Unbiased Recommender Learning from Biased Graded Implicit Feedback

Yuta Saito
Cornell University
ys552@cornell.edu

Suguru Yaginuma
MC Digital, Inc.
suguru@mcdigital.jp

Taketo Naito
SMN Corporation
taketo_naito@so-netmedia.co.jp

Kazuhide Nakata
Tokyo Institute of Technology
nakata.k.ac@m.titech.ac.jp

ABSTRACT

Binary user-behavior logs such as clicks or views, called *implicit feedback*, are often used to build recommender systems because of its general availability in real practice. Most existing studies formulate implicit feedback as binary relevance feedback. However, in numerous applications, implicit feedback is observed not only as a binary indicator but also in a graded form, such as the number of clicks and the dwell time observed after a click, which we call the *graded implicit feedback*. The grade information should appropriately be utilized, as it is considered a more direct relevance data compared to the mere implicit feedback. However, a challenge is that the grade information is observed only for the user–item pairs with implicit feedback, whereas the grade information is unobservable for the pairs without implicit feedback. Moreover, graded implicit feedback for some user–item pairs is more likely to be observed than for others, resulting in the *missing-not-at-random* (MNAR) problem. To the best of our knowledge, graded implicit feedback under the MNAR mechanism has not yet been investigated despite its prevalence in real-life recommender systems. In response, we formulate a recommendation with graded implicit feedback as a statistical estimation problem and define an ideal loss function of interest, which should ideally be optimized to maximize the user experience. Subsequently, we propose an unbiased estimator for the ideal loss, building on the *inverse propensity score* estimator. Finally, we conduct an empirical evaluation of the proposed method on a public real-world dataset.

ACM Reference Format:
Yuta Saito, Suguru Yaginuma, Taketo Naito, and Kazuhide Nakata. 2020. Unbiased Recommender Learning from Biased Graded Implicit Feedback. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3336191.3371783

1 INTRODUCTION

Recommender systems constitute a central part of numerous existing online platforms, such as e-commerce (e.g., Amazon and Etsy) and content recommendation platforms (e.g., Netflix, YouTube, and Spotify). The essential step in building such systems is to accurately estimate the relevance of user–item pairs. A version of user–item relevance feedback, called *implicit feedback*, is often used for solving this task. Implicit feedback is collected from natural behaviors of users in recommender systems, and it is generally available in large volumes, as expert annotation or crowd sourcing is unnecessary. Most existing studies focus on learning or evaluating recommenders using the *binary implicit feedback*, for example, an indicator of whether an item is clicked by a user [5, 10, 11, 14, 15, 18, 23]. However, in numerous application domains, implicit feedback involves not just binary relevance feedback, but it is observed in a *graded* form, such as the number of clicks [7]. The so-called *graded implicit feedback* can be considered as a more informative form of relevance feedback compared to the binary counterpart [7]. This is because graded implicit feedback contains the level of the relevance, and the grade information should be useful to better capture the user–item relevance. However, a difficulty is that the grade information is observable only for the user–item pairs with a click, whereas that for the pairs without a click is unobservable. Furthermore, the graded information for several pairs of users and items is more likely to be observed as compared with others. For example, if we deploy a popularity-based recommendation policy that recommends popular items with high probabilities in the past, the graded implicit feedback for the popular items is much easier to be observed compared with that for rare items [2, 23]. Therefore, the graded implicit feedback is often *missing-not-at-random* (MNAR), as depicted in Figure 1, and one has to address this biased missing mechanism to adequately improve the user satisfaction [18, 20].

Related Work. Some works employed graded implicit feedback to train recommendation models. Lerche and Jannach [9] modified the loss function of Bayesian personalized ranking (BPR) using confidence weights, which were estimated based on the difference in the reception times of two messages in a Twitter application. Another related work is Wang et al. [21], wherein a new set of triplets that comprised a user and two items was created based on the graded implicit feedback to train BPR, improving the recommendation accuracy of the vanilla BPR [7, 9]. However, all existing studies on the graded implicit feedback implicitly assumed the *missing-completely-at-random* mechanism, where the probability of observing the graded implicit feedback is **uniform** among all the user–item pairs. This assumption is unrealistic in many real-life

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-1-4503-6822-3/20/02…$15.00
https://doi.org/10.1145/3336191.3371783

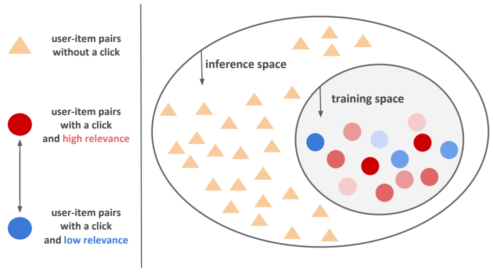

Figure 1: *Missing-not-at-random* (MNAR) problem of the graded implicit feedback. The training space comprises user–item pairs with a click ($Y = 1$), and therefore, does not represent the entire inference space, which comprises all the user–item pairs ($\mathcal{D}$).

situations, as implicit feedback is often confounded by various factors, such as the past recommendation policy [2, 13] and position bias [8, 22]. To the best of our knowledge, no theoretical framework exists in the literature that is able to handle the **MNAR** the graded implicit feedback, thereby limiting the real-world application of the methods leveraging this type of feedback.

Contributions. To better utilize graded implicit feedback, we formulate a recommendation using the MNAR graded implicit feedback from a statistical estimation perspective, which allows us to theoretically characterize the bias in using the graded implicit feedback. Based on our formulation, we first define an ideal loss function of interest that should be optimized to maximize the relevance in the platform. We then demonstrate that the naive procedure, which ignores the MNAR mechanism, is an inappropriate practice and is subject to a statistical bias. Subsequently, we propose an unbiased estimator for the ideal loss function inspired by the *inverse propensity score* (IPS) estimator in causal inference [6, 16–19, 23]. The proposed estimator is unbiased for the ideal loss function by weighting the value of the loss function for each user–item pair using the inverse of the *propensity score*. Furthermore, we demonstrate that the propensity score is estimatable in an unbiased fashion in our graded implicit feedback setting, even under the MNAR mechanism. Finally, we empirically demonstrate that the proposed unbiased learning procedure outperforms various existing baseline methods on real-world data.

2 PRELIMINARIES

2.1 Problem Setting

Let $u \in \mathcal{U}$ be a user ($|\mathcal{U}| = m$), $i \in \mathcal{I}$ be an item ($|\mathcal{I}| = n$), and $\mathcal{D} = \mathcal{U} \times \mathcal{I}$ be the set of all user-item pairs. Let $Y_{u,i}$ be a Bernoulli random variable, which represents the *binary implicit feedback* observed between u and i. If implicit feedback between u and i is observed, then $Y_{u,i}$ is equal to 1, otherwise, it takes 0. Here, $Y_{u,i} = 1$ represents the positive relationship between the user-item pair. In contrast, $Y_{u,i} = 0$ does not always imply a negative relationship. This is because implicit feedback is never observed when the user is unaware of the item, despite the fact that the item is potentially relevant to the user. To accurately model this implicit nature, we introduce two additional random variables. The first is the binary *relevance* random variable, denoted by $R_{u,i}$. This random variable represents the relevance between user u and item i, i.e., if item i is relevant to user u, $R_{u,i}$ is equal to 1, otherwise it is equal to 0. The second is the binary *awareness* random variable, denoted by $O_{u,i}$. This random variable represents whether a user u is aware of an item i.

Note that, in the binary implicit feedback setting, both the relevance and awareness random variables are **unobserved**, and only the binary implicit random variable is observable. Following the implicit feedback generation model used in previous studies [10, 15, 18], we assume the following relationships between the implicit feedback, relevance, and awareness random variables.

Assumption 2.1. *Positive implicit feedback between a user-item pair is observed if and only if the item is relevant to the user and the user is aware of the item, i.e.,* $Y_{u,i} = O_{u,i} \cdot R_{u,i}$.

This assumption implies that the item has to be relevant to the user ($R_{u,i} = 1$), **and** the user has to be aware of the item ($O_{u,i} = 1$) for the implicit feedback to be observed ($Y_{u,i} = 1$).

In the binary implicit feedback setting (BI), we can observe only the binary implicit feedback, and the training dataset for learning recommendation algorithms can be represented as:

$$\mathcal{D}_{BI} = \{(u, i, Y_{u,i})\}_{(u,i)\in\mathcal{D}}$$

In contrast, in our graded implicit feedback setting (GI), we assume that the confidence score of a user-item pair being relevant to each other ($\gamma_{u,i}$) is observable as grade information only for the pair with implicit feedback ($Y_{u,i} = 1$). Therefore, the training dataset for learning recommendation algorithms in the GI setting can be represented as:

$$\mathcal{D}_{GI} = \{(u, i, \gamma_{u,i}) \mid Y_{u,i} = 1\}_{(u,i)\in\mathcal{D}}$$

This GI setting is general in many application domains. For example, in the music recommendation problem, the dwell or consumption times that are observed after the click of a song are considered as interest signals with different strengths [4, 24]. Another example is the repeated consumption of items in an online media service, which should be interpreted as a stronger signal than a single consumption event [7]. We formulate these types of general real-world situations by treating the relevance parameter $\gamma_{u,i}$ as different levels of relevance signal. This formulation allows us to understand the bias of the MNAR graded implicit feedback from a theoretical perspective.

2.2 Ideal Loss Function of Interest

Here, we describe the objective of this study. To evaluate recommendation algorithms in the BI setting, top-N recommendation metrics such as the *mean average precision* (MAP), *discounted cumulative gain* (DCG), and *recall* are often used [12, 23]. To measure the improvement in user experience by recommendation algorithms, we rely on the following top-N recommendation metric defined with the ground-truth relevance level.

$$\mathcal{R}_{rel}\left(\widehat{\mathcal{Z}}\right) = \frac{1}{|\mathcal{U}|} \sum_{u\in\mathcal{U}} \sum_{i\in\mathcal{I}} \underbrace{\gamma_{u,i}}_{\text{relevance level}} \cdot c\left(\widehat{Z}_{u,i}\right) \qquad (1)$$

The focus of this study is to optimize the performance metric in Eq. (1). To achieve this goal, we aim at optimizing the following *ideal pointwise loss function*, which is designed to minimize the prediction loss for the ground-truth relevance parameter.

$$\mathcal{L}_{ideal}^{rel}(f) = \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \gamma_{u,i}\delta^{(1)}(f(u,i)) + \left(1-\gamma_{u,i}\right)\delta^{(0)}(f(u,i)) \quad (2)$$

where $f : \mathcal{D} \to [0,1]$ is a relevance predictor and $\delta^{(R)}$ denotes the local loss for the user-item pair (u,i). For example, if $\delta^{(R)}(f(u,i)) := (R - f(u,i))^2$, then Eq. (2) is referred to as the mean squared loss. In the following, we use $\delta^{(\cdot)}(f(u,i))$ to denote $\delta^{(\cdot)}_{u,i}$ for simplicity.

A relevance predictor f, minimizing the ideal loss defined using the relevance level in Eq. (2), is expected to lead to the desired values of the top-N recommendation metric in Eq. (1). However, we cannot handle the ideal loss function directly in our GI setting, as the ground-truth relevance parameter for the user-item pair without implicit feedback (i.e., (u,i) with $Y_{u,i}=0$) is unobservable. Therefore, accurately estimating the ideal loss function with only observable data (i.e., $\mathcal{D}_{GI}$) is important to maximize relevance from biased graded implicit feedback.

3 METHOD

In this section, we first show that the naive use of graded implicit feedback leads to a biased estimation of the ideal loss function. Then, we propose an approach to unbiasedly estimate the ideal loss function using only the biased graded implicit feedback.

3.1 Bias of Naive Estimator

To learn an arbitrary recommendation model on the graded implicit feedback dataset $\mathcal{D}_{GI}$, one can use the following naive estimator for the ideal loss function.

Definition 3.1. The naive estimator for the ideal loss function is

$$\begin{aligned}\widehat{\mathcal{L}}_{Naive}^{rel}(f) &:= \frac{1}{|\mathcal{D}_{GI}|} \sum_{(u,i)\in\mathcal{D}_{GI}} \gamma_{u,i}\delta^{(1)}_{u,i} + (1-\gamma_{u,i})\delta^{(0)}_{u,i} \\ &= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \frac{Y_{u,i}}{\pi} \left(\gamma_{u,i}\delta^{(1)}_{u,i} + (1-\gamma_{u,i})\delta^{(0)}_{u,i}\right) \quad (3)\end{aligned}$$

where $\pi = |\mathcal{D}_{GI}|/mn$ is the sparsity of the GI training dataset.

The above naive estimator is feasible, as it uses only the grade information for user-item pairs with observed implicit feedback ($Y_{u,i}=1$). However, the following proposition indicates that naive use of graded implicit feedback can lead to a sub-optimal recommender.

PROPOSITION 3.2. *The naive estimator in Eq. (3) is statistically biased against the ideal loss function of interest in Eq. (2), i.e., for some given f we have $\mathbb{E}[\widehat{\mathcal{L}}_{Naive}^{rel}(f)] \neq \mathcal{L}_{ideal}^{rel}(f)$, where the expectation is taken over the relevance and awareness random variables.*

PROOF. First, we calculate the expectation of the naive estimator below.

$$\begin{aligned}\mathbb{E}\left[\widehat{\mathcal{L}}_{Naive}^{rel}(f)\right] &= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \frac{\mathbb{E}[Y_{u,i}]}{\pi} \left(\gamma_{u,i}\delta^{(1)}_{u,i} + (1-\gamma_{u,i})\delta^{(0)}_{u,i}\right) \\ &= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \frac{\theta_{u,i}\gamma_{u,i}}{\pi} \left(\gamma_{u,i}\delta^{(1)}_{u,i} + (1-\gamma_{u,i})\delta^{(0)}_{u,i}\right)\end{aligned}$$

Therefore, for the naive estimator to satisfy the unbiasedness, the probability of observing implicit feedback must be a constant, i.e., $\theta_{u,i}\gamma_{u,i} = \pi, \ \forall (u,i) \in \mathcal{D}$. However, this necessary condition is violated in the MNAR situation, i.e, $\theta_{u,i}\gamma_{u,i}$ may not be a constant, and the naive estimator is subject to bias in the sense that its expectation is not equal to the ideal loss for some given f. □

As shown in Proposition 3.2, the expectation of the naive estimator is not necessarily equal to the ideal loss function. This is because the naive estimator ignores the distributional shift between the user-item pairs with and without implicit feedback. Therefore, the naive use of graded implicit feedback is an inappropriate approach to approximate the ideal loss function. Instead, one should rely on an estimator alleviating the bias alternative to using the naive estimator.

3.2 The Proposed Unbiased Estimator

To alleviate the bias of the graded implicit feedback, we propose the following unbiased estimator.

Definition 3.3. The unbiased estimator for the ideal loss function is defined as

$$\begin{aligned}\widehat{\mathcal{L}}_{UB}^{rel}(f) &:= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}_{GI}} \frac{1}{\theta_{u,i}} \left(\delta^{(1)}_{u,i} + \frac{1-\gamma_{u,i}}{\gamma_{u,i}}\delta^{(0)}_{u,i}\right) \\ &= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \frac{Y_{u,i}}{\theta_{u,i}} \left(\delta^{(1)}_{u,i} + \frac{1-\gamma_{u,i}}{\gamma_{u,i}}\delta^{(0)}_{u,i}\right), \quad (4)\end{aligned}$$

where the awareness parameter $\theta_{u,i}$ is interpreted as the propensity score in the GI setting.

The following proposition states that the proposed unbiased estimator is unbiased for the ideal loss function.

PROPOSITION 3.4. *The unbiased estimator in Eq. (4) is statistically unbiased against the ideal loss function of interest in Eq. (2), i.e., for any given relevance predictor f, we have $\mathbb{E}[\widehat{\mathcal{L}}_{UB}^{rel}(f)] = \mathcal{L}_{ideal}^{rel}(f)$, where the expectation is taken over the relevance and awareness random variables.*

PROOF.

$$\begin{aligned}\mathbb{E}\left[\widehat{\mathcal{L}}_{UB}^{rel}(f)\right] &= \mathbb{E}\left[\frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \frac{Y_{u,i}}{\theta_{u,i}} \left(\delta^{(1)}_{u,i} + \frac{1-\gamma_{u,i}}{\gamma_{u,i}}\delta^{(0)}_{u,i}\right)\right] \\ &= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \frac{\mathbb{E}[Y_{u,i}]}{\theta_{u,i}} \left(\delta^{(1)}_{u,i} + \frac{1-\gamma_{u,i}}{\gamma_{u,i}}\delta^{(0)}_{u,i}\right) \\ &= \frac{1}{mn} \sum_{(u,i)\in\mathcal{D}} \gamma_{u,i}\delta^{(1)}_{u,i} + (1-\gamma_{u,i})\delta^{(0)}_{u,i} \\ &= \mathcal{L}_{ideal}^{rel}(f)\end{aligned}$$

□

Proposition 3.4 suggests that the proposed unbiased estimator is *statistically unbiased* for the ideal loss function; one can unbiasedly approximate the ideal loss function of interest using only biased graded implicit feedback with this estimator.

3.3 Unbiased Propensity Estimation

The proposed unbiased estimator requires the true value of the awareness parameter $\theta_{u,i}$. However, in our setting, the ground-truth awareness parameters are unobserved and need to be estimated from the data. For estimating this parameter, we should optimize the following loss function.

$$\mathcal{L}^{score}_{ideal}(g) := \frac{1}{mn} \sum_{(u,i) \in \mathcal{D}} \theta_{u,i} \delta^{(1)}(g(u,i)) + (1 - \theta_{u,i}) \delta^{(0)}(g(u,i)) \quad (5)$$

where $g : \mathcal{D} \rightarrow [0,1]$ is an awareness parameter estimator.

However, the realizations of the awareness random variables are unobserved in $\mathcal{D}_{GI}$, and the direct optimization of Eq. (5) is infeasible. Therefore, we propose the following unbiased estimator to estimate the propensity score with only available data.

Definition 3.5. The unbiased estimator for propensity estimation is defined as follows

$$\widehat{\mathcal{L}}^{score}_{UB}(g) = \frac{1}{mn} \sum_{(u,i) \in \mathcal{D}} \frac{Y_{u,i}}{\gamma_{u,i}} \delta^{(1)}_{u,i} + \left(1 - \frac{Y_{u,i}}{\gamma_{u,i}}\right) \delta^{(0)}_{u,i} \quad (6)$$

Note that Eq. (6) depends on only observable data. This is because the ground-truth relevance parameters can be used for the user-item pair with implicit feedback ($Y_{u,i} = 1$), whereas those for the pair without implicit feedback ($Y_{u,i} = 0$) are unnecessary to calculate Eq. (6).

The following proposition provides a theoretical justification of the unbiased estimator in Eq. (6).

Proposition 3.6. *The unbiased estimator for the propensity estimation in Eq. (6) is statistically unbiased for the loss function in Eq. (5), i.e., for any given g, we have $\mathbb{E}[\widehat{\mathcal{L}}^{score}_{UB}(g)] = \mathcal{L}^{score}_{ideal}(g)$.*

Proof.

$$\begin{aligned}
\mathbb{E}\left[\frac{Y_{u,i}}{\gamma_{u,i}} \delta^{(1)}_{u,i} + \left(1 - \frac{Y_{u,i}}{\gamma_{u,i}}\right) \delta^{(0)}_{u,i}\right] &= \frac{\mathbb{E}[Y_{u,i}]}{\gamma_{u,i}} \delta^{(1)}_{u,i} + \left(1 - \frac{\mathbb{E}[Y_{u,i}]}{\gamma_{u,i}}\right) \delta^{(0)}_{u,i} \\
&= \frac{\gamma_{u,i}\theta_{u,i}}{\gamma_{u,i}} \delta^{(1)}_{u,i} + \left(1 - \frac{\gamma_{u,i}\theta_{u,i}}{\gamma_{u,i}}\right) \delta^{(0)}_{u,i} \\
&= \theta_{u,i} \delta^{(1)}_{u,i} + \left(1 - \theta_{u,i}\right) \delta^{(0)}_{u,i}
\end{aligned}$$

Therefore, we obtain the unbiasedness of the estimator from the linearity of the expectation operator. □

Thus, by using the unbiased estimator for the propensity estimation, the unobserved awareness parameters can be estimated in an unbiased fashion from only observable data.

4 EXPERIMENTAL EVALUATION

In this section, we empirically evaluate the proposed unbiased learning framework on a public real-world dataset.

4.1 Experimental Setting

4.1.1 Dataset and Preprocessing. We use the Yahoo! R3[1] dataset and employ the following preprocessing procedure.

(1) Transform five-star ratings into grade information following a methodology used in the information retrieval domain [1, 3]:

$$\gamma_{u,i} = \epsilon + (1 - \epsilon) \frac{2^r - 1}{2^{r_{max}} - 1}$$

where $r \in \{1, 2, 3, 4, 5\}$ denotes a five-star rating, and r_{max} is the maximum possible rating, which is 5 in our case. Additionally, $\epsilon \in [0, 1]$ controls the noise level in the grade information. We apply $\epsilon = 0.1$ for the training datasets and $\epsilon = 0$ for the test datasets to evaluate the recommenders with the ground-truth (noise-free) relevance in the test sets.

(2) Sample binary relevance variable $R_{u,i}$ by performing the following Bernoulli sampling:

$$R_{u,i} \sim Bern(\gamma_{u,i}), \ \forall (u,i) \in \mathcal{D}$$

where $Bern(\cdot)$ denotes Bernoulli distribution.

(3) Define the awareness variable for a user–item pair as follows:

$$O_{u,i} = \begin{cases} 1 & \text{(if item } i \text{ is rated by user } u) \\ 0 & \text{(if item } i \text{ is not rated by user } u) \end{cases}$$

(4) Define dataset as: $\mathcal{D}_{GI} = \{(u, i, \gamma_{u,i}) \mid Y_{u,i} = 1\}$, where $Y_{u,i} = O_{u,i} \cdot R_{u,i}$. Note that $R_{u,i}$ and $O_{u,i}$ are unobservable in our setting and are not used for training recommenders.

4.1.2 Compared Methods. We compare the following methods: (i) **Item popularity model (ItemPop)**: This method always recommends the k most clicked items in the training set; therefore, it is not personalized. (ii) **Naive matrix factorization (NaiveMF)**: This model predicts the graded implicit feedback (i.e., $\gamma_{u,i}$) by optimizing the naive loss in Eq. (3) via MF [5]. It ignores the distributional shift, and thus the loss function is biased. (iii) **Exposure matrix factorization (ExpoMF)**: This model is a state-of-the-art method for the MNAR binary implicit feedback [10]. This method is based on a probabilistic model that uses the exposure variable and optimizes its parameters using an EM-algorithm-like procedure. (iv) **Relevance matrix factorization (Rel-MF) [18]:** This method is another state-of-the-art method for the MNAR binary implicit feedback. It obtains the model parameters by optimizing the unbiased loss function for the ideal loss function defined with only the binary implicit feedback signals. (v) **Matrix factorization with unbiased estimator (MF-UB)**: This is our proposed method, which obtains model parameters by optimizing the proposed unbiased loss function in Eq. (4). We estimate the propensity score based on the unbiased propensity estimation procedure described in the previous section.

4.2 Results

We used DCG@K, Recall@K, and MAP@K to measure the recommendation quality in the test sets. The values of K were set to $\{3, 5\}$ for all metrics. Table 1 summarizes the ranking performances of all methods on Yahoo! R3. It shows that the proposed MF-UB method outperforms other baseline methods in all settings by utilizing the graded implicit feedback in a theoretically grounded fashion. Specifically, MF-UB outperformed the best baselines by 3.5% for DCG@5, 3.0% for Recall@5, and 3.9% for MAP@5. The results suggest that

[1]http://webscope.sandbox.yahoo.com/

Table 1: Evaluating the ranking performance of the proposed method against the baselines

Methods	DCG@K		Recall@K		MAP@K	
	$K = 3$	$K = 5$	$K = 3$	$K = 5$	$K = 3$	$K = 5$
ItemPop	0.26400	0.35476	0.30034	0.49539	0.30379	0.50232
NaiveMF	0.33509	0.42596	0.37393	0.56866	0.40950	0.64196
ExpoMF	0.32805	0.41591	0.36626	0.55477	0.40376	0.63031
Rel-MF	0.33523	0.42498	0.37479	0.56744	0.40909	0.64072
MF-UB (ours)	**0.34856**	**0.44004**	**0.38845**	**0.58493**	**0.42872**	**0.66757**

one can significantly improve the recommendation quality by appropriately using the biased graded implicit feedback in the right manner.

5 CONCLUSION

We explored a theoretically sophisticated method for improving user experience with biased graded implicit feedback. In particular, we formulated the problem as a statistical estimation problem to address the inherent bias of the graded implicit feedback. We then demonstrated that the naive use of the graded implicit feedback could result in a sub-optimal recommendation, as it ignores the distributional shift between clicked and unclicked events. To address the bias of the naive estimator, we proposed an unbiased estimator for the ideal loss function, which can be estimated using observable graded implicit feedback. We also experimentally demonstrated that the proposed unbiased estimation approach outperformed the widely used baselines by effectively utilizing additional grade information to infer the relevance.

REFERENCES

[1] Qingyao Ai, Keping Bi, Cheng Luo, Jiafeng Guo, and W Bruce Croft. 2018. Unbiased learning to rank with unbiased propensity estimation. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval.* 385–394.

[2] Stephen Bonner and Flavian Vasile. 2018. Causal Embeddings for Recommendation. In *Proceedings of the 12th ACM Conference on Recommender Systems* (Vancouver, British Columbia, Canada) *(RecSys '18)*. 104–112.

[3] Olivier Chapelle, Donald Metlzer, Ya Zhang, and Pierre Grinspan. 2009. Expected reciprocal rank for graded relevance. In *Proceedings of the 18th ACM conference on Information and knowledge management.* 621–630.

[4] Paolo Dragone, Rishabh Mehrotra, and Mounia Lalmas. 2019. Deriving User- and Content-specific Rewards for Contextual Bandits. In *The World Wide Web Conference.* 2680–2686.

[5] Yifan Hu, Yehuda Koren, and Chris Volinsky. 2008. Collaborative filtering for implicit feedback datasets. In *2008 Eighth IEEE International Conference on Data Mining.* Ieee, 263–272.

[6] Guido W Imbens and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences.* Cambridge University Press.

[7] Dietmar Jannach, Lukas Lerche, and Markus Zanker. 2018. Recommending based on implicit feedback. In *Social Information Access.* Springer, 510–569.

[8] Thorsten Joachims, Adith Swaminathan, and Tobias Schnabel. 2017. Unbiased learning-to-rank with biased feedback. In *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining.* ACM, 781–789.

[9] Lukas Lerche and Dietmar Jannach. 2014. Using graded implicit feedback for bayesian personalized ranking. In *Proceedings of the 8th ACM Conference on Recommender systems.* ACM, 353–356.

[10] Dawen Liang, Laurent Charlin, James McInerney, and David M Blei. 2016. Modeling user exposure in recommendation. In *Proceedings of the 25th International Conference on World Wide Web.* International World Wide Web Conferences Steering Committee, 951–961.

[11] Steffen Rendle. 2010. Factorization machines. In *2010 IEEE International Conference on Data Mining.* IEEE, 995–1000.

[12] Steffen Rendle. 2019. Evaluation Metrics for Item Recommendation under Sampling. *arXiv preprint arXiv:1912.02263* (2019).

[13] Yuta Saito. 2020. Asymmetric Tri-training for Debiasing Missing-Not-At-Random Explicit Feedback. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval.* 309–318.

[14] Yuta Saito. 2020. Doubly Robust Estimator for Ranking Metrics with Post-Click Conversions. In *Fourteenth ACM Conference on Recommender Systems.* 92–100.

[15] Yuta Saito. 2020. Unbiased Pairwise Learning from Biased Implicit Feedback. In *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval.* 5–12.

[16] Yuta Saito, Gota Morisihta, and Shota Yasui. 2020. Dual Learning Algorithm for Delayed Conversions. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval.* 1849–1852.

[17] Yuta Saito, Takuma Udagawa, Haruka Kiyohara, Kazuki Mogi, Yusuke Narita, and Kei Tateno. 2021. Evaluating the Robustness of Off-Policy Evaluation. *arXiv preprint arXiv:2108.13703* (2021).

[18] Yuta Saito, Suguru Yaginuma, Yuta Nishino, Hayato Sakata, and Kazuhide Nakata. 2020. Unbiased Recommender Learning from Missing-Not-At-Random Implicit Feedback. In *Proceedings of the 13th International Conference on Web Search and Data Mining.* 501–509.

[19] Tobias Schnabel, Adith Swaminathan, Ashudeep Singh, Navin Chandak, and Thorsten Joachims. 2016. Recommendations as Treatments: Debiasing Learning and Evaluation. In *Proceedings of The 33rd International Conference on Machine Learning*, Vol. 48. PMLR, 1670–1679.

[20] Harald Steck. 2010. 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.* ACM, 713–722.

[21] Sheng Wang, Xiaobo Zhou, Ziqi Wang, and Ming Zhang. 2012. Please spread: recommending tweets for retweeting with implicit feedback. In *Proceedings of the 2012 workshop on Data-driven user behavioral modelling and mining from social media.* 19–22.

[22] Xuanhui Wang, Nadav Golbandi, Michael Bendersky, Donald Metzler, and Marc Najork. 2018. Position bias estimation for unbiased learning to rank in personal search. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining.* ACM, 610–618.

[23] Longqi Yang, Yin Cui, Yuan Xuan, Chenyang Wang, Serge Belongie, and Deborah Estrin. 2018. Unbiased Offline Recommender Evaluation for Missing-not-at-random Implicit Feedback. In *Proceedings of the 12th ACM Conference on Recommender Systems (RecSys '18).* 279–287.

[24] Xing Yi, Liangjie Hong, Erheng Zhong, Nanthan Nan Liu, and Suju Rajan. 2014. Beyond clicks: dwell time for personalization. In *Proceedings of the 8th ACM Conference on Recommender systems.* 113–120.