A Rate-Distortion Framework for Compressing User Interaction Sequences with Adaptive Preference Disentanglement*

Wooseung Kang[1] and Sang-Min Choi[1†]

Gyeongsang National University, Jinju-si, Republic of Korea
kung0615@gnu.ac.kr, jerassi@gnu.ac.kr

Abstract

As data continues to grow, research on search and recommendation technologies has become increasingly active. To enable personalized recommendations, platforms often utilize diverse user–system interaction data. However, the direct use of this information raises significant privacy concerns, as it can lead to the leakage of sensitive user data. In this paper, we propose a method to mitigate such privacy risks by compressing user interaction sequences into compact representations suitable for model learning. Leveraging the rate–distortion theory, the proposed approach adaptively adjusts the compression rate based on the variability within user interaction sequences, achieving an effective trade-off between personalization performance and privacy preservation.

1 Introduction

Modern digital platforms, such as e-commerce sites and content streaming services, must provide users with relevant information through efficient search and recommendation systems to navigate the vast amount of available data. Recommender systems aim to predict a user's next item of interest by learning from their diverse interaction histories, such as past clicks and purchases [?]. However, these interaction data contain sensitive personal information, which requires robust privacy-preserving measures. To this end, a common practice is to store user data not in its raw form but as learned embeddings [?]. Despite this, recent studies have shown that these embeddings are vulnerable to information leakage, making it possible to infer the original user interactions [?]. Therefore, this paper proposes a novel framework for learning compressed representations of user interaction sequences based on Rate-Distortion theory. The core of our methodology is to disentangle the user's preferences into long- and short-term components and apply an adaptive compression rate to each. This approach enhances privacy by introducing significant distortion that makes the original sequence difficult to reconstruct, while simultaneously ensuring memory efficiency for storage.

2 Our Approach

We propose a Rate-Distortion framework to compress user interaction sequences. As in figure **??**, we apply a Fast Fourier Transform(FFT) to disentangle the sequence into low-frequency (long-term) and high-frequency (short-term) preference components [?]. Each component is

*Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-6, December 16-18, 2025, Sapporo, Japan.

†Corresponding author

Mobisec2025_.png

Figure 1: Example of Our method

then fed into a separate MLP-based Rate-Distortion Encoder, which compresses it into a probabilistic latent variable, Z_{low} and Z_{high}. These variables are concatenated to form the input for a downstream recommendation model. The framework is trained by minimizing the total loss:

$$\mathcal{L}_{\text{Total}} = \mathcal{D} + \beta_{\text{low}} \cdot \mathcal{R}_{\text{low}} + \beta_{\text{high}} \cdot \mathcal{R}_{\text{high}} \tag{1}$$

where $\mathcal{D}$ is the primary recommendation loss. The rate losses, $\mathcal{R}_{\text{low}}$ and $\mathcal{R}_{\text{high}}$, are are calculated using the Kullback-Leibler divergence. This loss achieves compression by forcing the distribution of each latent variable to match a simple prior(P)—the standard normal distribution $\mathcal{N}(0, 1)$ The hyperparameters β_{low} and β_{high} control the compression-accuracy trade-off for long-term and short-term preferences, respectively.

3 Conclusion

The proposed framework is expected to (1) reduce privacy risks by limiting reconstructable user details in latent space, (2) maintain recommendation quality by preserving dominant preference frequencies, and (3) enable lightweight model training through adaptive compression.

Acknowledgments

This research was supported by the Regional Innovation System & Education(RISE) program through the RISE Center, Gyeongsangnam-do, funded by the Ministry of Education(MOE) and the Gyeongsangnam-do Provincial Government, Republic of Korea.(2025-RISE-16-001) Additionally, this work was supported by The Glocal University 30 Project Fund of Gyeongsang National University in 2025.