from the same group of users. Clearly these two distributions are very different. In the user selected ratings, there are far more items with high ratings than that in the randomly selected songs. This is compelling evidence showing that the assumption that all the users would rate all the inspected items or select random items to rate is unlikely to be true. The investigation of the Yahoo!LaunchCast data indicates that users are more likely to rate items they do love and hate, but not neutral [16, 26].

To further demonstrate the risk of incorrect parameter estimation and biased rating prediction when ignoring the response information, we show an intuitive example of five users’ rating on five items in Table 1, where the ratings are skewed to either 4 or 5. Clearly, user-based approaches [1, 4] and item-based approaches [3, 11, 25] are more likely to predict rating values in the range of 4 to 5. Similarly, ignoring response information will cause the one-class issue for model-based approaches [10, 17, 18]. In a real-world recommender system, the case may not be as extreme as is in Table 1. Nevertheless, the effect is similar. By ignoring the response information, we will learn a model that has bias.

Currently, there are two main streams of work trying to solve the above response ignorance problem. One line of work try to model the above phenomena as a one-class collaborative filtering task [10, 17, 18]. A heuristic weight in the range of 0 to 1 is introduced to calibrate the loss on those unseen ratings, where the rating scores are set to zeros [17, 18]. Embedding user information is also adopted to optimize the weight on the unseen ratings via users’ similarity [10]. However, these methods do not model the users’ missing response information together with the ratings. The other line of work model the response ignorance through missing data theory [13]. The multinomial mixture model is adopted to model the non-random response [16]. The work is also extended for collaborative ranking [15]. These methods model users’ response patterns and ratings via multinomial mixture model, but they discard the effectiveness and interpretability of the matrix factorization approaches [8, 22].

To bridge this gap, we are the first to integrate the users’ response patterns into PMF to establish a unified framework, which we refer to as Response Aware PMF (RAPMF). The response models we propose include the rating dominating response model, and a generalized one, the context-aware response model. We demonstrate the advantages of our proposed RAPMF through detailed and fair experimental comparison.

The rest of the paper is organized as follows. In Section 2, we motivate the explicit modeling of user responses from a probabilistic point of view. In Section 3, we present how to incorporate response models into PMF and elaborate the proposed RAPMF model. Empirical study and comparison with previous work is conducted in Section 4. The paper is concluded in Section 5.

2 Response and Missing Theory

Modeling response patterns have a strong incentive from statistical missing data theory [13]. The response patterns can be hidden [2, 6] or explicit. In recommender system case, it is explicit. In the following, we show that without modeling the response patterns properly, we may learn a bias model.

2.1 Setup and Notation

Assume that we are given a partially observed \( N \times M \) matrix \( X \), where \( N \) is the number of users and \( M \) is the number of items, the \((i, j)\) element of \( X \) denotes the rating assigned by user \( i \) to item \( j \) in the scale of 1 to \( D \). Collaborative filtering approaches try to recover the original full matrix \( X_{full} \) to predict users’ preferences.

In the matrix \( X \), an unobserved entry is denoted by 0. Alternatively, we denote all the observations as a set of triplets \((i, j, x) \in Q\). Moreover, we define a companion response indicator matrix \( R \) to denote whether the corresponding rating is observed in \( X \). If \( X_{ij} \neq 0 \),