修　士　論　文

A Research on Cross-Domain Recommendation under Non-Overlapping Settings by Integrating Gaussian Mixture User Preference Distributions and Optimal Transport

非重複設定下におけるガウス混合モデルユーザ嗜好分布と最適輸送を統合したクロスドメイン推薦に関する研究

指導教員　　**鈴村 豊太郎 教授**

東京大学大学院 情報理工学系研究科
電子情報学専攻

氏 名　**48-246453 肖　子吟**

提 出 日　　2026 年 1 月 22 日

Abstract

Recommendation systems are essential tools for helping users discover items of interest in various domains, such as e-commerce, streaming services, and social media. With the rapid growth of online platforms, the need for effective recommendation systems has become increasingly important.

However, traditional recommendation systems often face challenges such as data sparsity and cold-start problems, which can limit their effectiveness. Cross-domain recommendation has emerged as a promising approach to address these challenges by leveraging information from multiple domains to improve recommendation accuracy. Nevertheless, existing cross-domain recommendation methods often assume overlapping users or items between domains, which may not hold in many real-world scenarios. Also, they usually represent user preferences as point estimates, which may not capture the full complexity of user behavior.

To address these limitations, in this thesis, we propose a novel framework that models user preferences as Gaussian Mixture Models (GMMs) and integrates optimal transport techniques for cross-domain recommendation under non-overlapping settings. By representing user preferences as distributions, our approach can capture the uncertainty and diversity of user behavior more effectively and fine-grainedly. The use of optimal transport allows us to align user preference distributions across domains, facilitating knowledge transfer even in the absence of overlapping users or items.

We evaluate our proposed method on Amazon datasets across multiple domains, comparing its performance against state-of-the-art cross-domain recommendation techniques and traditional single-domain methods. We also conduct ablation studies to assess the effectiveness of representing user preferences as distributions and the impact of boosting target domain recommendation performance using source domain information. The experimental results demonstrate that our approach significantly outperforms existing methods in terms of recommendation accuracy, highlighting the effectiveness of modeling user preferences as distributions and leveraging optimal transport for cross-domain knowledge transfer.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1   Background

These days, information overload has become a significant challenge for users in various online platforms, such as e-commerce websites, streaming services, and social media[1, 2, 3]. Recommender systems play a crucial role in addressing this challenge by providing personalized guidance to users by selecting items with the highest predicted utility from a large space of options.[4, 5, 6]. By analyzing user-item interactions, recommender systems can help users discover relevant content, products, or services, thereby enhancing user satisfaction and engagement. Besides improving user experience, recommender systems also benefit businesses by increasing sales, customer retention, and overall platform activity[7]. As the volume of online content continues to grow, the importance of effective recommender systems becomes even more pronounced. Despite their significance, recommender systems face several challenges that can hinder their performance. One of the primary challenges is data sparsity, where users interact with only a small fraction of the available items, making it difficult to accurately model user preferences. This issue is particularly prevalent in new or niche platforms with limited user interactions. Another challenge is the cold-start problem, which occurs when new users or items are introduced to the system. In such cases, there is insufficient data to make reliable recommendations, leading to suboptimal user experiences. Both data sparsity and cold-start problems are caused by the limited availability of user-item interaction data, which can significantly impact the effectiveness of recommender systems. To address these challendges, cross-domain recommendation has emerged as a promising approach that leverages information from multiple domains to improve recommendation accuracy. Cross-domain recommendation aims to enhance the performance of recommender systems by transferring knowledge from a source domain to a target domain. This approach is particularly useful in scenarios where the target domain suffers from data sparsity or cold-start issues. By utilizing user-item interactions from a related source domain, cross-domain recommendation can help alleviate these challenges and provide more accurate recommendations. However, existing cross-domain recommendation methods often assume overlapping users or items between the source and target domains, and rely on these overlaps to build connections between the two domains in the training stage. However, in many real-world scenarios, such overlaps may not exist, limiting the applicability of these

methods. Additionally, traditional recommender systems typically represent user preferences as point estimates, which may not capture the full complexity and uncertainty of user behavior. To overcome these limitations, there is a need for novel approaches that can effectively model user preferences and facilitate knowledge transfer between non-overlapping domains during the training process. In this thesis, we propose a novel framework that models user preferences as Gaussian Mixture Models (GMMs) and integrates optimal transport techniques for cross-domain recommendation under non-overlapping settings. By representing user preferences as distributions, our approach can capture the uncertainty and diversity of user behavior more effectively. The use of optimal transport allows us to align user preference distributions across domains, facilitating knowledge transfer even in the absence of overlapping users or items. The proposed framework consists of two main components: (1) modeling user preferences as GMMs to capture the complexity of user behavior, and (2) employing optimal transport to align these distributions between the source and target domains. This combination enables our method to leverage information from the source domain to enhance recommendation accuracy in the target domain, even when there are no overlapping users or items.

1.2   Contributions

The main contributions of this thesis are as follows:

- We introduce a novel framework for cross-domain recommendation under non-overlapping settings by modeling

- We propose representing user preferences as Gaussian Mixture Models (GMMs) to capture the uncertainty and diversity of user behavior more effectively.

- We integrate optimal transport techniques to align user preference distributions across domains, enabling knowledge transfer

- We conduct extensive experiments on Amazon datasets across multiple domains to evaluate the performance of our proposed method. The results demonstrate that our approach significantly outperforms existing cross-domain recommendation techniques and traditional single-domain methods in terms of recommendation accuracy.

- We perform ablation studies to assess the effectiveness of representing user preferences as distributions and the impact of boosting target domain recommendation performance using source domain information.

- We provide insights into the benefits of modeling user preferences as distributions and leveraging optimal transport for cross-domain knowledge transfer, highlighting the potential of our approach for addressing challenges in recommender systems.

1.3   Thesis Organization

The remainder of this thesis is organized as follows:

- In Chapter 2, we provide an overview of the fundamental concepts and techniques relevant to this research, including recommender systems, cross-domain recommendation, Gaussian Mixture Models, and optimal transport.

- In Chapter 3, we review related work on recommender systems, cross-domain recommendation, and optimal transport techniques.

- In Chapter 4, we present our proposed framework for cross-domain recommendation under non-overlapping settings, detailing the modeling of user preferences as GMMs and the integration of optimal transport for knowledge transfer.

- In Chapter 5, we describe the experimental setup, datasets, and evaluation metrics used to assess the performance of our proposed method.

- In Chapter 6, we present and discuss the experimental results, including comparisons with existing methods and ablation studies.

- Finally, in Chapter 7, we summarize the main findings of this thesis and discuss potential directions for future research.

Chapter 2

Preliminaries

2.1   Recommender Systems

Recommender systems can be mathmatically modeled as a function $R : U \times I \rightarrow S$, where U is the set of users, I is the set of items, and S is the set of possible scores or ratings. The goal of a recommender system is to predict the score $s \in S$ that a user $u \in U$ would give to an item $i \in I$. This prediction can be represented as $\hat{s} = R(u, i)$. The final task of recommender systems is to generate a ranked list of items for each user based on the predicted scores.

2.2   Cross-Domain Recommendation

Cross-domain recommendation aims to leverage user preferences and behaviors from one or more source domains to improve recommendation performance in a target domain. Formally, let D_s be the source domain with user set U_s and item set I_s, and D_t be the target domain with user set U_t and item set I_t. The objective is to learn a recommendation function $R_t : U_t \times I_t \rightarrow S$ for the target domain by utilizing information from the source domain(s) D_s. The infromation that can be transferred from the source domain to the target domain includes user-item interactions, user profiles, item attributes, and latent factors learned from the source domain. The challenge in cross-domain recommendation lies in effectively transferring knowledge while addressing issues such as domain heterogeneity, data sparsity, and cold-start problems.

2.3   Gaussian Mixture Models

A Gaussian Mixture Model (GMM) is a probabilistic model that assumes data is generated from a mixture of several Gaussian distributions. Formally, a GMM can be represented as:

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x | \mu_k, \Sigma_k) \tag{2.1}$$

where K is the number of Gaussian components, π_k are the mixture weights satisfying $\sum_{k=1}^{K} \pi_k = 1$, and $\mathcal{N}(x|\mu_k, \Sigma_k)$ is the Gaussian distribution with mean μ_k and covariance matrix Σ_k.

2.4   Optimal Transport

Optimal transport is a mathematical framework for comparing and transforming probability distributions. Given two probability distributions μ and ν defined on spaces X and Y, respectively, the optimal transport problem seeks to find a mapping $T : X \to Y$ that minimizes the cost of transporting mass from μ to ν. The cost function $c(x,y)$ quantifies the expense of moving mass from point $x \in X$ to point $y \in Y$. The optimal transport problem can be formulated as:

$$\min_{T} \int_X c(x, T(x))d\mu(x) \tag{2.2}$$

subject to the constraint that the pushforward measure $T_{\#}\mu = \nu$. Optimal transport has been widely used in various applications, including image processing, machine learning, and economics, due to its ability to capture the geometric structure of probability distributions. In our proposed method, we are not using optimal transport in the standard sense above, but using a related concept called Wasserstein distance to measure the distance between two probability distributions. The Wasserstein distance is defined as:

$$W_p(\mu, \nu) = \left(\inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y)^p d\gamma(x, y) \right)^{1/p} \tag{2.3}$$

where $\Pi(\mu, \nu)$ is the set of all joint distributions (couplings) with marginals μ and ν, and $c(x,y)$ is the cost function. The Wasserstein distance provides a meaningful way to compare probability distributions, taking into account the underlying geometry of the data. With the distance between two probability distributions defined, we can replace the conventional pointwise distance (e.g., Euclidean distance) between vectors with the Wasserstein distance between the corresponding distributions in our proposed method.

Chapter 3

Related Work

3.1   Recommender Systems

This section reviews the existing literature on recommender systems, categorizing the related work into several key areas based on the techniques and approaches employed. Each subsection delves into specific methodologies and advancements within the field.

3.1.1   Traditional Recommender Systems

In the early stages of recommender systems, traditional techniques such as collaborative filtering, content-based filtering, and hybrid methods were predominantly used. Collaborative filtering (CF) is one of the foundational approaches in recommender systems, operating on the principle that users with similar preferences will like similar items, or that items liked by similar users will be preferred by a given user[8, 9].

CF can be divded into memory based and model-based methods based on how recommendations are generated. Memory-based methods utilize user-item interaction data directly to compute similarities between users or items, while model-based methods employ machine learning algorithms to learn latent factors from the interaction data.

3.1.1.1   Memory-based CF

By user or item similarities, memory-based CF can be further categorized into user-based and item-based approaches. User-based CF recommends items to a user based on the preferences of similar users, while item-based CF suggests items similar to those the user has previously liked.

GroupLens[10] is one of the earliest and most influential memory-based CF systems, which introduced user-based collaborative filtering using Pearson correlation to compute user similarities. It demonstrated the effectiveness of CF in providing personalized recommendations and laid the groundwork for subsequent research in the field.

Sarwar et al.[11] proposed an item-based collaborative filtering approach that focused on item similarities rather than user similarities. This method improved scalability and recommendation accuracy

by leveraging item-item correlations, making it more suitable for large-scale recommender systems. This approach is widely adopted in commercial recommender systems due to its efficiency and effectiveness even today.

3.1.1.2   Model-based CF

Model-based CF methods utilize machine learning techniques to learn latent representations of users and items from interaction data.

Breese et al.[12] conducted a comprehensive analysis of various model-based CF techniques, including Bayesian networks and clustering methods. Their work highlighted the advantages of model-based approaches in handling data sparsity and improving recommendation accuracy. Bayesian networks, in particular, provided a probabilistic framework for modeling user preferences and item characteristics. This method laid the foundation for later model-based CF techniques as it proposed a general probabilistic framework for recommendation.

Ungar and Foster[13] introduced a clustering-based CF method that grouped users with similar preferences and made recommendations based on cluster-level preferences. This approach effectively addressed the scalability issue in CF by reducing the dimensionality of the user-item interaction space.

Koren et al.[14] proposed matrix factorization (MF) techniques for recommender systems, which became one of the most popular model-based CF methods. MF decomposes the user-item interaction matrix into low-rank latent factor matrices, capturing the underlying structure of user preferences and item characteristics. This method demonstrated significant improvements in recommendation accuracy and scalability, particularly in large-scale datasets.

Salakhutdinov and Mnih[15] introduced Bayesian Probabilistic Matrix Factorization (BPMF), which incorporated Bayesian inference into the MF framework. BPMF provided a principled way to handle uncertainty in user preferences and item characteristics, leading to improved recommendation performance.

These model-based methods mainly represent user preferences as discrete vectors in a latent space, but there are also some methods that represent user preferences as probability distributions. Considering that our proposed method also represents user preferences as distributions, we introduce some of these related works below.

Hofmann [16] proposed Probabilistic Latent Semantic Analysis (PLSA) for recommender systems, which models user-item interactions using a mixture of latent topics. PLSA represents user preferences as probability distributions over latent topics, allowing for a more nuanced representation of user behavior.

Marlin [17] introduced a probabilistic approach to model user rating profiles, capturing the uncertainty in user preferences. This method utilized Gaussian distributions to represent user ratings, enabling more accurate recommendations by accounting for variability in user behavior.

Blei, Ng, and Jordan [18] proposed Latent Dirichlet Allocation (LDA), a generative probabilistic model that represents documents as mixtures of topics. Although originally designed for text modeling, LDA has been adapted for recommender systems to model user preferences as distributions over latent topics, providing a richer representation of user behavior.

3.1.1.3   Content-based Filtering

Content-based filtering (CBF) recommends items to users based on the attributes of items and the user's past interactions with similar items.

Salton et al.[19] introduced the vector space model for information retrieval, which laid the foundation for content-based recommender systems. This model represents items and user profiles as vectors in a high-dimensional space, allowing for the computation of similarity scores between users and items.

Pazzani et al.[20] proposed a content-based recommender system that constructs user profiles based on item attributes and user feedback. Their approach utilized machine learning techniques to learn user preferences and recommend items that align with those preferences.

3.1.1.4   Hybrid Methods

Hybrid recommender systems combine both collaborative filtering and content-based filtering techniques to leverage the strengths of each approach and mitigate their respective weaknesses.

Burke[5] provided a comprehensive survey of hybrid recommender systems, categorizing various hybridization strategies such as weighted, switching, and feature combination methods. This work highlighted the benefits of hybrid approaches in improving recommendation accuracy and addressing challenges like data sparsity and cold-start problems.

Pazzani[21] proposed a hybrid recommender system that integrates collaborative filtering and content-based filtering to enhance recommendation quality. This system utilized user-item interaction data and item attributes to generate personalized recommendations, demonstrating improved performance over single-method approaches.

Melville et al.[22] introduced a content-boosted collaborative filtering approach that incorporates content information into the collaborative filtering framework. This method effectively addresses data sparsity issues by leveraging item attributes to enhance user-item interaction data. The core idea is to use content-based predictions to fill in missing values in the user-item interaction matrix, thereby improving the overall recommendation accuracy.

Billsus et al.[23] developed a hybrid news recommender system that combines collaborative filtering and content-based filtering techniques. Their approach utilized user feedback and item content to generate personalized news recommendations, demonstrating the effectiveness of hybrid methods in real-world applications. This system is a typical example of application-driven hybrid recommender systems.

3.1.2   Supervised Machine Learning in Recommender Systems

Supervised machine learning techniques have also been applied to recommender systems to enhance recommendation accuracy. Algorithms such as decision trees, support vector machines (SVM), and ensemble methods have been utilized to model user preferences based on various features. These techniques often incorporate user demographics, item attributes, and contextual information to improve recommendation quality.

Basilico et al.[24] proposed a unified framework that combines collaborative filtering and content-based filtering using supervised learning techniques. Their approach employed decision trees to model

user preferences, demonstrating improved recommendation performance by leveraging both user-item interactions and item attributes.

Render et al.[25] introduced Factorization Machines (FMs), a supervised learning algorithm that generalizes matrix factorization techniques. FMs can model interactions between features in high-dimensional sparse data, making them well-suited for recommender systems. This method has been shown to outperform traditional CF and CBF approaches in various recommendation tasks.

Burges et al.[26] developed a learning-to-rank framework for recommender systems using SVMs. Their approach focused on optimizing the ranking of recommended items based on user preferences, leading to improved recommendation quality.

Later, Burges et al.[27] extended their work on learning-to-rank by introducing RankNet, LambdaRank, and LambdaMART algorithms. These methods utilized gradient-based optimization techniques to directly optimize ranking metrics, further enhancing the performance of recommender systems.

In industrial applications, supervised learning techniques have been widely adopted to build effective recommender systems. And decision tree and ensemble methods are particularly popular due to their interpretability and ability to handle complex feature interactions.

He, et al.[28] developed a recommender system for Facebook Ads using gradient boosting decision trees. Their approach effectively modeled user preferences based on various features, leading to improved ad targeting and user engagement. This work exemplifies the successful application of supervised learning techniques in large-scale industrial recommender systems using decision trees.

Besides, some improved ensemble methods like XGBoost[29] and LightGBM[30] have also been applied in recommender systems to enhance recommendation accuracy and scalability. These methods leverage gradient boosting techniques to build robust models that can effectively capture complex user-item interactions.

3.1.3   Deep Learning-based Recommender Systems

Deep learning has revolutionized recommender systems by enabling the modeling of complex user-item interactions and capturing non-linear relationships. Neural Collaborative Filtering (NCF) employs multi-layer perceptrons to learn user-item interaction functions, outperforming traditional MF methods. Convolutional Neural Networks (CNNs) have been utilized to extract features from item content, such as images and text, enhancing recommendation quality. Recurrent Neural Networks (RNNs) have been applied to model sequential user behavior, capturing temporal dynamics in user preferences.

3.1.3.1   Neural Collaborative Filtering

Neural Collaborative Filtering (NCF) is a deep learning-based approach that utilizes neural networks to model user-item interactions.

He et al.[31] proposed the NCF framework, which formulates collaborative filtering as a neural interaction learning problem. Instead of assuming a fixed interaction function such as the inner product used in traditional matrix factorization, NCF leverages neural networks to learn flexible and nonlinear user – item interaction functions from data. Specifically, users and items are first represented as low-dimensional embedding vectors, which are then combined through neural architectures to capture complex interaction patterns that cannot be modeled by linear similarity measures.

Within the NCF framework, the authors introduced several representative model instantiations, including Generalized Matrix Factorization (GMF), which extends classical matrix factorization by generalizing the inner product into a learnable function, and MLP-based collaborative filtering, which employs multi-layer perceptrons to model high-order nonlinear interactions between user and item embeddings. Furthermore, the NeuMF model integrates GMF and MLP components into a unified architecture, enabling the framework to simultaneously capture both linear and nonlinear interaction signals.

By replacing manually designed interaction functions with data-driven neural networks, NCF significantly enhances the expressive capacity of collaborative filtering models while retaining the fundamental user – item interaction paradigm. This neural interaction learning framework has since been widely adopted as a standard baseline in recommender system research, particularly for evaluating the effectiveness of neural network – based interaction modeling.

3.1.3.2   Sequential Recommender Systems

Sequential Recommender Systems leverage sequential patterns in user behavior to provide personalized recommendations. This approach captures the temporal dynamics of user preferences by modeling the order of user interactions with items and are widely used in various applications, such as e-commerce, music streaming, and video platforms. The main techniques used in sequential recommender systems include Recurrent Neural Networks (RNNs), LSTM networks, and Transformer-based models.

Hidasi et al.[32] proposed a session-based recommender system using RNNs to model user behavior within sessions. Their approach effectively captures the sequential nature of user interactions, leading to improved recommendation accuracy.

The mainstream sequential recommender systems now are Transformer-based models. Kang and McAuley[33] introduced SASRec, a self-attentive sequential recommender system that utilizes the Transformer architecture to model user behavior. SASRec captures long-range dependencies in user interactions, enabling more accurate recommendations. SASRec utilizes the encoder part of the Transformer architecture to model user interaction sequences. By employing self-attention mechanisms, SASRec can effectively capture complex dependencies and patterns in user behavior over time. This allows the model to consider the entire sequence of user interactions when making recommendations, rather than relying solely on recent actions. The self-attention mechanism enables SASRec to weigh the importance of different items in the user's interaction history, allowing it to focus on relevant past behaviors that are most indicative of future preferences. This capability to model long-range dependencies and complex interaction patterns significantly enhances the recommendation quality, making SASRec a powerful approach for sequential recommendation tasks.

BERT4Rec[34] is another Transformer-based sequential recommender system that employs bidirectional self-attention to model user behavior. BERT4Rec captures contextual information from both past and future interactions, leading to enhanced recommendation performance. This structure is similar to the decoder part of the original Transformer architecture, where the model is trained to predict masked items in the user interaction sequence. By leveraging bidirectional self-attention, BERT4Rec can effectively capture complex dependencies and contextual information from both past and future interactions, enabling it to generate more accurate recommendations. This bidirectional modeling allows BERT4Rec to consider the entire sequence of user interactions when making predictions, rather than relying solely on preceding items. As a result, BERT4Rec can better understand user preferences and

provide more relevant recommendations, making it a powerful approach for sequential recommendation tasks.

3.1.3.3   Graph-based Recommender Systems

With the increasing popularity of graph neural networks (GNNs), graph-based recommender systems have gained significant attention in recent years. These systems leverage the inherent graph structure of user-item interactions to capture complex relationships and dependencies. The advantages of graph-based recommender systems include their ability to model high-order connectivity, incorporate side information, and handle dynamic interactions.

Van Den Berg et al.[35] proposed Graph Convolutional Matrix Completion (GCMC), which utilizes graph convolutional networks to model user-item interactions. GCMC effectively captures high-order connectivity in the interaction graph, leading to improved recommendation accuracy.

Ying et al.[36] introduced PinSage, a graph-based recommender system that combines GNNs with random walks to generate item embeddings. PinSage effectively incorporates side information and captures complex relationships in the interaction graph, enhancing recommendation quality. This method has been successfully deployed in large-scale industrial applications, such as Pinterest, demonstrating its scalability and effectiveness.

Wang et al.[37] proposed Neural Graph Collaborative Filtering (NGCF), which integrates GNNs into the collaborative filtering framework. NGCF captures high-order connectivity and complex user-item interactions, leading to significant improvements in recommendation performance.

He et al. [38] introduced LightGCN, a simplified graph convolutional network for recommender systems. LightGCN removes unnecessary components from traditional GNNs, focusing solely on neighborhood aggregation to enhance recommendation accuracy while reducing computational complexity. This streamlined architecture allows LightGCN to efficiently capture high-order connectivity in user-item interaction graphs, leading to improved recommendation performance with lower resource consumption. This method has become a popular baseline in graph-based recommender system research due to its effectiveness and efficiency.

Wu et al.[39] proposed SR-GNN, a session-based recommender system that utilizes GNNs to model user behavior within sessions. SR-GNN effectively captures complex relationships in the interaction graph, leading to enhanced recommendation accuracy. This work is a representation of combining graph-based methods with sequential recommender systems to leverage the strengths of both approaches.

3.1.3.4   Generative Recommender Systems

Generative Recommender Systems are newly emerging as a promising approach to enhance recommendation quality by leveraging generative models to capture complex user-item interactions and generate personalized recommendations. These systems utilize techniques such as Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and Normalizing Flows to model the underlying distribution of user preferences and item characteristics. Generative recommender systems offer several advantages, including the ability to model uncertainty in user preferences, generate diverse recommendations, and handle data sparsity effectively. This benefits can also be seen in our proposed method, which models user preferences as distributions.

Sedhain et al.[40] proposed AutoRec, a VAE-based recommender system that learns latent representations of users and items from interaction data. AutoRec effectively captures complex user-item interactions, leading to improved recommendation accuracy.

Liang et al.[41] introduced a VAE-based collaborative filtering approach that models user preferences as latent variables. This method captures uncertainty in user behavior, enhancing recommendation quality.

Wang et al.[42] proposed IRGAN, a GAN-based recommender system that unifies generative and discriminative models for information retrieval tasks. IRGAN effectively generates personalized recommendations by modeling user-item interactions through adversarial training.

Chae et al.[43] introduced CFGAN, a collaborative filtering framework based on GANs. CFGAN generates user-item interaction data to improve recommendation accuracy, particularly in scenarios with limited interaction data.

Besides these typical generative models, diffusion models have also been recently applied in recommender systems. Wang et al.[44] proposed DiffRec, a diffusion-based recommender system that leverages diffusion processes to model user-item interactions. DiffRec effectively captures complex relationships in the interaction data, leading to enhanced recommendation performance. This work represents the application of diffusion models in recommender systems, showcasing their potential to improve recommendation quality by modeling intricate user-item dynamics.

3.2   Cross-Domain Recommender Systems

3.2.1   Overlapping CDR

3.2.2   Non-overlapping CDR

3.3   Distributional User Preference Modeling

3.3.1   Multi-interest User Modeling

3.3.2   Probabilistic Preference Modeling

3.4   Optimal Transport for Representation Alignment

3.4.1   OT Basics in ML

3.4.2   OT in Domain Adaptation

Chapter 4

Proposed Method: Distributional Preference Modeling for Cross-Domain Recommendation

Chapter 5

Experiments

Chapter 6

Results

Chapter 7

Conclusion

References

[1] D. Bawden and L. Robinson. Information overload: An overview. In *Oxford Encyclopedia of Political Decision Making*. Oxford University Press, Oxford, June 2020. This is a draft of a chapter that has been published by Oxford University Press in the book Oxford Encyclopedia of Political Decision Making.

[2] Jun Lv and Xuan Liu. The impact of information overload of e-commerce platform on consumer return intention: Considering the moderating role of perceived environmental effectiveness. *International Journal of Environmental Research and Public Health*, Vol. 19, No. 13, 2022.

[3] Guihua Zhang, Junwei Cao, and Dong Liu. Examining the influence of information overload on consumers' purchase in live streaming: A heuristic-systematic model perspective. *PLOS ONE*, Vol. 18, No. 8, pp. 1–16, 08 2023.

[4] Paul Resnick and Hal R Varian. Recommender systems. *Communications of the ACM*, Vol. 40, No. 3, pp. 56–58, 1997.

[5] Robin Burke. Hybrid recommender systems: Survey and experiments. *User modeling and user-adapted interaction*, Vol. 12, No. 4, pp. 331–370, 2002.

[6] Robin Burke, Alexander Felfernig, and Mehmet H. Göker. Recommender systems: An overview. *AI Magazine*, Vol. 32, No. 3, pp. 13–18, Jun. 2011.

[7] Dietmar Jannach and Michael Jugovac. Measuring the business value of recommender systems. *ACM Trans. Manage. Inf. Syst.*, Vol. 10, No. 4, December 2019.

[8] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez. Recommender systems survey. *Knowledge-Based Systems*, Vol. 46, pp. 109–132, 2013.

[9] Xiaoyuan Su and Taghi M. Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in Artificial Intelligence*, Vol. 2009, No. 1, p. 421425, 2009.

[10] Paul Resnick, Neophytos Iacovou, Mitesh Suchak, Peter Bergstrom, and John Riedl. Grouplens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, CSCW '94, p. 175–186, New York, NY, USA, 1994. Association for Computing Machinery.

[11] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on World Wide Web*, pp. 285–295, 2001.

[12] John S. Breese, David Heckerman, and Carl Kadie. Empirical analysis of predictive algorithms for collaborative filtering, 2013.

[13] Lyle H Ungar and Dean P Foster. Clustering methods for collaborative filtering. In *AAAI workshop on recommendation systems*, Vol. 1, pp. 114–129. Menlo Park, CA, 1998.

[14] Yehuda Koren, Robert Bell, and Chris Volinsky. Matrix factorization techniques for recommender systems. *Computer*, Vol. 42, No. 8, pp. 30–37, 2009.

[15] Ruslan Salakhutdinov and Andriy Mnih. Bayesian probabilistic matrix factorization using markov chain monte carlo. In *Proceedings of the 25th International Conference on Machine Learning*, ICML '08, p. 880–887, New York, NY, USA, 2008. Association for Computing Machinery.

[16] Thomas Hofmann. Latent semantic models for collaborative filtering. *ACM Trans. Inf. Syst.*, Vol. 22, No. 1, p. 89–115, January 2004.

[17] Benjamin M Marlin. Modeling user rating profiles for collaborative filtering. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems*, Vol. 16. MIT Press, 2003.

[18] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, Vol. 3, No. Jan, pp. 993–1022, 2003.

[19] Gerard Salton, Chris Buckley, and Edward A Fox. Automatic query formulations in information retrieval. *Journal of the American society for information science*, Vol. 34, No. 4, pp. 262–280, 1983.

[20] Michael Pazzani and Daniel Billsus. Learning and revising user profiles: The identification of interesting web sites. *Machine learning*, Vol. 27, No. 3, pp. 313–331, 1997.

[21] Michael J Pazzani. A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, Vol. 13, No. 5, pp. 393–408, 1999.

[22] Prem Melville, Raymond J Mooney, Ramadass Nagarajan, et al. Content-boosted collaborative filtering for improved recommendations. *Aaai/iaai*, Vol. 23, pp. 187–192, 2002.

[23] Daniel Billsus and Michael J Pazzani. User modeling for adaptive news access. *User modeling and user-adapted interaction*, Vol. 10, No. 2, pp. 147–180, 2000.

[24] Justin Basilico and Thomas Hofmann. Unifying collaborative and content-based filtering. In *Proceedings of the twenty-first international conference on Machine learning*, p. 9, 2004.

[25] Steffen Rendle. Factorization machines. In *2010 IEEE International conference on data mining*, pp. 995–1000. IEEE, 2010.

[26] Chris Burges, Tal Shaked, Erin Renshaw, Ari Lazier, Matt Deeds, Nicole Hamilton, and Greg Hullender. Learning to rank using gradient descent. In *Proceedings of the 22nd international conference on Machine learning*, pp. 89–96, 2005.

[27] Christopher JC Burges. From ranknet to lambdarank to lambdamart: An overview. *Learning*, Vol. 11, No. 23-581, p. 81, 2010.

[28] Xinran He, Junfeng Pan, Ou Jin, Tianbing Xu, Bo Liu, Tao Xu, Yanxin Shi, Antoine Atallah, Ralf Herbrich, Stuart Bowers, et al. Practical lessons from predicting clicks on ads at facebook. In *Proceedings of the eighth international workshop on data mining for online advertising*, pp. 1–9, 2014.

[29] Tianqi Chen. Xgboost: A scalable tree boosting system. *Cornell University*, 2016.

[30] Guolin Ke, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, Vol. 30, , 2017.

[31] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural collaborative filtering. In *Proceedings of the 26th international conference on world wide web*, pp. 173–182, 2017.

[32] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*, 2015.

[33] Wang-Cheng Kang and Julian McAuley. Self-attentive sequential recommendation. In *2018 IEEE international conference on data mining (ICDM)*, pp. 197–206. IEEE, 2018.

[34] Fei Sun, Jun Liu, Jian Wu, Changhua Pei, Xiao Lin, Wenwu Ou, and Peng Jiang. Bert4rec: Sequential recommendation with bidirectional encoder representations from transformer. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management*, CIKM '19, p. 1441–1450, New York, NY, USA, 2019. Association for Computing Machinery.

[35] Rianne Van Den Berg, N Kipf Thomas, and Max Welling. Graph convolutional matrix completion. *arXiv preprint arXiv:1706.02263*, Vol. 2, No. 8, p. 9, 2017.

[36] Rex Ying, Ruining He, Kaifeng Chen, Pong Eksombatchai, William L Hamilton, and Jure Leskovec. Graph convolutional neural networks for web-scale recommender systems. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 974–983, 2018.

[37] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR'19, p. 165–174, New York, NY, USA, 2019. Association for Computing Machinery.

[38] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, YongDong Zhang, and Meng Wang. Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '20, p. 639–648, New York, NY, USA, 2020. Association for Computing Machinery.

[39] Shu Wu, Yuyuan Tang, Yanqiao Zhu, Liang Wang, Xing Xie, and Tieniu Tan. Session-based recommendation with graph neural networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33, No. 01, pp. 346–353, Jul. 2019.

[40] Suvash Sedhain, Aditya Krishna Menon, Scott Sanner, and Lexing Xie. Autorec: Autoencoders meet collaborative filtering. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, p. 111–112, New York, NY, USA, 2015. Association for Computing Machinery.

[41] Dawen Liang, Rahul G. Krishnan, Matthew D. Hoffman, and Tony Jebara. Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, p. 689–698, Republic and Canton of Geneva, CHE, 2018. International World Wide Web Conferences Steering Committee.

[42] Jun Wang, Lantao Yu, Weinan Zhang, Yu Gong, Yinghui Xu, Benyou Wang, Peng Zhang, and Dell Zhang. Irgan: A minimax game for unifying generative and discriminative information retrieval models. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '17, p. 515–524, New York, NY, USA, 2017. Association for Computing Machinery.

[43] Dong-Kyu Chae, Jin-Soo Kang, Sang-Wook Kim, and Jung-Tae Lee. Cfgan: A generic collaborative filtering framework based on generative adversarial networks. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, CIKM '18, p. 137–146, New York, NY, USA, 2018. Association for Computing Machinery.

[44] Wenjie Wang, Yiyan Xu, Fuli Feng, Xinyu Lin, Xiangnan He, and Tat-Seng Chua. Diffusion recommender model. In *Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '23, p. 832–841, New York, NY, USA, 2023. Association for Computing Machinery.

Appendix

A Graph-Enhanced EEG Foundation Models

Publications

- 王 力敏, 脇 聡志 and 鈴村 豊太郎. 大規模マルチモーダルモデルを用いたグラフィックレイアウトの自動生成に向けて. 2024 年度人工知能学会全国大会. 2024.

- Limin Wang, Toyotaro Suzumura, and Hiroki Kanezashi. Graph-Enhanced EEG Foundation Model. AAAI-25 Workshop on Large Language Models and Generative AI for Health. 2025.

Acknowledgment

In addition to funding, the completion of this master's thesis was made possible through the guidance and support from many individuals, to whom I would like to express my deepest gratitude.

First and foremost, I would like to thank Professor Toyotaro Suzumura, my academic advisor, for his extensive guidance and support throughout my research. He provided me with invaluable advice and direction in deciding on a research theme, selecting conferences and managing my schedule. His insights and suggestions on how to present my research in a clear and compelling manner were particularly helpful. I also appreciate his efforts to foster a positive atmosphere in the lab and a productive research environment, and his willingness to review my papers despite his busy schedule.

I would also like to thank Dr. Hiroki Kanezashi, a Research Scientist in our laboratory, for his numerous contributions to my research. He provided me with technical support, helped me with presentations, and reviewed my papers. His practical tips on how to conduct experiments efficiently were especially helpful, and his advice on how to improve my results was invaluable. I appreciate his kindness and willingness to help me with various tasks, from setting up computational environments to handling travel arrangements.

I am also profoundly grateful to Dr. Masatoshi Hanai, an Assistant Professor in our laboratory, for his guidance and support. His suggestions on how to approach my research and his valuable insights on the research topics were particularly helpful. I also appreciate his comprehensive support, from helping me set up the research environment and equipments, to engaging in daily conversations that were both enjoyable and informative, which has greatly enhanced my research experience.

Furthermore, I would like to thank Satoshi Waki, a senior alumnus of our laboratory, for his mentorship and guidance. He provided me with valuable advice on how to choose a research topic and how to find directions, and his insights on how to conduct research were particularly helpful. His willingness to engage in discussions and provide feedback on my ideas was invaluable, and his advice on the research life and the future career goals was especially helpful.

Last but certainly not least, I would like to acknowledge the unwavering support and encouragement of my family, without whom I would not have been able to complete my research.

Once again, I would like to express my heartfelt gratitude to everyone who has supported me throughout my research life. I am deeply grateful for your kindness, guidance, and encouragement, and I am honored to have had the opportunity to work with such a wonderful group of people. Thank you.