edu.cn.*

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

selection bias of ratings [17], [18], etc. Rating recency means that users form their opinions from a small number of latest ratings [8], [9], [10]. Selection bias means that users tend to assign ratings to items that they hate or like [17], [18]. The macro aspect refers to the aggregate rating or opinions under the assimilate-contrast effects. To the best of our knowledge, an analytical framework to characterize product estimation accuracy under assimilate-contrast effects is still missing. We are aware that Zhang *et al.* [15], [16] provided some numerical simulation of their model to show that under assimilate-contrast effects, the aggregation of online ratings may evolve toward a value that is inaccurate in reflecting the intrinsic quality of a product. However, two questions are still under: *(1) What are the conditions to guarantee the convergence of aggregate ratings or opinions? (2) How does assimilate-contrast effects influence the accuracy of product quality estimation?* Answering these questions would benefit rating aggregation rules design and rating displaying strategies design, etc.

To fill in the aforementioned gap, this paper aims to present an in-depth study of assimilate-contrast effects in online rating systems. There are three challenges. *How to balance the model complexity vs. model accuracy tradeoff?* The model needs to accurately capture important factors missed by previous models and at the same time be neat enough to support analytical studies of assimilate-contrast effects. *How to design an analytical framework to track the dynamics of aggregate ratings or opinions and characterized the convergences of them?* The analytical framework needs to establish sufficient conditions to guarantee the convergence and reveal the impact of various factors on the convergence. *How to apply our model to improve product recommendation?* One needs to appropriately parameterize our model to estimate user preference and design inference algorithm to infer model parameters. We have addressed these challenges. Figure 1 illustrates the high-level idea of our method. More specifically, we first formulate a personalized rating model to quantify the assimilate-contrast effects. We then develop two independent extensions of this model: (1) extension to characterize the user population rating behavior, which enables us to analyze the evolving dynamics of aggregate opinion; (2) extension to rating factorization, which enables us to conduct rating prediction and recommendation. Our contributions are:

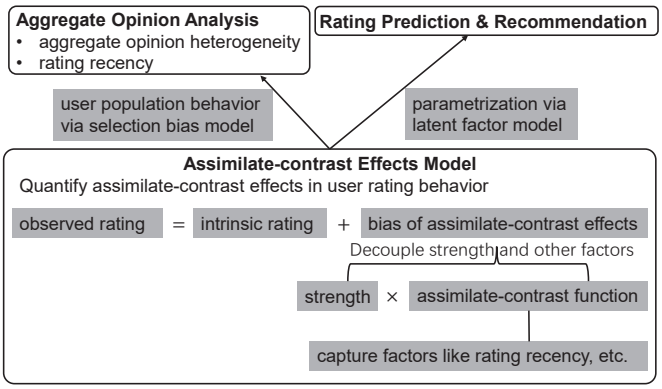

Fig. 1: Illustrating high-level idea of our method.

- We formulate a probabilistic model to quantify assimilate-contrast effects in the micro aspect. Our model captures important factors like rating recency, selection bias, etc., in user rating behavior, which are missed in previous models [15], [16]. Furthermore, our model is neat enough to enable the analytical study of assimilate-contrast effects.
- In the macro aspect, we apply stochastic approximation theory to show that when the rating bias satisfies mild contraction conditions, the aggregate ratings converges under aggregate opinion heterogeneity. We also apply non-stationary Markov chain theory to show that when the strength of assimilate-contrast satisfies mild stable conditions, the aggregate ratings converges under rating recency. We also derive an equation to characterize the converged aggregate ratings. These convergence conditions reveal important factors like selection bias, rating recency, aggregation rules, etc., on the convergence of aggregate ratings or opinions. These insights guide the design of rating aggregation rules, etc.
- We parameterize our model and apply regularized least square method to infer model parameters. To demonstrate the versatility of our model, we apply it to rating prediction and product recommendation. Extensive experiments on four public datasets show that our model improves the accuracy of rating prediction and recommendation over a number of baselines, i.e., HIALF [16], LF [19] and Herd [20], etc., under various metrics like RMSE, NDCG, etc. We also show that our model can be straightforwardly applied to improve other models like the spiral of silence model [21].

The remaining parts of this paper organize as follows. Section 2 presents a probabilistic model to quantify the assimilate-contrast effects. Section 3 presents an analytical framework to study the converges of aggregate ratings or opinions. Section 4 parametrizes our model and applies it to rating prediction. Section 5 applies our model product recommendation. Section 6 presents extensions of our model. Section 7 presents the related work and Section 8 concludes.

2 Online Rating Model

We first present the baseline model of unbiased ratings. Then, we present a model to quantify the assimilate-contrast effects. Finally, we present a model to quantify the selection bias of users. Figure 2 illustrates the logic dependency among key notations of this paper.

2.1 Overview the Assimilate-Contrast Effects Model

Consider an online rating system with a set $\mathcal{U} \subseteq \mathbb{N}$ of users and a set $\mathcal{I} \subseteq \mathbb{N}$ of items. An item represents a Book in Amazon, a Hotel in TripAdvisor, etc. Users evaluate the quality of products based on an $M \in \mathbb{N}_+$ level cardinal rating metric denoted by $\mathcal{M} \triangleq \{1, \ldots, M\}$, where $M \in \mathbb{N}_+$. A higher rating to an item implies that a user is more satisfied with an item. For example, the rating metric in Amazon and Tripadvisor is $\mathcal{M} = \{1 =$ "Terrible", 2 = "Poor", 3 = "Average", 4 = "Good", 5 = "Excellent"$\}$.

Let $R_{i,k} \in \mathcal{M}$ denote the k-th observed rating of item i, where $k \in \mathbb{N}_+$. Let $t_{i,k} \in \mathbb{R}_+$ denote the arrival time (or time stamp) associated with $R_{i,k}$. Let $\mathcal{H}_{i,k}$ denote a

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

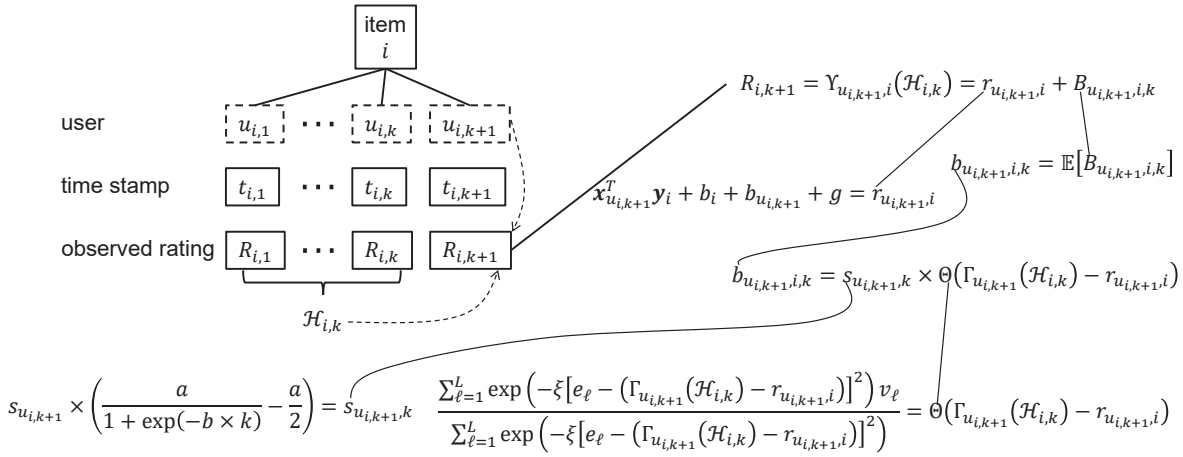

Fig. 2: Illustrating the logic dependency among key notations.

set of all historical ratings of item i up to the k-th rating, formally $\mathcal{H}_{i,k} \triangleq \{(t_{i,1}, R_{i,1}), \ldots, (t_{i,k}, R_{i,k})\}, \forall k \in \mathbb{N}_+$. For presentation convenience, we set $\mathcal{H}_{i,0} \triangleq \emptyset$ by default.

Consider a user $u \in \mathcal{U}$ who has not assigned ratings to item i. Let $\Upsilon_{u,i}(\mathcal{H}_{i,k}) \in \mathcal{M}$ denote the rating that user u would assign to item i under assimilate-contrast effects caused by rating history $\mathcal{H}_{i,k}$. Formally, we model $\Upsilon_{u,i}(\mathcal{H}_{i,k})$ as follows:

$$\Upsilon_{u,i}(\mathcal{H}_{i,k}) = \underbrace{r_{u,i}}_{\text{intrinsic rating}} + \underbrace{B_{u,i,k}}_{\text{bias of assimilate-contrast effects}}, \quad (1)$$

where $r_{u,i}$ denote the intrinsic rating (will be modeled in Section 2.2) and $B_{u,i,k} \in \{1 - r_{u,i}, \ldots, M - r_{u,i}\}$ denotes the rating bias caused by assimilate-contrast effects. Equation (1) models that under the influence of rating history $\mathcal{H}_{i,k}$, the rating assigned by a user is a combination of intrinsic rating and rating bias caused by assimilate-contrast effects. Unbiased rating is a special case of our model with $B_{u,i,k}=0$. This paper focuses on the general setting that $B_{u,i,k}$ is a random variable capturing uncertainty in rating bias.

2.2 Modeling Intrinsic Ratings

Intrinsic rating. Let $r_{u,i} \in \mathcal{M}$ denote the intrinsic/unbiased rating of user $u \in \mathcal{U}$ toward item $i \in \mathcal{I}$. Namely, $r_{u,i}$ characterizes user u's intrinsic overall opinion toward item i. Denote the fraction of users whose intrinsic overall opinions toward item i is $m \in \mathcal{M}$ as $o_{i,m} \triangleq \sum_{u \in \mathcal{U}} \mathbb{1}_{\{r_{u,i}=m\}} / |\mathcal{U}|, \forall m \in \mathcal{M}$. For each item i, it holds that $\sum_{m \in \mathcal{M}} o_{i,m} = 1$. Note that $o_{i,m}$ is unobservable because only a subset of users assign ratings to item i and assigned ratings may contain certain bias as we will model in Section 2.1. For simplicity, define the intrinsic opinion vector as $\boldsymbol{o}_i \triangleq [o_{i,1}, \ldots, o_{i,M}], \forall i \in \mathcal{I}$. The intrinsic opinion vector $\boldsymbol{o}_i$ characterizes the collective opinion of the whole user population $\mathcal{U}$ toward item i. For example, $\boldsymbol{o}_i = [0.01, 0.04, 0.1, 0.35, 0.5]$ means that 50% of users hold an intrinsic opinion of 5 toward item i. Let $\mathcal{O}$ denote a set of all the possible collective opinion vectors. Formally, $\mathcal{O}$ can be expressed as: $\mathcal{O} \triangleq \{\boldsymbol{o} | \boldsymbol{o} \in [0,1]^M, \sum_{m \in \mathcal{M}} o_m = 1\}$.

We model an opinion aggregation rule as a function $A : \mathcal{O} \to [1, M]$. Namely, the function A aggregates the collective opinion to produce an indicator on the overall quality of an item. We refer to $A(\boldsymbol{o}_i)$ as the intrinsic quality of item i. For example, the rating aggregation rule in Amazon and TripAdvisor is the *average scoring rule*, which can be expressed as $A(\boldsymbol{o}_i) = \sum_{m=1}^{M} m o_{i,m}$. Furthermore, consider $\boldsymbol{o}_i = [0.01, 0.02, 0.1, 0.35, 0.5]$, we have $A(\boldsymbol{o}_i) = 4.25$.

Selection bias. A user arrived at an item may not be drawn from the user population uniformly at random, but instead the arrival of this user may be biased toward the item exposure policy of items, self-selection behavior of users, etc. The selection bias characterizes the non-uniform random arrival of users to an item. Let $u_{i,k} \in \mathcal{U}$ denote the user who assigns rating the rating $R_{i,k}$. The $k+1$-th rating of item i is:

$$R_{i,k+1} = \Upsilon_{u_{i,k+1},i}(\mathcal{H}_{i,k}).$$

We consider a large user population. Each user $u_{i,k}$ is drawn from the whole user population. We use a probability distribution $\mathcal{D}_u$ to characterize the uncertainty in this process of drawing users. The support of $\mathcal{D}_u$ is $\mathcal{U}$, and drawing a user is modeled as generating a sample from $\mathcal{D}_u$, i.e., $u_{i,k} \sim \mathcal{D}_u$. We consider the case that $u_{i,k}, \forall k \in \mathbb{N}_+$ are independently generated from $\mathcal{D}_u$. Define intrinsic opinion $\widetilde{o}_{i,m}$ as

$$\widetilde{o}_{i,m} \triangleq \mathbb{P}_{u_{i,k} \sim \mathcal{D}_u}[r_{u_{i,k},i} = m],$$

where $m \in \mathcal{M}$. One can observe that $\widetilde{o}_{i,m} \in [0,1]$ and $\sum_{m \in \mathcal{M}} \widetilde{o}_{i,m} = 1$. For simplicity, denote $\widetilde{\boldsymbol{o}}_i \triangleq [\widetilde{o}_{i,1}, \ldots, \widetilde{o}_{i,M}]$. For example, when each $u_{i,k}$ is sampled from the whole user population uniformly at random, the $\widetilde{o}_{i,m}$ satisfies $\widetilde{o}_{i,m} = o_{i,m}$. By nature of the selection bias, i.e., users tend to assign ratings to the item that they hate or like [17], [18], the intrinsic opinion vector $\widetilde{\boldsymbol{o}}_i$ quantifies the selection bias of the user population. The case $\widetilde{\boldsymbol{o}}_i \neq \boldsymbol{o}_i$ reflects that the user population has selection bias. In general, as the number of ratings is larger enough, the aggregate intrinsic ratings converges to the intrinsic opinion $\widetilde{\boldsymbol{o}}_i$, i.e., $\lim_{k \to \infty} \sum_{j=1}^{k} \mathbb{1}_{\{r_{u_{i,j},i}=m\}} / k = \widetilde{o}_{i,m}$.

2.3 Modeling Bias of Assimilate-contrast Effects

Modeling the rating bias $B_{u,i,k}$. We use a two layer process to model the rating bias $B_{u,i,k}$ caused by assimilate-contrast effects. In the first layer, user u forms an initial opinion

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

toward item i from historical ratings $\mathcal{H}_{i,k}$. Note that this layer is supported by several survey studies, which revealed that users form opinions toward an item from historical ratings before adopting an item [8], [9], [10]. We quantify this initial opinion via function $\Gamma_u(\mathcal{H}_{i,k}) \in [1, M]$. By the nature of assimilate-contrast effects, the function $\Gamma_u(\mathcal{H}_{i,k})$ is influenced by important factors like rating recency, aggregate opinion heterogeneity, etc. We will discuss how to model these important factors in the function $\Gamma_u(\mathcal{H}_{i,k})$ later. A larger $\Gamma_u(\mathcal{H}_{i,k})$ means that the initial opinion of user u is more positive toward item i. In the second layer, user u adopts the item and experiences the quality of the item. We model the experienced quality of item i as $r_{u,i}$. Due to assimilate-contrast effects, the initial opinion $\Gamma_u(\mathcal{H}_{i,k})$ and the experienced quality $r_{u,i}$ result in the rating bias $B_{u,i,k}$. Formally, $B_{u,i,k}$ follows a probability distribution denoted by $\mathcal{D}_B(s_{u,k}, \Gamma_u(\mathcal{H}_{i,k}), r_{u,i})$, where $s_{u,k} \in [0, 1]$ denotes the strength of assimilate-contrast effects. The $s_{u,k}$ increases as the strength of assimilate-contrast effects increase, i.e., a user is more prone to assimilate-contrast effects. We further model the expectation of $B_{u,i,k}$ as follows:

$$b_{u,i,k} = \mathbb{E}_{B_{u,i,k} \sim \mathcal{D}_B(s_{u,k}, \Gamma_u(\mathcal{H}_{i,k}), r_{u,i})}[B_{u,i,k}] \quad (2)$$

$$= \underbrace{s_{u,k}}_{\text{strength}} \times \underbrace{\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})}_{\text{assimilate-contrast function}}, \quad (3)$$

where $\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})$ denotes the assimilate contrast function. Namely, the assimilate-contrast function $\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})$ models the base bias and the strength $s_{u,k}$ scales this base bias. Larger strength $s_{u,k}$ leads to larger bias. The case $s_{u,k} = 0$ models that a user is not influenced by assimilate-contrast effects. The strength $s_{u,k}$ varies across user u and the number of ratings k capturing that different users may have different strengths of assimilate-contrast effects and a user may have different strengths of assimilate-contrast effects when the number of ratings is different.

One can parametrize $\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})$ to capture assimilate-contrast phenomenon. For example, to capture the empirical findings in [16], it can be parametrized such that $\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})$ is close to $\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i}$ when $|\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i}|$ is small (i.e., the product is slightly under or over rated) capturing the assimilate effects. And $\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})$ is much smaller (or larger) than $\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i}$ when $\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i}$ is much larger (or smaller) than 0, i.e., an item is highly over rated (or under) rated, capturing contrast effects. Unless we parametrize it explicitly, we focus on the general setting that Θ is nonparametric. We assume that Θ is a continuous function.

Modeling initial opinion formation. Now we model the initial opinion formation, i.e., the function $\Gamma_u(\mathcal{H}_{i,k})$. Users may form initial opinions from the aggregation of historical ratings or a small number of latest ratings.

- **Aggregate-opinion-based initial opinion formation.** We first consider the case that users form initial opinions from the aggregation of historical ratings. Denote the collective opinion summarized from the rating history $\mathcal{H}_{i,k}$ as $\boldsymbol{h}_{i,k} \triangleq [h_{i,k,1}, \ldots, h_{i,k,M}]$, where $h_{i,k,m} \in [0, 1]$ and $\sum_{m \in \mathcal{M}} h_{i,k,m} = 1$. The $\boldsymbol{h}_{i,k}$ is public to all users. We consider a *class* of weighted aggregation rules to summarize historical ratings:

$$h_{i,k,m} = \frac{\sum_{j=1}^{k} \alpha_j \mathbb{1}_{\{R_{i,j}=m\}}}{\sum_{j=1}^{k} \alpha_j}, \quad \forall m \in \mathcal{M}, k \in \mathbb{N}_+, \quad (4)$$

where $\alpha_j \in \mathbb{R}_+$ denotes the weight associated with j-th rating, and $\mathbb{1}$ is an indicator function. For example, $\alpha_j = 1, \forall j$, is deployed in Amazon and TripAdvisor, which corresponds to the "*average rating rule*". Under this average rating rule, we have $h_{i,k,m} = \sum_{j=1}^{k} \mathbb{I}_{\{R_{i,j}=m\}}/k$, which is the fraction of historical ratings equal m. Note that $\boldsymbol{h}_{i,k}$ is displayed to all users. We capture the aggregate opinion heterogeneity in initial opinion formation as:

$$\Gamma_u(\mathcal{H}_{i,k}) = \frac{\sum_{m \in \mathcal{M}} m \beta_m h_{i,k,m}}{\sum_{m \in \mathcal{M}} \beta_m h_{i,k,m}}, \quad (5)$$

where the weight $\beta_m \in [0, 1]$ models how a user weighs the opinion associated with each rating level. For example, a user may assign a large weight to low ratings representing that she is sensitive to negative opinions. There are several possible parametric forms of the weights:

$$\beta_m = \exp(-\gamma m), \;\; \beta_m = m^{\gamma}, \;\; \beta_m = \ln(m^{\gamma} + 1).$$

where $\gamma \in \mathbb{R}$. The above selections of β_m cover three typical types of functions, i.e., exponential, polynomial and logarithmic. These functions capture the opinion heterogeneity from strong to weak and they form a reasonably large function space to fit the ground truth β_m.

- **Rating-recency-based initial opinion formation.** Now, we consider the case that users form initial opinions from a small number of latest ratings. Let $n \in \mathbb{N}_+$ denote the number of latest ratings that users refer to for initial opinion formation. We capture rating recency in initial opinion formation as:

$$\Gamma_u(\mathcal{H}_{i,k}) = \frac{\sum_{j=1}^{n} \eta_{k,j} R_{i,k-j+1} \mathbb{1}_{\{k-j+1 \geq 1\}}}{\sum_{j=1}^{n} \eta_{k,j} \mathbb{1}_{\{k-j+1 \geq 1\}}}, \quad (6)$$

where the weight satisfies $\eta_{k,j} \in \mathbb{R}_+$ and we set $R_{i,j} = 0$ for all $j < 1$ by default. There are several possible forms of the weight $\eta_{k,j}$.

Order-based rating recency. One can use the following forms of $\eta_{k,j}$ to model arrival-order-based initial opinion formation:

$$\eta_{k,j} = \exp(-\zeta j), \;\; \eta_{k,j} = j^{-\zeta}, \;\; \eta_{k,j} = 1/(\zeta \ln(j + 1)),$$

where $\zeta \in \mathbb{R}$.

Time-stamp-based rating recency. One can use the following forms of $\eta_{k,j}$ to model arrival-time-stamp-based initial opinion formation:

$$\eta_{k,j} = \exp(-\rho(t_u - t_{i,k+1-j})^{\delta}), \;\; \eta_{k,j} = (t_u - t_{i,k+1-j})^{-\rho},$$
$$\eta_{k,j} = 1/(\rho \ln(t_u - t_{i,k+1-j} + 1)).$$

where t_u denotes the arrival time of user u, $\delta \in \mathbb{R}, \rho \in \mathbb{R}$, we set $t_{i,j} = 0$ by default for all $j < 1$. The selections of $\eta_{k,j}$ covers three typical types of functions, i.e., exponential, polynomial form and logarithmic.

The above selections of $\eta_{k,j}$ cover three typical types of functions, i.e., exponential, polynomial and logarithmic. These functions capture the strength of rating recency from strong to weak and they form a reasonably large function space to fit the ground truth $\eta_{k,j}$.

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

3 Theoretical Analysis

In this section, we first analyze the convergence of historical aggregate opinion under the case that users form initial opinions from historical aggregate opinions. Then we analyze the case that users form initial opinions from a small number of latest ratings. Finally, we analyze the hybrid case mixing of both types of users. *All proofs are in our supplementary file.*

3.1 Aggregate Opinion Heterogeneity

We consider the case that the initial opinion is formed from the historical aggregate opinion, i.e., $\Gamma_u(\mathcal{H}_{i,k})$ is derived in Eq. (5). Table 1 summarizes key notations defined to characterize aggregate opinion heterogeneity.

TABLE 1: Key notations for characterizing the aggregate opinion heterogeneity.

$\boldsymbol{e}_m$	the M-dimensional standard base vector
$\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$	the expected overall opinion bias of the whole user population caused by the assimilate-contrast effects
$\widetilde{\mathcal{D}}_B(\cdot)$	the distribution of $B_{u,i,k}$ under aggregate opinion
ω_k	the adjusted weight $\alpha_k/\sum_{j=1}^{k}\alpha_j$
$\boldsymbol{W}_{i,k+1}$	the stochastic noise of the opinion vector
ρ_k	the contraction factor

Opinion dynamics via non-stationary stochastic difference equation. Let $\boldsymbol{e}_m \triangleq [e_{m,1},\ldots,e_{m,M}], \forall m \in \mathcal{M}$, denote an M-dimensional standard basis vector, where the m-th entry is $e_{m,m}=1$ and all the other entries are zero, i.e., $e_{m,m'}=0, \forall m' \neq m$. Observe that $\Gamma_u(\mathcal{H}_{i,k})$ satisfying Equation (5) implies that $\Gamma_u(\mathcal{H}_{i,k})$ is a function of $\boldsymbol{h}_{i,k}$. Then it follows that for each given tuple (u,i,k), the rating bias $B_{u,i,k}$ follows a probability distribution, which is determined by historical aggregate opinion $\boldsymbol{h}_{i,k}$, the strength of assimilate-contrast effects $s_{u,k}$ and the intrinsic rating $r_{u,i}$. For clarity, we denote this distribution of $B_{u,i,k}$ as $\widetilde{\mathcal{D}}_B(\boldsymbol{h}_{i,k}, s_{u,k}, r_{u,i})$. Denote $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k}) \in [0,1]^M$ as:

$$\begin{aligned}\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k}) &\triangleq \mathbb{E}[B_{u,i,k}|\mathcal{H}_{i,k}] \\ &= \mathbb{E}_{u\sim\mathcal{D}_u}\Big[\mathbb{E}_{B_{u,i,k}\sim\widetilde{\mathcal{D}}_B(\boldsymbol{h}_{i,k},s_{u,k},r_{u,i})}\left[\boldsymbol{e}_{B_{u,i,k}}\right]\Big]. \quad (7)\end{aligned}$$

The $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$ characterizes the expected overall opinion bias of the whole user population caused by the assimilate-contrast effects. This expected overall opinion bias $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$ is determined by historical aggregate opinion $\boldsymbol{h}_{i,k}$ and the number of ratings k. In the following lemma, we quantify the impact of $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$ on the dynamics of historical aggregate opinion.

Lemma 1. *The historical aggregate opinion $\boldsymbol{h}_{i,k}$ satisfies:*

$$\boldsymbol{h}_{i,k+1} = (1-\omega_{k+1})\boldsymbol{h}_{i,k} + \omega_{k+1}\left[\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k}) + \widetilde{\boldsymbol{o}}_i + \boldsymbol{W}_{i,k+1}\right],$$

where $\omega_k \triangleq \alpha_k/\sum_{j=1}^{k}\alpha_j$ and $\boldsymbol{W}_{i,k+1}$ is defined as

$$\boldsymbol{W}_{i,k+1} \triangleq \boldsymbol{e}_{R_{i,k+1}} - \boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k}) - \widetilde{\boldsymbol{o}}_i.$$

Furthermore, $\boldsymbol{W}_{i,k+1}$ satisfies $\mathbb{E}[\boldsymbol{W}_{i,k+1}|\mathcal{H}_{i,k}] = 0$.

Remark: Lemma 1 states that the evolving dynamics of the historical aggregate opinion $\boldsymbol{h}_{i,k}$ is governed by a non-stationary stochastic difference equation. This non-stationary stochastic difference equation quantifies the impact of rating aggregation rule (i.e., weights α_j), intrinsic opinion vector $\widetilde{\boldsymbol{o}}_i$ and the expected overall opinion bias $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$ on the dynamics of $\boldsymbol{h}_{i,k}$. Furthermore, this non-stationary stochastic difference equation serves as an important building block for one to further analyze the dynamics of historical aggregate opinion analytically.

Convergence of opinions. Based on Lemma 1, now we study the convergence of historical aggregate opinion analytically via stochastic approximation techniques. The following theorem states sufficient conditions under which historical aggregate opinion converges to the intrinsic opinion.

Theorem 1. *Suppose ω_k satisfies*

$$\sum_{k=1}^{\infty}\omega_k = \infty, \qquad \sum_{k=1}^{\infty}\omega_k^2 < \infty. \quad (8)$$

Suppose $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$ satisfies

$$\|\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})\| \leq \rho_k \|\boldsymbol{h}_{i,k} - \widetilde{\boldsymbol{o}}_i\|, \quad (9)$$

where $\rho_k \in [0,1)$ and $\sup_{k\in\mathbb{N}}\rho_k < 1$. Then, $\boldsymbol{h}_{i,k}$ converges to $\widetilde{\boldsymbol{o}}_i$ almost surely, i.e., $\mathbb{P}\left[\lim_{k\to\infty}\boldsymbol{h}_{i,k} = \widetilde{\boldsymbol{o}}_i\right] = 1$. Furthermore, we have $\mathbb{P}\left[\lim_{k\to\infty}A(\boldsymbol{h}_{i,k}) = A(\widetilde{\boldsymbol{o}}_i)\right] = 1$.

Remark: In Theorem 1, Condition (8) characterizes a class of rating aggregation rules, under which the historical aggregate opinion can converge. To illustrate Condition (8), consider the example of "average scoring rule", i.e., $\alpha_j = 1$ and $\omega_k = 1/k$. One can easily check that this average scoring rule satisfies Condition (8). Condition (9) states that the overall opinion bias caused by the assimilate-contrast effects is closer to the intrinsic opinion than the historical aggregate opinion. In other words, users do not fully follow the historical aggregate opinion. This condition is consistent with the assimilate-contrast effects, because under the assimilate-contrast effects if an item is highly over or under rated, users follow the historical aggregate opinion to a low degree, while if an item is slightly over or under rated, users may follow the historical aggregate opinion to a certain high degree. Theorem 1 holds for any selection bias and it implies the following corollary.

Corollary 1. *Under the conditions of Theorem 1. If $\widetilde{\boldsymbol{o}}_i = \boldsymbol{o}_i$, we have $\mathbb{P}\left[\lim_{k\to\infty}\boldsymbol{h}_{i,k} = \boldsymbol{o}_i\right] = 1$ and $\mathbb{P}\left[\lim_{k\to\infty}A(\boldsymbol{h}_{i,k}) = A(\boldsymbol{o}_i)\right] = 1$.*

Remark. Corollary 1 states that under the same conditions as Theorem 1, if there is no selection bias, the intrinsic collective opinion of the whole user population can be revealed by the historical aggregate opinion when the number of ratings is sufficiently large. When not all users form initial opinion from the historical aggregate opinion, i.e., some may form initial opinions from a small number of latest ratings, Condition (9) may not hold. The following theorem generalizes Theorem 1 to handle this case.

Theorem 2. *Suppose Condition (8) holds. Suppose the opinion bias $\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k})$ satisfies:*

$$\|\boldsymbol{\Psi}_{i,k}(\boldsymbol{h}_{i,k}) + \widetilde{\boldsymbol{o}}_i - \boldsymbol{v}\| \leq \rho_k \|\boldsymbol{h}_{i,k} - \boldsymbol{v}\|,$$

where $\boldsymbol{v} \in [0,1]^M$. Then, it holds that $\mathbb{P}\left[\lim_{k\to\infty}\boldsymbol{h}_{i,k} = \boldsymbol{v}\right] = 1$. Furthermore, we have $\mathbb{P}\left[\lim_{k\to\infty}A(\boldsymbol{h}_{i,k}) = \boldsymbol{v}\right] = 1$.

Remark: Theorem 2 characterizes the convergence of historical aggregate opinion when there is a certain deviation (quantified by $\boldsymbol{v} - \widetilde{\boldsymbol{o}}_i$) in the expected overall opinion bias.

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

Theorem 1 is a special case of the above theorem when there is no deviation, i.e., $\boldsymbol{v} - \widetilde{\boldsymbol{o}}_i = \boldsymbol{0}$.

3.2 Rating Recency

We consider the case that users form initial opinions from the latest n ratings, i.e., $\Gamma_u(\mathcal{H}_{i,k})$ is derived in Equation (6). In this case, the expected overall opinion bias is not a function of $\boldsymbol{h}_{i,k}$, i.e., Equation (7) does not hold anymore. Hence, the non-stationary stochastic difference equation stated in Lemma 1 can not be applied to quantify the dynamics of historical aggregate opinion $\boldsymbol{h}_{i,k}$. We will use non-stationary Markov chain to characterize the dynamics of $\boldsymbol{h}_{i,k}$ and study its convergence. Table 2 summarizes key notations defined to characterize aggregate opinion under rating recency.

TABLE 2: Key notations for characterizing aggregate opinion under rating recency.

$\widehat{\mathcal{D}}_B(\cdot)$	the distribution of $B_{u,i,k}$ under the Markov model
$\boldsymbol{P}^{(i,\tau)}$	the state transition matrix in time slot τ
$\mathcal{S}$	the state space
$\mathcal{N}_{\boldsymbol{s}}$	the neighboring set of state $\boldsymbol{s}$
$\boldsymbol{p}_k$	the landing probability after k steps of transition
$\boldsymbol{\pi}$	the stationary landing probability
$s_{u,\infty}$	the limiting assimilate-contrast effects strength
$\mathcal{R}$	recurrent state
$\mathcal{T}$	transient state

Order dependent initial opinion. We first consider the case that users' initial opinions only depend on the arrival order of the latest n ratings. Namely, the weight $\eta_{k,j}$ derived in Equation (6) only depends on the arrival order j and it does not depend on the arrival time stamp of ratings. We will generalize to consider arrival time stamps later. The initial opinion $\Gamma_u(\mathcal{H}_{i,k})$ is determined by ratings $(R_{i,k-n+1}, \ldots, R_{i,k})$. Recall that the rating bias $B_{u,i,k}$ follows a distribution $\mathcal{D}_B(s_{u,k}, \Gamma_u(\mathcal{H}_{i,k}), r_{u,i})$. In other words, the distribution of $B_{u,i,k}$ is determined by $s_{u,k}, r_{u,i}$ and $(R_{i,k-n+1}, \ldots, R_{i,k})$. We denote this distribution $\widehat{\mathcal{D}}_B(s_{u,k}, r_{u,i}, R_{i,k-n+1}, \ldots, R_{i,k})$.

• **Opinion dynamics via non-stationary MC.** We consider a discrete time system indexed by $\tau \in \mathbb{N}_+$. In the τ-th time slot, we are given a rating history $\mathcal{H}_{i,\tau}$. We formulate a discrete time finite state non-stationary Markov chain denoted by $\mathbb{M} = (\mathcal{S}, \{\boldsymbol{P}^{(i,\tau)} : \tau \in \mathbb{N}_+\})$ to characterize the dynamics of the historical aggregate opinion of item i. Here, the $\mathcal{S}$ denotes the state space and $\boldsymbol{P}^{(i,\tau)}$ denotes the state transition matrix in time slot τ.

When the number of ratings of item i is no less than n, i.e., $k \geq n$, we denote n latest ratings of item i by $(R_{i,k-n+1}, \ldots, R_{i,k})$. When the number of ratings of item i is less than n, we denote n latest ratings of item i by

$$(\underbrace{0, \ldots, 0}_{(n-k) \text{ missing ratings}}, \underbrace{R_{i,1}, \ldots, R_{i,k}}_{\text{latest } k \text{ ratings}}),$$

where 0 denotes a default rating representing that a rating is missing. The state space $\mathcal{S}$ contains all possible outcomes of latest n ratings

$$\mathcal{S} = \bigcup_{j=0}^{n} \underbrace{\{0\} \times \cdots \times \{0\}}_{(n-j) \text{ missing ratings}} \times \underbrace{\mathcal{M} \times \cdots \times \mathcal{M}}_{\text{latest } j \text{ observed ratings}}.$$

We say the Markov chain $\mathbb{M}$ is at state $\boldsymbol{s} \triangleq [s_1, \ldots, s_n] \in \mathcal{S}$, if the n latest historical ratings of item i are $s_1, \ldots, s_n$ in chronological order. For example, item i is initialized with no ratings, namely, the initial state is $\boldsymbol{s} = [0, \ldots, 0]$. After the first consumer posts the review $R_{i,1}$, the state becomes i.e., $\boldsymbol{s} = [0, \ldots, 0, R_{i,1}]$.

The matrix $\boldsymbol{P}^{(i,\tau)}$ characterizes all possible state transitions in time slot τ, i.e.,

$$\boldsymbol{P}^{(i,\tau)} \triangleq \left[P^{(i,\tau)}(\tilde{\boldsymbol{s}}|\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{S}, \tilde{\boldsymbol{s}} \in \mathcal{S}\right].$$

Each transition probability $P^{(i,\tau)}(\tilde{\boldsymbol{s}}|\boldsymbol{s})$ is defined as $P^{(i,\tau)}(\tilde{\boldsymbol{s}}|\boldsymbol{s}) \triangleq \mathbb{P}[S_{\tau+1} = \tilde{\boldsymbol{s}}|S_\tau = \boldsymbol{s}]$, where S_τ denotes the state of the Markov chain in time slot τ. To derive the closed-form formula of $P^{(i,\tau)}(\tilde{\boldsymbol{s}}|\boldsymbol{s})$, we define neighbors of a state.

Definition 1. *For each state $\boldsymbol{s} \in \mathcal{S}$, define its neighboring set as*

$$\mathcal{N}_{\boldsymbol{s}} \triangleq \left\{[s_2, \ldots, s_N, m] \middle| m \in \mathcal{M}\right\}.$$

Each state is only possible to transit to one of its neighbors, namely, $P(\tilde{\boldsymbol{s}}|\boldsymbol{s}) = 0, \forall \tilde{\boldsymbol{s}} \notin \mathcal{N}_{\boldsymbol{s}}$. In time slot τ, the next state of the Markov chain $\mathbb{M}$ is governed by rating $R_{i,\tau+1}$. Formally, the probability of transiting to each neighbor node in time slot τ can be expressed as:

$$\begin{aligned}
&P^{(i,\tau)}(\tilde{\boldsymbol{s}}|\boldsymbol{s}) \\
=&\mathbb{P}[R_{i,\tau+1} = \tilde{s}_n | (R_{i,\tau-n+1}, \ldots, R_{i,\tau}) = \boldsymbol{s}], \tilde{s}_n \in \mathcal{N}_{\boldsymbol{s}} \\
=&\mathbb{P}_{u\sim\mathcal{D}_u}\Big[\mathbb{P}_{B_{u,i,k}\sim\widehat{\mathcal{D}}_B(s_{u,k}, r_{u,i}, R_{i,k-n+1}, \ldots, R_{i,k})}[B_{u,i,k} = \tilde{s}_n - r_{u,i}]\Big].
\end{aligned}$$

The transition probability $P^{(i,\tau)}(\tilde{\boldsymbol{s}}|\boldsymbol{s})$ is determined by current state $\boldsymbol{s}$ and the index of the time slot τ.

Convergence of aggregate opinions. Let $\boldsymbol{p}_0 \triangleq [p_{0,\boldsymbol{s}} : \boldsymbol{s} \in \mathcal{S}]$ denote the initial state distribution of the Markov chain. Recall that item i is initialized with no ratings. Hence, the initial state distribution can be expressed as

$$p_{0,\boldsymbol{s}} = \begin{cases} 1, & \text{if } \boldsymbol{s} = [0, \ldots, 0], \\ 0, & \text{otherwise.} \end{cases}$$

Let $\boldsymbol{p}_k$ denote the landing probability of the Markov chain, which is the state distribution of the Markov chain after k steps of transition. The landing probability $\boldsymbol{p}_k$ can be derived as $\boldsymbol{p}_k = \boldsymbol{p}_0 \prod_{\tau=1}^{k} \boldsymbol{P}^{(i,\tau)}$. We next connect the convergence of landing probability to the convergence of historical aggregate opinion.

Theorem 3. *Suppose the landing probability vector $\boldsymbol{p}_k$ satisfies that $\lim_{k\to\infty} \boldsymbol{p}_k = \boldsymbol{\pi}$, where $\boldsymbol{\pi} \in [0,1]^{|\mathcal{S}|}$. Then, we have*

$$\lim_{k\to\infty} \boldsymbol{h}_{i,k} = \sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s}) \boldsymbol{e}_{s_n}, \lim_{k\to\infty} A(\boldsymbol{h}_{i,k}) = A\left(\sum_{\boldsymbol{s}\in\mathcal{S}} \pi(\boldsymbol{s}) \boldsymbol{e}_{s_n}\right), \quad (10)$$

where $\boldsymbol{e}_{s_n}$ denotes a basis vector.

Remark: Theorem 3 states that the convergence of the landing probability implies the convergence of historical collective opinions and production quality estimator. To analyze the convergence of landing probability, the next assumptions eliminate some corner cases for the purpose of making the presentation simple.

Assumption 1. *The $\tilde{\boldsymbol{o}}_i$ satisfies $\tilde{o}_{i,m} > 0, \forall m \in \mathcal{M}$.*

Remark: Assumption 1 states that there are no redundant rating levels. Technically, if there are some redundant rating

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

levels, i.e., $\tilde{o}_{i,m} = 0$ for some $m \in \mathcal{M}$, one can eliminate them to make Assumption 1 hold, through which our analysis applies.

Assumption 2. *For each user $u \in \mathcal{U}$, the strength of assimilate-contrast effects converges, $\lim_{k\to\infty} s_{u,k} = s_{u,\infty}$, where $s_{u,\infty} \in \mathbb{R}_+$.*

Remark: Assumption 2 states that the strength of assimilate-contrast effects becomes stable as the number of ratings becomes sufficiently large.

Theorem 4. *Suppose Assumptions 1 and 2 hold. Suppose the support of the distribution $\widehat{\mathcal{D}}_B(s_{u,\infty}, r_{u,i}, \boldsymbol{s})$ is $\{1 - r_{u,i}, \ldots, M - r_{u,i}\}$ for all $\boldsymbol{s} \in \mathcal{M}^n$. Then, the Markov chain $\mathbb{M}$ is strongly ergodic, implying that there exists a unique $\boldsymbol{\pi} \in [0,1]^M$, such that $\lim_{k\to\infty} \boldsymbol{p}_k = \boldsymbol{\pi}$.*

Theorem 4 states sufficient conditions under which the landing probability converges to a unique value implying the convergence of historical aggregate opinions and product quality estimator. The condition on $\widehat{\mathcal{D}}_B(s_{u,\infty}, r_{u,i}, \boldsymbol{s})$ means that there is no redundant rating levels when the number of ratings is sufficiently large. We establish this theorem via the notion of strong ergodicity in non-stationary Markov chains. Theorem 4 only establishes the convergence result, but it does not reveal which value the landing probability will converge to. The following theorem studies this point.

Theorem 5. *Under the same condition as Theorem 4. There exists a matrix $\boldsymbol{P}^{(i,\infty)}$ such that $\lim_{k\to\infty} \boldsymbol{P}^{(i,k)} = \boldsymbol{P}^{(i,\infty)}$. Suppose $\mathbb{P}[R_{i,k} = m] > \epsilon, \forall m \in \mathcal{M}, k \in \mathbb{N}_+$, where $\epsilon \in (0,1)$. The limiting value of the landing probability $\boldsymbol{\pi} = \lim_{k\to\infty} \boldsymbol{p}_k$ satisfies*

$$\begin{cases} \pi(\boldsymbol{s}) = 0, & \text{if } \boldsymbol{s} \in \mathcal{T}, \\ \pi(\boldsymbol{s}) > 0, & \text{if } \boldsymbol{s} \in \mathcal{R}, \end{cases}$$

where $\mathcal{R} \triangleq \mathcal{M}^n$ and $\mathcal{T} \triangleq \mathcal{S} \setminus \mathcal{R}$. Furthermore, $\boldsymbol{\pi}_{\mathcal{R}} \triangleq [\pi(\boldsymbol{s}) : \boldsymbol{s} \in \mathcal{R}]$ is a unique solution of

$$\boldsymbol{\pi}_{\mathcal{R}} = \boldsymbol{\pi}_{\mathcal{R}} \boldsymbol{P}_{\mathcal{R}}, \quad \sum\nolimits_{\boldsymbol{s}\in\mathcal{R}} \pi(\boldsymbol{s}) = 1, \tag{11}$$

where $\boldsymbol{P}_{\mathcal{R}} \triangleq [P^{(i,\infty)}[\tilde{\boldsymbol{s}}|\boldsymbol{s}] : \boldsymbol{s} \in \mathcal{R}, \tilde{\boldsymbol{s}} \in \mathcal{R}]$.

Remark: Theorem 5 characterizes the structure of the limiting value of the landing probability. It states that this limiting value satisfies a linear equation array. Furthermore, this equation array is fully characterized by the limiting transition matrix $\boldsymbol{P}^{(i,\infty)}$. The limiting matrix $\boldsymbol{P}^{(i,\infty)}$ is determined by the distribution $\widehat{\mathcal{D}}_B(s_{u,\infty}, r_{u,i}, \boldsymbol{s}), \forall \boldsymbol{s} \in \mathcal{M}^n$. Through this we know how to compute the limiting value of the landing probability analytically.

Time dependent initial opinion. When users form initial opinions based on the time stamps of ratings, one can extend the above non-stationary Markov chain to study the convergence of historical aggregate opinions. One can augment the state to include the arrival time of latest n ratings. Note that in practice the arrival time stamps are discrete. One can use a probabilistic model to characterize the distribution arrival times. Through this one can have similar convergence results as the case of order dependent initial opinion formation.

3.3 Hybrid Model

Our model thus far has analyzed the case that all users form initial opinions either from the historical aggregate opinion or from a small number of latest historical ratings. In practice, it may happen that some users form initial opinions from the historical aggregate opinion, while others from a small number of historical ratings. The general case is challenging to analyze. But when one type of users dominate, i.e., of a large fraction, one can extend the above results to handle it, such as Theorem 1.

4 Application I: Rating Prediction

To demonstrate the versatility of our model, we parametrize our model to predict subsequent ratings. We apply regularized least square to infer model parameters. Extensive experiments on four datasets demonstrate that our model can improve the accuracy of LF [19], Herd [20] and the HIALF [16].

4.1 Applying Our Model to Rating Prediction

We consider the following rating prediction problem: *given a set of historical ratings of items, predict subsequent ratings of items.* To apply our model to address this problem, we next first parameterize our model and then infer model parameters.

Model parameterization. The HIALF algorithm was proposed by Zhang *et al.* [16], which is the first algorithm exploring assimilate-contrast effects to improve rating prediction. We will use HIALF as a major comparison baseline. For a fair comparison with HIALF [16], we parameterize $r_{u,i}$ via the classical latent factor (LF) model [19]:

$$r_{u,i} = \boldsymbol{x}_u^T \boldsymbol{y}_i + b_i + b_u + g. \tag{12}$$

In Equation (12), $\boldsymbol{x}_u \in \mathbb{R}^\kappa$ and $\boldsymbol{y}_i \in \mathbb{R}^\kappa$ represent vectors of latent features for user u and item i, where $\kappa \in \mathbb{N}_+$. The $b_u \in \mathbb{R}$ and $b_i \in \mathbb{R}$ model user and item bias respectively. The $g \in \mathbb{R}$ models the constant shift or residual. All parameters in Equation (12) are unknown and will be inferred from the data. Similarly, we parametrize the strength of assimilate-contrast effects as:

$$s_{u,k} = s_u \times \left(\frac{a}{1 + \exp(-b \times k)} - \frac{a}{2} \right),$$

where $s_u \in \mathbb{R}, a \in \mathbb{R}$ and $b \in \mathbb{R}$ are unknown parameters to be inferred from data. We parameterize bias function Θ as follows:

$$\Theta(\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i}) = \frac{\sum_{\ell=1}^{L} \exp(-\xi[e_\ell - (\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})]^2) v_\ell}{\sum_{\ell=1}^{L} \exp(-\xi[e_\ell - (\Gamma_u(\mathcal{H}_{i,k}) - r_{u,i})]^2)},$$

where $\{e_1, \ldots, e_L\} = \{-4, -3.5, \ldots, 3.5, 4\}$. Namely, we have $L = 17$. The parameter $\xi \in \mathbb{R}$ is a hyper parameter and $v_1, \ldots, v_L$ are unknown parameters to be inferred from data. Note this parametrization method was used in [16], and we choose it for a fair comparison with HIALF. We parameterize $\Gamma_u(\mathcal{H}_{i,k})$ use Equations (5) and (6) and its parameters are treated as hyper parameters. We will show that under simple selections of hyper parameters of $\Gamma_u(\mathcal{H}_{i,k})$, our model can outperform all the baselines. This implies that if one selects them with finer tuning, our model can

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

achieve better performance. We summarize all parameters to be inferred from data as

$$\mathcal{Z} \triangleq \left\{g, \{b_i, \boldsymbol{y}_i\}_{i\in\mathcal{I}}, \{b_u, \boldsymbol{x}_u, s_u\}_{u\in\mathcal{U}}, a, b, \{v\}_{l\in\{1,\ldots,17\}}\right\}.$$

All hyper parameters will be given before model training.

Model inference. Consider a training rating dataset, in which item i has $K_i \geq 0$ historical ratings. We aim to infer $\mathcal{Z}$ from the training rating dataset. In particular, we use regularized least square method to infer $\mathcal{Z}$:

$$\min_{\mathcal{Z}} \sum\nolimits_{i\in\mathcal{I}} \sum\nolimits_{k=1}^{K_i} \Big[\left(r_{u_{i,k},i} + b_{u_{i,k},i,k} - R_{i,k}\right)^2 + \lambda_{\text{LF}} \left(b_{u_{i,k}}^2 + b_i^2 + \|\boldsymbol{x}_{u_{i,k}}\|_2^2 + \|\boldsymbol{y}_i\|_2^2\right) + \lambda_k(a^2+b^2) + \lambda_s s_{u_{i,k}}^2 + \lambda_\Theta \sum\nolimits_{\ell=1}^{L} v_\ell^2 \Big].$$

We use the stochastic gradient descent (SGD) algorithm to learn model parameters $\mathcal{Z}$. Note that the SGD was widely used to improve training efficiency [22], [23], [24].

Rating prediction. Let $\widehat{\mathcal{Z}}$ denote the inferred model parameter set. To illustrate, suppose we are going to predict the k-th rating of item i, i.e., $R_{i,k}$. Note that $u_{k,i}$ is the user who assign the rating $R_{i,k}$. We are given the ID of the user denoted by $u_{k,i}$. Then we predict the rating $R_{i,k}$ as $\widehat{r}_{u_{k,i},i} + \widehat{b}_{u_{k,i},i,k}$, which are computed using our model with the inferred parameters $\widehat{\mathcal{Z}}$. We denote our rating prediction method as AC-RP (Assimilate-Contrast effects aware Rating Prediction).

4.2 Evaluation Settings

The dataset. We use four public datasets to evaluate the accuracy of our AC-RP method, whose overall statistics are summarized in Table 3. The dataset from Amazon was published in [25] and it contains historical ratings of movies in Amazon. The dataset from Google Local was published in [26], [27] and it contains reviews about businesses from Google Local (Google Maps). The dataset from TripAdvisor was published in [28], and it contains historical ratings of hotels in TripAdvisor. The dataset from Yelp was downloaded from the link [1] and it contains historical ratings for restaurants in Yelp.

TABLE 3: Overall statistics of four datasets

category	# items	# users	# ratings
Amazon-movie	208,321	2,088,620	4,607,047
Googlelocal	4,567,431	3,116,785	10,601,852
TripAdvisor	1,705	623,567	871,689
Yelp	60,785	366,715	1,569,264

For a fair comparison with HIALF, we use the same method as HIALF [16] to extract the training dataset and testing dataset. Similar with HIALF [16], for each item in Table 3, we only select items with a medium positive average rating, i.e., an average rating in the range $[3.9, 4.1]$ out. For each selected item we extract all its ratings and the associated users out. We then remove items with less than 50 training ratings, to avoid over-fitting. Table 4 summarizes the overall statistics of the selected data. One can observe that after this selection, some datasets still contain around two hundred thousands of users. Comparing the number of ratings with the number of users, one can observe that the rating matrix is very sparse. Similar with HIALF [16], we further aggregate users with twenty ratings (or fifty) as a big user for Googlelocal and TripAdvisor (or Amazon-movie and Yelp). Furthermore, for each selected item, we use its last 25 ratings as test ratings and use all other ratings as training ratings. We train the models on the training dataset, and validate the model on the testing dataset.

1. https://www.yelp.com/dataset

TABLE 4: Overall statistics of selected rating dataset.

category	# items	# users	# ratings
Amazon-movie	1,118	198,209	271,295
Googlelocal	445	38,553	53,740
TripAdvisor	309	134,484	150,201
Yelp	736	82,509	157,463

Comparison baseline & metrics. We use the root mean squared error (RMSE) to quantify the testing accuracy:

$$RMSE = \sqrt{\frac{\sum_{i\in\mathcal{I}} \sum_{k=K_i-24}^{K_i} (\widehat{r}_{u_{k,i},i} + \widehat{b}_{u_{i,k},i,k} - R_{i,k})^2}{25|\mathcal{I}|}}, \quad (13)$$

where $\widehat{b}_{u_{i,k},i,k}$ denotes an estimation of $b_{u_{i,k},i,k}$, which is computed using our model with the inferred parameters $\widehat{\mathcal{Z}}$. We compare our AC-RP method with the following three baselines.

- HIALF [16]. The HIALF algorithm was proposed by Zhang *et al.* [16], which is the first algorithm exploring assimilate-contrast effects to improve rating prediction.
- Herd [20]. The Herd algorithm was proposed by Zhang *et al.*, which explores herding effects to improve rating prediction.
- LF [19]. The LF is widely used [29], [30], [16]. Note that the LF is the baseline for our AC-RP and HIALF, because both of them use LF to parameterize the unbiased rating of users $r_{u,i}$.

Parameter setting. Table 5 shows hyper parameters that are fixed throughout experiments. These hyper parameters are also used in HIALF [16], we thus select them following a similar principle in HIALF [16], in order to prevent over fitting. Following previous work [29], [30], we choose the dimension of latent features as $\kappa = 5$. We set $\xi = 10$.

TABLE 5: Hyper parameters fixed throughout experiments.

λ_{LF}	λ_k	λ_s	λ_Θ	κ	ξ
0.1	0.0001	0.01	0.0005	5	10

For other hyper parameters associated with the initial opinion function $\Gamma_u(\mathcal{H}_{i,k})$, we do not present them here. The reason is that it is redundant to present them here. We study different instances of $\Gamma_u(\mathcal{H}_{i,k})$. Each instance has hyper parameters, and we select three values of each hyper parameter to study its impact. To save some space, we omit these hyper parameters here.

4.3 Rating Recency for Rating Prediction

In this section, we evaluate the benefit of exploiting rating recency for rating prediction tasks. Recall that under rating recency, users form initial opinions from n latest ratings.

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

We consider the initial opinion formation model derived in Equation (6).

Impact of n. The initial opinion is the simple average of n latest ratings, i.e., $\eta_{k,j} = 1, \forall j$. Table 6 shows the RMSE of LF, Herd, HIALF and our AC-RP method. In Table 6, the column $n = 20, 40, 80$ corresponds to the RMSE of our AC-RP method with $n = 20, 40, 80$. One can observe that when $n = 20$, our method has a smaller RMSE than LF, Herd and HIALF. This statement also holds when $n = 40, 80$. Namely, under some simple selections of n, our method has a higher rating prediction accuracy than LF, Herd and HIALF. Note that this improvement of rating prediction accuracy by exploiting rating recency is supported by survey studies, which identified that users tend to read a small number of latest reviews or ratings to form initial opinions [9]. The RMSE of our method varies as we increase n from 20 to 80, This implies that n is an important factor for the rating prediction accuracy and one needs to exploit the rating recency carefully for rating prediction. The above improvement on the rating prediction accuracy is achieved at simple selections of n. One may further improve the rating prediction accuracy by finer tuning of n.

TABLE 6: Impact of n on rating prediction ($\eta_{k,j}$=1, $\forall j$).

category	Herd	LF	HIALF
Amazon-movie	1.4871	1.2342	1.2209
Googlelocal	1.1412	1.0594	1.0616
TripAdvisor	1.1838	0.9266	0.9208
Yelp	1.4200	1.2008	1.2006
category	**n=20**	**n=40**	**n=80**
Amazon-movie	1.2060	1.2060	1.2122
Googlelocal	1.0548	1.0578	1.0581
TripAdvisor	0.9164	0.9145	0.9154
Yelp	1.1956	1.1958	1.1968

Impact of arrival order. The are two types of weights on ratings, i.e., based on arrival order and based on arrival time stamp of ratings. Here we study the weight which is based on arrival order. We fix $n = 20$. To study the impact of arrival order, we consider three types of weights associated with the arrival order of ratings, i.e., exponential, polynomial and logarithmic in the arrival order, which are stated in Table 7. Table 7 shows that the RMSE of our AC-RP method under these three types of weights. Consider that the weight $\eta_{k,j}$ is exponential in the arrival order, i.e., $\eta_{k,j} = \exp(-\zeta j)$. One can observe that as we vary the parameter ζ of $\eta_{k,j} = \exp(-\zeta j)$ from 0.01 to 0.0001, the RMSE can be further reduced over the unweighted case, i.e., $\zeta = 0$. Similar observations can be found when the weight $\eta_{k,j}$ is polynomial or logarithmic in the arrival order of ratings. This implies that one can further improve the rating prediction accuracy via tuning the weight of ratings based on the arrival order. This improvement on rating prediction accuracy is supported by that users tend to assign larger weights to more recent ratings [9] Furthermore, this improvement is achieved at simple selections of the weight of ratings. One can further improve the rating prediction accuracy by finer tuning of weights.

Due to page limit, we present experiments on the impact of arrival time stamp and aggregate opinion heterogeneity in our supplementary file.

TABLE 7: Impact of arrival order on rating prediction (n=20).

		$\eta_{k,j} = \exp(-\zeta j)$		
category	**ζ=0**	**ζ=0.01**	**ζ=0.001**	**ζ=0.0001**
Amazon-movie	1.2060	**1.2053**	1.2109	**1.2059**
Googlelocal	1.0548	**1.0544**	1.0551	1.0569
TripAdvisor	0.9164	0.9166	**0.9162**	**0.9160**
Yelp	1.1956	1.1964	1.1960	**1.1946**

$\eta_{k,j} = j^{-\zeta}$			$\eta_{k,j} = 1/[\zeta \ln(j+1)]$		
ζ=0.5	**ζ=1**	**ζ=2**	**ζ=0.5**	**ζ=1**	**ζ=2**
1.2079	1.2110	1.2151	**1.2043**	1.2092	1.2083
1.0539	**1.0536**	**1.0538**	**1.0546**	**1.0546**	**1.0538**
0.9163	0.9177	0.9221	**0.9162**	0.9177	**0.9163**
1.1941	**1.1942**	1.1977	**1.1948**	**1.1937**	**1.1938**

5 Application II: Recommendation

In this section, we further apply the parametrization of our model (please refer to Section 4) to recommendation tasks. Extensive experiments on four datasets demonstrate that our model can improve the recommendation accuracy of LF [19] and the HIALF [16].

5.1 Applying Our Model to Recommendation

We consider the following recommendation problem: *given a set of historical ratings of each item, predict intrinsic ratings of subsequent users toward items.* Note that the intrinsic rating reveals a user's true preference toward an item and it serves as a generic metric for many recommendation tasks. For example, for top-k recommendations, one just needs to rank items based on the intrinsic ratings of a user to items and then select top k items. We first parameterize our model using the same method as Section 4. Then use the same method as Section 4 to infer parameters of our model. We still use $\widehat{\mathcal{Z}}$ to denote the inferred model parameter set. To illustrate, suppose we are going to predict the intrinsic rating of user u toward item i, i.e., $r_{u,i}$. We estimate the intrinsic rating $r_{u,i}$ as

$$\widehat{r}_{u,i} = \widehat{g} + \widehat{b}_i + \widehat{b}_u + \widehat{\boldsymbol{x}}_u^T \widehat{\boldsymbol{y}}_i, \tag{14}$$

where $\widehat{g}, \widehat{b}_i, \widehat{b}_u, \widehat{\boldsymbol{x}}_u$ and $\widehat{\boldsymbol{y}}_i$ are inferred model parameters. We denote our recommendation method as AC-Rec (Assimilate-Contrast effects aware Recommendation).

5.2 Evaluation Settings

We use the same datasets as Section 4. We use the same method as Section 4 to process data, except: (1) we do not aggregate users with rating less than a certain number as a big user, as here we aim for personalized intrinsic rating; (2) we take the set of ratings without historical ratings as the ground truth (e.g., the first rating of each item) for testing and we train our models with the rest of ratings. Let $\mathcal{I}_{u,\text{test}} \subseteq \mathcal{I}$ denote a set of items in the testing dataset.

We consider the following three performance metrics:

- **RMSE:** RMSE is a widely used to evaluate the performance of recommendation algorithms [16]. Note that each

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

item in the testing data set only has one rating, i.e., the first rating $R_{i,1}$. We thus express the RMSE as:

$$RMSE = \sqrt{\frac{\sum_{i\in\mathcal{I}}(\widehat{r}_{u_{i,1},i} - R_{i,1})^2}{|\mathcal{I}_{u,\text{test}}|}}. \quad (15)$$

Smaller RMSE implies a higher recommendation accuracy.

- **Relative Cumulative Reciprocal Rank (RCRR).** RCRR [16] is a ranking based metric. For each item $u \in \mathcal{U}$, we construct a ranking list, in which items are ranked based on the estimated rating $\widehat{r}_{u,i}$ (expressed in Equation (14)) in descending order. Let $\text{Rank}_{u,i}$ denote the rank of item i in user u's ranking list. For example, if item i is ranked top one, then $\text{Rank}_{u,i} = 1$. For products in the testing dataset, denote the set of items adopted by user u as $\mathcal{I}_{u,\text{test}} \subseteq \mathcal{I}$. The RCRR with respect to user u as:

$$RCRR_u \triangleq \frac{1}{|\mathcal{I}_{u,\text{test}}|}\sum\nolimits_{i\in\mathcal{I}_{u,\text{test}}} \frac{1}{\text{Rank}_{u,i}/|\mathcal{I}|}$$

Larger $RCRR_u$ implies that products adopted by user u are more likely to be ranked higher, i.e., a higher accuracy. Let $\mathcal{U}_{\text{test}}$ denote a set of all users in the testing dataset. We aim at evaluate the average relative cumulative reciprocal rank over all users in the testing dataset:

$$\overline{RCRR} \triangleq \frac{1}{|\mathcal{U}_{\text{test}}|}\sum\nolimits_{u\in\mathcal{U}_{\text{test}}} RCRR_u.$$

- **Normalized discounted cumulative gain (NDCG).** NDCG is also a ranking based metric. For brevity, one can refer to [21] for the formula of this metric. For the TripAdvisor metric, we consider the NDCG@5 which measures the accuracy of top-5 ranking list. For the other datasets, we consider the NDCG@10 which measures the accuracy of top-10 ranking list. The reason is that in TripAdvisor, each user only rates a small number of items.

We compare our AC-Rec method with LF and HIALF. Similar with [16], we do not compare with the Herd model, because it is not developed for recommendation. Due to that the rating matrix is sparse, we select some large regularization hyper parameters to prevent over fitting following similar principle in HIALF [16]. Table 8 shows the regularization hyper parameters. All the other hyper parameters are the same as Sec. 4. Due to similar reason with Section 4 we omit hyper parameters associated with the initial opinion function $\Gamma_u(\mathcal{H}_{i,k})$.

TABLE 8: Regularization hyper parameters.

category	λ_{LF}	λ_κ	λ_s	λ_Θ
Amazon-movie	10	0.01	1	0.05
Googlelocal	100	0.1	10	0.5
TripAdvisor	100	0.1	10	0.5
Yelp	100	0.1	10	0.5

5.3 Rating Recency for Recommendation

In this section, we evaluate the benefit of exploiting rating recency for recommendation tasks. Recall that under rating recency, users form initial opinions from n latest ratings. We consider the initial opinion formation model derived in Equation (6).

Impact of n. The initial opinion is the simple average of n latest ratings, i.e., $\eta_{k,j} = 1, \forall j$. Table 9 shows the RMSE, $\overline{RCRR}$ and NDCG of LF, HIALF and our AC-Rec method. In Table 9, the column $n = 20, 40, 80$ corresponds to the RMSE, $\overline{RCRR}$ and NDCG of our AC-Rec method with $n = 20, 40, 80$. One can observe that when $n = 20$, our method has a smaller RMSE than LF and HIALF. This statement also holds when $n = 40, 80$. Namely, under some simple selections of n, our method outperforms LF and HIALF with respect to RMSE. Similarly, Table 9 also shows that our AC-Rec method outperforms LF and HIALF with respect to $\overline{RCRR}$ and NDCG. Note that this improvement of recommendation accuracy by exploiting rating recency is supported by that users tend to read a small number of latest reviews or ratings to form initial opinions [9]. The RMSE, $\overline{RCRR}$ and NDCG of our method varies as we increase n from 20 to 80. This implies that n is an important factor for the recommendation accuracy. The above improvement on the recommendation accuracy is achieved at simple selections of n. One may further improve the recommendation accuracy by finer tuning of n. In this experiment, the initial opinion is the simple average of n latest ratings. Users may have different weights to different ratings, i.e., more weights on more recent ratings. We next explore this direction.

TABLE 9: Impact of n on recommendation ($\eta_{k,j} = 1$).

category	LF	HIALF	n=20	n=40	n=80
			RMSE		
Amazon-movie	1.1503	1.1491	1.1454	1.1484	1.1486
Googlelocal	1.2021	1.1913	1.1865	1.1921	1.1749
TripAdvisor	1.2639	1.2563	1.2550	1.2684	1.2512
Yelp	0.9371	0.9364	0.9327	0.9359	0.9354
			$\overline{RCRR}$		
Amazon-movie	1.9619	1.9820	2.0230	2.0538	2.0895
Googlelocal	1.3953	1.7649	1.9628	1.9622	1.9582
TripAdvisor	1.8425	1.8464	1.8471	1.8493	1.8500
Yelp	1.4288	1.8380	2.0455	2.0502	2.0636
			NDCG		
Amazon-movie	0.9957	0.9957	0.9967	0.9969	0.9966
Googlelocal	0.9943	0.9946	0.9971	0.9960	0.9958
TripAdvisor	0.9990	0.9990	0.9990	0.9990	0.9990
Yelp	0.9900	0.9897	0.9906	0.9911	0.9919

Impact of arrival order. Here we study the weight which is based on arrival order of ratings. We fix $n = 20$. We consider three types of weights associated with the arrival order of ratings, i.e., exponential, polynomial and logarithmic in the arrival order, which are stated in Table 10. Table 10 shows the RMSE, $\overline{RCRR}$ and NDCG of our AC-Rec method under these three types of weights. Consider that the weight $\eta_{k,j}$ is exponential in the arrival order, i.e., $\eta_{k,j} = \exp(-\zeta j)$. One can observe that as we vary the parameter ζ of $\eta_{k,j} = \exp(-\zeta j)$ from 0.01 to 0.0001, the RMSE can be further reduced over the unweighted case, i.e., $\zeta = 0$. Similar observations can be found when the weight $\eta_{k,j}$ is polynomial or logarithmic in the arrival order of ratings. This implies that one can further improve the RMSE via tuning the weight of ratings based on the arrival order. Similar improvement on the recommendation accuracy can be found with respect to $\overline{RCRR}$ and NDCG. This improvement on recommendation accuracy is supported by that users tend to assign larger weights to more recent ratings [9].

Impact of arrival time stamp. We fix $n = 20$. We consider

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

TABLE 10: Impact of arrival order on recommendation accuracy.

		$\eta_{k,j} = \exp(-\zeta j)$			$\eta_{k,j} = j^{-\zeta}$			$\eta_{k,j} = 1/[\zeta ln(j+1)]$		
category	**ζ=0**	**ζ=0.01**	**ζ=0.001**	**ζ=0.0001**	**ζ=0.5**	**ζ=1**	**ζ=2**	**ζ=0.5**	**ζ=1**	**ζ=2**
					RMSE					
Amazon-movie	1.1454	1.1459	1.1477	1.1484	1.1477	1.1479	1.1502	1.1467	**1.1436**	1.1488
Googlelocal	1.1865	**1.1859**	**1.1841**	**1.1831**	**1.1842**	**1.1858**	**1.1864**	**1.1839**	1.1871	1.1865
TripAdvisor	1.2550	1.2553	1.2565	**1.2546**	**1.2522**	**1.2531**	**1.2496**	1.2553	**1.2525**	**1.2427**
Yelp	0.9327	0.9340	0.9331	0.9343	0.9340	0.9361	0.9344	0.9328	0.9328	**0.9317**
					$\overline{RCRR}$					
Amazon-movie	2.0230	2.0728	2.0460	2.0690	2.0618	2.0675	2.0690	2.0670	2.0605	2.0513
Googlelocal	1.9628	1.9603	1.9607	1.9618	1.9532	1.9539	1.9599	1.9609	1.9673	1.9625
TripAdvisor	1.8471	1.8496	1.8501	1.8491	1.8490	1.8485	1.8490	1.8491	1.8496	1.8496
Yelp	1.0455	2.0683	1.9703	2.0757	2.0594	2.0539	2.0329	2.0464	2.0464	2.0459
					NDCG					
Amazon-movie	0.9957	0.9973	0.9974	0.9972	0.9968	0.9962	0.9967	0.9963	0.9963	0.9963
Googlelocal	0.9946	0.9972	0.9960	0.9954	0.9970	0.9968	0.9966	0.9957	0.9955	0.9961
TripAdvisor	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	
Yelp	0.9897	0.9916	0.9901	0.9911	0.9914	0.9897	0.9919	0.9928	0.9917	0.9943

three types of weights associated with the arrival time stamp of ratings, i.e., exponential, polynomial and logarithmic in the arrival time stamp, which are stated in Table 11. Table 11 shows that the RMSE, $\overline{RCRR}$ and NDCG of our AC-Rec method under these three types of weights with respect to time stamp. Consider that the weight $\eta_{k,j}$ is exponential in the arrival time stamp, i.e., $\eta_{k,j} = \exp(-\rho(t_u - t_{i,k+1-j})^{0.4})$. One can observe that as we vary the parameter ρ of $\eta_{k,j} = \exp(-\rho(t_u - t_{i,k+1-j})^{0.4})$ from 0.01 to 0.0001, the RMSE can be further reduced over the unweighted case, i.e., $\rho = 0$. Similar observations can be found when the weight $\eta_{k,j}$ is polynomial or logarithmic in the arrival time stamp of ratings. This implies that one can further improve the RMSE via tuning the weight of ratings based on the arrival time stamp. Similar, Table 11 also shows that one can further improve the $\overline{RCRR}$ and NDCG via tuning the weight of ratings based on the arrival time stamp. This improvement on recommendation accuracy is also supported by that users tend to assign larger weights to more recent ratings [9].

Due to page limit, we present the impact of aggregate opinion heterogeneity in the supplementary file.

6 EXTENSIONS

In this section, we show the generality of our proposed model. We show that our model can be applied to enhance other models beyond the LF model. In particular, we show that our proposed initial opinion formation model can be applied to enhance one of the latest rating behavior model, i.e., the spiral process model [21].

Spiral process model (SPM) [21] is one of the latest rating behavior model. One of its key component is the perceived global opinion climate of a product, which is modeled as the historical average rating of a product. We enhance the SPM model to capture rating recency or aggregate opinion heterogeneity by replacing the historical average rating with our proposed initial opinion function. Please refer to [21] for other components of the SPM model and the training of SPM model. We apply the experiment methodology of [21] on the dataset presented in Table 4. For brevity, we only present the experiment results with respect to the $\overline{RCRR}$ metric.

Table 12 shows that the $\overline{RCRR}$ of the enhanced variant of the SPM model by capturing rating recency. One can observe that these enhanced variants outperforms the original SPM model except over the Amazon-movie dataset. Table 13 shows the $\overline{RCRR}$ of the enhanced variant of the SPM model by capturing aggregate opinion heterogeneity. One can observe that these enhanced variants outperform the original SPM model.

7 RELATED WORK

The rating bias was studied extensively [31], [30]. A number of works studied the rating bias caused by item category [31], system interfaces [32], recommendation algorithms [33], evolving dynamics of user preferences [24], or the change in user experience [30], etc. Some recent works showed that historical ratings influence user ratings, resulting in biased ratings [11], [12], [13], [14]. This paper falls into the research line of rating bias caused by historical ratings.

The study on rating bias caused by historical ratings was initiated by several controlled experiments [11], [12], [13]. Those controlled experiments showed the existence of this rating bias. They also revealed one pattern of this rating bias interpreted by the herding effects, which states that users tend to assign higher ratings when higher historical ratings are displayed to them. Two notable models that quantify this herding effects are: (1) polynomial regression based model [34]; (2) additive generative based model [20]. Empirical studies showed that these two models have sufficient expressive power to capture pattern of herding effects. However, these two models are not neat enough to support analytical studies of the evolving dynamics of aggregate ratings under herding effects. Xie *et al.* [35] filled in this gap by proposing a neater model while retaining certain expressive power, i.e., the linear model, of herding effects, and provided theoretical analysis on the evolving dynamics of aggregate ratings. Compared to these works, we study a more sophisticated pattern of rating bias interpreted by

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

TABLE 11: Impact of arrival time on recommendation.

		$\eta_{k,j}=\exp(-\rho(t_u-t_{i,k+1-j})^{0.4})$			$\eta_{k,j}=(t_u-t_{i,k+1-j})^{-\rho}$			$\eta_{k,j}=1/\rho\ln(t_u-t_{i,k+1-j}+1)$		
category	ρ **=0**	ρ**=0.01**	ρ**=0.001**	ρ**=0.0001**	ρ**=0.2**	ρ**=0.4**	ρ**=0.8**	ρ**=0.2**	ρ**=0.4**	ρ**=0.8**
						RMSE				
Amazon-movie	1.1454	1.1476	1.1499	1.1476	1.1460	1.1486	1.1466	**1.1453**	1.1506	1.1474
Googlelocal	1.1865	**1.1846**	1.1901	**1.1835**	**1.1802**	1.1909	1.1884	**1.1864**	**1.1851**	**1.1827**
TripAdvisor	1.2550	1.2569	1.2558	1.2590	**1.2463**	1.2579	1.2586	1.2585	**1.2493**	**1.2494**
Yelp	0.9327	0.9347	0.9341	0.9350	0.9345	0.9329	0.9357	0.9332	0.9346	0.9354
						$\overline{RCRR}$				
Amazon-movie	2.0230	2.0547	2.0547	2.0609	2.0524	2.0613	2.0636	2.0582	2.0565	2.0626
Googlelocal	1.9628	1.9600	1.9487	1.9591	1.9583	1.9582	1.9532	1.9591	1.9552	1.9581
TripAdvisor	1.8471	1.8479	1.8509	1.8493	1.8489	1.8492	1.8486	1.8486	1.8499	1.8485
Yelp	1.0455	2.0506	2.0055	2.0656	2.0660	2.0758	2.0451	2.0680	1.9702	2.0414
						NDCG				
Amazon-movie	0.9957	0.9963	0.9972	0.9967	0.9967	0.9966	0.9967	0.9960	0.9966	0.9965
Googlelocal	0.9946	0.9958	0.9975	0.9957	0.9965	0.9958	0.9962	0.9967	0.9961	0.9962
TripAdvisor	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990	0.9990
Yelp	0.9897	0.9909	0.9916	0.9908	0.9905	0.9920	0.9903	0.9918	0.9913	0.9909

TABLE 12: Impact of n on recommendation accuracy of variants of SPM ($\eta_{k,j}=1$, $\overline{RCRR}$)

category	**SPM**	n**=20**	n**=40**	n**=80**
Amazon-movie	7.3290	6.3701	5.7046	4.3163
Googlelocal	0.4904	**0.5034**	0.3295	0.3487
TripAdvisor	1.3479	0.2832	**1.3774**	1.1054
Yelp	1.4488	**2.3578**	1.4686	**2.0569**

TABLE 13: Exponential weights in opinion level for recommendation accuracy of variants of SPM ($n=20, \overline{RCRR}$).

		$\beta_m=\exp(-\gamma m)$		
category	**SPM**	γ**=0.1**	γ**=0.01**	γ**=0.001**
Amazon-movie	7.3290	**41.019**	**43.205**	2.6254
Googlelocal	0.4904	**0.5002**	**0.5034**	0.3826
TripAdvisor	1.3479	0.9537	1.2283	**1.5802**
Yelp	1.4488	**2.2268**	**1.6879**	**1.6295**

the assimilate-contrast effects. Built on the non-stationary stochastic difference equation and non-stationary high order Markov chain, our model supports analytical studies of the evolving dynamics of aggregate ratings under assimilate-contrast effects.

Via extensive data analysis, Zhang *et al.* [15], [16] identified a more sophisticated pattern of rating bias interpreted by the assimilate and contrast phenomenon. This rating bias states that when a product is slightly over or under rated, users tend to imitate historical ratings, while over or under rated to much, users tend to deviate a lot from the historical ratings. They captured assimilate-contrast effects via adding a bias term to the latent factor model. Experiment studies showed that their model is more accurate than the herding effects model [20] in characterizing rating behavior. We further study the assimilate-contrast effects to address several limitations of the model by Zhang *et al.* [15], [16]. Our model captures more important factors like rating recency than their model, achieving a higher accuracy. Built on the non-stationary stochastic difference equation and non-stationary high order Markov chain, our model is still neat enough to support analytical studies of the evolving dynamics of aggregate ratings under assimilate-contrast effects. Note that their model does not support analytical studies of the evolving dynamics of aggregate ratings.

Recent notable works on rating bias caused by historical ratings include the following ones. Via extensive data analysis, Lin *et al.* [21] identified the spiral of silence phenomenon in online ratings. They quantified spiral of silence effects via combing the matrix factorization model and the Gaussian mixture model. Inspired by the theory of message-based persuasion in psychology [36], [37], Xie *et al.* [28] developed a homogeneous Markov chain model to quantify the rating bias caused by persuasion effects of historical ratings. Different from the assimilate-contrast effects, spiral of silence characterizes the missing pattern of ratings, while message-based persuasion characterizes how users imitate recent ratings. Note that these models do not support analytical studies of the evolving dynamics of aggregate ratings, while our model supports them. We achieve this by building our model on the non-stationary stochastic difference equation and non-stationary high order Markov chain.

8 CONCLUSION

This paper proposes a mathematical model to quantify assimilate-contrast effects in online rating systems. Our model captures important factors like rating recency that are missed by previous models. Furthermore, our model attains a good balance between model complexity and model accuracy, such that it is neat enough to support analytical study of assimilate-contrast effects. We apply stochastic approximation theory to show that when the rating bias satisfies mild contraction conditions, the aggregate rating converges under aggregate opinion heterogeneity. . We also apply non-stationary Markov chain theory to show that when the strength of assimilate-contrast effects satisfies mild stable conditions, the aggregate rating converges under rating recency. We also apply our model rating prediction tasks and recommendation tasks. Extensive experiment results show that our model can improve the accuracy of rating prediction and recommendation over previous models significantly under various metrics like RMSE, NDCG, etc. We also demonstrate the flexibility of our model.

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

ACKNOWLEDGMENT

This research was supported in part by the National Natural Science Foundation of China (No. 61902042, No. 62176029), Chongqing Talents: Exceptional Young Talents Project (cstc2021ycjhbgzxm0195), the Chinese Academy of Sciences "Light of West China" Program, the Key Cooperation Project of Chongqing Municipal Education Commission (HZ2021008, HZ2021017), and the "Fertilizer Robot" project of Chongqing Committee on Agriculture and Rural Affairs.

REFERENCES

[1] P. Resnick and H. R. Varian, "Recommender systems," *Communications of the ACM*, vol. 40, no. 3, pp. 56–58, 1997.

[2] G. Lackermair, D. Kailer, and K. Kanmaz, "Importance of online product reviews from a consumer's perspective," *Advances in Economics and Business*, vol. 1, no. 1, pp. 1–5, 2013.

[3] M. Li, L. Huang, C.-H. Tan, and K.-K. Wei, "Helpfulness of online product reviews as seen by consumers: Source and content features," *International Journal of Electronic Commerce*, vol. 17, no. 4, pp. 101–136, 2013.

[4] Y. Zhao, S. Yang, V. Narayan, and Y. Zhao, "Modeling consumer learning from online product reviews," *Marketing Science*, vol. 32, no. 1, pp. 153–169, 2013.

[5] J. Berger, *Bad Reviews Can Boost Sales. Here's Why*. Harvard Business Review, 2012.

[6] M. Luca, "Reviews, reputation, and revenue: The case of yelp. com," 2016.

[7] S. Zaroban, *Product reviews boost revenue per online visit 62%*. Digital Commerce, 2015.

[8] BrightLocal, *Local Consumer Review Survey*. BrightLocal, 2016.

[9] S. Rudolph, *50 Stats You Need to Know About Online Reviews [Infographic]*. Business 2 Community, 2016.

[10] K. Shrestha, *50 Stats You Need to Know About Online Reviews [Infographic]*. Vendasta, 2016.

[11] G. Adomavicius, J. C. Bockstedt, S. P. Curley, and J. Zhang, "Understanding effects of personalized vs. aggregate ratings on user preferences." in *IntRS@ RecSys*, 2016, pp. 14–21.

[12] L. Muchnik, S. Aral, and S. J. Taylor, "Social influence bias: A randomized experiment," *Science*, vol. 341, no. 6146, pp. 647–651, 2013.

[13] M. J. Salganik, P. S. Dodds, and D. J. Watts, "Experimental study of inequality and unpredictability in an artificial cultural market," *science*, vol. 311, no. 5762, pp. 854–856, 2006.

[14] T. Weninger, T. J. Johnston, and M. Glenski, "Random voting effects in social-digital spaces: A case study of reddit post submissions," in *Proceedings of the 26th ACM conference on hypertext & social media*, 2015, pp. 293–297.

[15] X. Zhang, J. Zhao, and J. C. Lui, "Modeling the assimilation-contrast effects in online product rating systems: Debiasing and recommendations," in *Proceedings of the Eleventh ACM Conference on Recommender Systems*, 2017, p. 98–106.

[16] X. Zhang, H. Xie, J. Zhao, and J. C. Lui, "Understanding assimilation-contrast effects in online rating systems: modelling, debiasing, and applications," *ACM Transactions on Information Systems (TOIS)*, vol. 38, no. 1, pp. 1–25, 2019.

[17] B. Marlin, R. S. Zemel, S. Roweis, and M. Slaney, "Collaborative filtering and the missing at random assumption," *arXiv preprint arXiv:1206.5267*, 2012.

[18] H. Steck, "Training and testing of recommender systems on data missing not at random," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: Association for Computing Machinery, 2010, p. 713–722.

[19] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook," in *Recommender systems handbook*. Springer, 2011, pp. 1–35.

[20] T. Wang, D. Wang, and F. Wang, "Quantifying herding effects in crowd wisdom," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: Association for Computing Machinery, 2014, p. 1087–1096.

[21] C. Lin, D. Liu, H. Tong, and Y. Xiao, "Spiral of silence and its application in recommender systems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 6, pp. 2934–2947, 2022.

[22] R. Forsati, I. Barjasteh, F. Masrour, A.-H. Esfahanian, and H. Radha, "Pushtrust: An efficient recommendation algorithm by leveraging trust and distrust relations." New York, NY, USA: Association for Computing Machinery, 2015, p. 51–58.

[23] Y. Koren, "Factorization meets the neighborhood: A multifaceted collaborative filtering model," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: Association for Computing Machinery, 2008, p. 426–434.

[24] ——, "Collaborative filtering with temporal dynamics," in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY, USA: Association for Computing Machinery, 2009, p. 447–456.

[25] J. McAuley, R. Pandey, and J. Leskovec, "Inferring networks of substitutable and complementary products." New York, NY, USA: Association for Computing Machinery, 2015, p. 785–794.

[26] R. He, W.-C. Kang, and J. McAuley, "Translation-based recommendation." New York, NY, USA: Association for Computing Machinery, 2017, p. 161–169.

[27] R. Pasricha and J. McAuley, "Translation-based factorization machines for sequential recommendation," ser. RecSys '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 63–71.

[28] H. Xie, M. Zhong, Y. Li, and J. C. S. Lui, "Understanding persuasion cascades in online product rating systems: Modeling, analysis, and inference," *ACM Transactions Knowledge Discovery from Data*, vol. 15, no. 3, 2021.

[29] G. Ling, M. R. Lyu, and I. King, "Ratings meet reviews, a combined approach to recommend." New York, NY, USA: Association for Computing Machinery, 2014, p. 105–112.

[30] J. J. McAuley and J. Leskovec, "From amateurs to connoisseurs: Modeling the evolution of user expertise through online reviews," in *Proceedings of the 22nd International Conference on World Wide Web*. New York, NY, USA: Association for Computing Machinery, 2013, p. 897–908.

[31] F. Guo and D. B. Dunson, "Uncovering systematic bias in ratings across categories: A bayesian approach," in *Proceedings of the 9th ACM Conference on Recommender Systems*. New York, NY, USA: Association for Computing Machinery, 2015, p. 317–320.

[32] D. Cosley, S. K. Lam, I. Albert, J. A. Konstan, and J. Riedl, "Is seeing believing? how recommender system interfaces affect users' opinions," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2003, p. 585–592.

[33] P. Shafto and O. Nasraoui, "Human-recommender systems: From benchmark data to benchmark cognitive models," in *Proceedings of the 10th ACM Conference on Recommender Systems*. New York, NY, USA: Association for Computing Machinery, 2016, p. 127–130.

[34] S. Krishnan, J. Patel, M. J. Franklin, and K. Goldberg, "A methodology for learning, analyzing, and mitigating social influence bias in recommender systems," in *Proceedings of the 8th ACM Conference on Recommender Systems*. New York, NY, USA: Association for Computing Machinery, 2014, p. 137–144.

[35] H. Xie and M. Zhong, "Robust product rating rules against herding effects: Theory and applications," in *IEEE International Conference on Data Mining*, 2020, pp. 1352–1357.

[36] C. I. Hovland, I. L. Janis, and H. H. Kelley, "Communication and persuasion; psychological studies of opinion change." 1953.

[37] W. Wood, "Attitude change: Persuasion and social influence," *Annual review of psychology*, vol. 51, no. 1, pp. 539–570, 2000.

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in IEEE Transactions on Knowledge and Data Engineering. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TKDE.2023.3292352

PLACE PHOTO HERE

Hong Xie is currently a researcher at Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences. He received Ph.D. degree in the Department of Computer Science and Engineering at The Chinese University of Hong Kong in 2015, proudly under the supervision of Prof. John C.S. Lui. He received his B.Eng. degree from the School of Computer Science and Technology at The University of Science and Technology of China in 2010. Hong Xie was a postdoctoral research fellow at Department of Computing Science and Engineering, The Chinese University of Hong Kong (CUHK) hosted by Prof. John C.S. Lui, and a postdoctoral research fellow at School of Computing, Nationial University of Singapore hosted by Prof. Richard T.B. Ma . He was also a faculty member at Chongqing University. He is a member of CCF, a member of IEEE and a member of ACM.

PLACE PHOTO HERE

Mingze Zhong completed his Master degree at College of Computer Science, Chongqing University, under the supervision of Prof. Hong Xie. He received his B.Eng. degree from Chongqing University of Posts and Telecommunications, China. His research interests include machine learning and recommendation systems. He is a research assistant at Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences.

PLACE PHOTO HERE

Xiaoying Zhang is currently a researcher at Bytedance. She completed her Ph.D. degree in the Department of Computer Science and Engineering at The Chinese University of Hong Kong, under the supervision of Prof. John C.S. Lui. She received his B.Eng. degree from the School of Computer Science and Technology at The University of Science and Technology of China. Her research interests include data science and recommender systems.

PLACE PHOTO HERE

Xiaoyu Shi received the BS degree in computer science from the PLA Information Engineering University, Zhengzhou, China, in 2007, and the PhD degree in computer science from the University of Electronic Science and Technology of China, Chengdu, China, in 2015. He joined the Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China, in 2015, as a associate professor of computer science and engineering. His research interests include recommender system, cloud computing, artificial intelligence and big data applications.

PLACE PHOTO HERE

Jiang Zhong is currently a full professor and associate dean at college computer science, Chongqing University. He received the Ph.D degree in computer science from Chongqing University. His research interest include data mining, recommendation systems, knowledge graph, etc.

PLACE PHOTO HERE

Mingsheng Shang Mingsheng Shang received the PhD degree in computer science from the University of Electronic Science and Technology of China (UESTC). He joined the Chongqing Institute of Green and Intelligent Technology, Chinese Academy of Sciences, Chongqing, China, in 2015, and is currently a professor of computer science and engineering. His research interests include data mining, complex networks and cloud computing.

Authorized licensed use limited to: CHENGDU BRANCH OF THE NATIONAL SCIENCE LIBRARY CAS. Downloaded on July 24,2023 at 08:33:34 UTC from IEEE Xplore. Restrictions apply.
© 2023 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.