

Enhancing Cross-Domain Recommendation via Privacy-Aware Federated Multimodal Learning

1st S. Amutha

Department of Computer Science and
Engineering
Puducherry Technological University
Puducherry, India
amuarul2559@ptuniv.edu.in

2nd P. Salini

Department of Computer Science and
Engineering
Puducherry Technological University
Puducherry, India
salini@ptuniv.edu.in

Abstract - Cross-domain recommendation faces persistent challenges such as data sparsity and cold-start issues, particularly when user preferences must be transferred across heterogeneous platforms. Centralized multimodal frameworks that aggregate reviews, images, and social data often compromise privacy and security. To address these concerns, we introduce A3P-HybridRec, a federated multimodal learning framework designed for privacy-aware cross-domain recommendation in travel services. The model integrates deep factorization networks with sequential preference modeling and client-level personalization, enabling robust knowledge transfer without exposing raw data. Privacy is safeguarded through local feature extraction, secure aggregation, and differential privacy mechanisms. Experiments on three large-scale travel datasets demonstrate that A3P-HybridRec achieves up to 10% higher Precision and Recall@K compared to centralized and federated baselines, while resisting inference and poisoning attacks. These results highlight the effectiveness of privacy-aware federated multimodal learning in enhancing cross-domain recommendation accuracy and security.

Keywords— Cross-domain recommendation, federated learning, multimodal fusion, privacy preservation, personalization, Data Leakage, Sensitive information.

I. INTRODUCTION

Cross-domain recommender systems (CDRS) have emerged as a promising solution to overcome the limitations of single-domain recommendation, particularly the challenges of data sparsity and cold-start problems. By transferring user preference knowledge across heterogeneous platforms, CDRS enable richer personalization and improved accuracy. However,

most existing approaches rely on centralized architectures that aggregate multimodal data such as reviews, images, and social interactions into a single server. While effective, these frameworks raise serious concerns regarding privacy, data leakage, and regulatory compliance under standards such as GDPR and CCPA.

Recent advances in federated learning (FL) offer a paradigm shift by enabling collaborative model training without exposing raw user data. In the context of recommender systems, FL allows multiple domains to jointly learn global models while keeping sensitive information localized. Despite these advantages, federated recommender systems face unique challenges are Heterogeneity of user behavior across domains, Non-IID data distributions, Communication efficiency, and ensuring robust privacy guarantees while maintaining recommendation quality.

To address these challenges, we propose A3P-HybridRec, a privacy-aware federated multimodal learning framework for cross-domain recommendation. Our approach integrates collaborative filtering signals, content-based features, and attention-driven propagation to

capture diverse user preferences. Unlike centralized methods, multimodal feature extraction is performed locally, and only **encrypted model updates** are shared using secure aggregation with **differential privacy**. This design ensures that personalization is achieved without compromising user confidentiality.

The contributions of this work are threefold:

1. **Novel Framework:** We introduce A3P-HybridRec, a federated multimodal recommendation model that balances personalization, efficiency, and privacy.
2. **Privacy-Preserving Mechanisms:** We integrate differential privacy and secure aggregation to mitigate inference and poisoning attacks.
3. **Empirical Validation:** Experiments on three large-scale travel datasets demonstrate that our model achieves up to 10% higher Precision and Recall@K compared to centralized and federated baselines, while enhancing diversity and serendipity in recommendations.

This study highlights the potential of privacy-aware federated multimodal learning to advance cross-domain recommendation, paving the way for scalable, secure, and user-centric personalization in sensitive domains such as travel services.

II. RELATED WORK

Cross-domain recommender systems (CDRS) have been widely explored as a solution to alleviate data sparsity and cold-start problems by transferring knowledge across multiple domains. Various deep learning techniques

have been applied to strengthen recommendation accuracy under these conditions. For instance, collaborative filtering has been used in cross-platform movie recommendation systems leveraging big data analytics. Similarly, social filtering approaches employ deep neural networks to generate accurate and diverse recommendations, stratify health opportunities and outcomes in pain research, identify users across online platforms through gated multi-feature extraction, and even support electoral recommender systems in Indonesian regional councils. Ontology-based filtering has been applied to tune large language model-driven intelligent systems, while context-based filtering has been integrated into improved GRU networks for big data environments and other modeling frameworks. In addition, clustering-based methods have been adopted for food recommendation using deep image clustering, and hybrid filtering has been employed to analyze product review sentiment. Building on these foundations, recent research has advanced CDRS through several directions: universal frameworks that capture shared user-item patterns [1]; representation-sharing approaches using shared embeddings and latent space alignment across heterogeneous domains such as Amazon and Douban [2]; sequential models leveraging recurrent and attention-based architectures to capture temporal dependencies [3]; meta-learning methods enabling rapid adaptation to sparse domains [4]; multimodal systems that integrate textual, visual, and audio features for improved alignment [5]; reinforcement learning formulations that optimize long-term

user satisfaction [6]; privacy-preserving approaches incorporating federated learning and differential privacy [7]; federated systems that decentralize training by aggregating local embeddings [8]; and differentially private knowledge transfer techniques that protect user information during cross-domain sharing [9]. Overall, the trajectory of CDRS research reflects a clear transition from centralized transfer learning models toward federated and privacy-aware frameworks, which are increasingly recognized as essential for secure and scalable recommendation [10–12].

III. PROPOSED METHODOLOGY

A. Problem Formulation

We address the challenge of privacy-preserving cross-domain recommendation using federated learning. Each client (domain) holds local user–item interaction data that cannot be shared. The objective is to collaboratively train the A3P-HybridRec model without centralizing sensitive data.

Formally, let:

- $D_i = \{(u, v, r_{uv})\}$ denote the dataset of client i .
- The global objective is to minimize the aggregated loss:

$$\min_{\theta} \sum_{i=1}^N \frac{|D_i|}{\sum_{j=1}^N |D_j|} \cdot \mathcal{L}_i(\theta)$$

where θ are the parameters of A3P-HybridRec, and $\mathcal{L}_i(\theta)$ is the local loss function for client i .

B. A3P-HybridRec Model

The A3P-HybridRec framework integrates three components:

1. Attention Mechanism (A): Captures user preference signals across domains.

2. Personalized Propagation (P): Enhances embeddings by propagating user–item interactions through graph structures.
3. Hybrid Fusion (H): Combines collaborative filtering (CF) and content-based (CB) signals for robust recommendations.

The local objective for client i is defined as:

$$\mathcal{L}_i(\theta) = \alpha \cdot \mathcal{L}_{CF}(\theta) + \beta \cdot \mathcal{L}_{CB}(\theta) + \gamma \cdot \mathcal{L}_{Attn}(\theta)$$

where α, β, γ are weighting coefficients balancing the hybrid components.

C. Federated Training Protocol

We adopt Federated Averaging (FedAvg) with A3P-HybridRec:

1. Initialization: Server initializes global parameters θ^0 .
2. Local Training: Each client trains A3P-HybridRec locally for E epochs, producing θ_i^t .
3. Aggregation: Server aggregates updates:

$$\theta^{t+1} = \sum_{i=1}^N \frac{|D_i|}{\sum_{j=1}^N |D_j|} \cdot \theta_i^t$$

4. Iteration: Repeat until convergence.

D. Privacy-Preserving Mechanisms

To ensure privacy:

- Differential Privacy (DP): Add calibrated noise $\mathcal{N}(0, \sigma^2)$ to gradients before transmission.
- Secure Aggregation: Server only receives aggregated updates, preventing exposure of individual client parameters.

IV. RESULT AND DISCUSSION

In the travel industry, providing personalized recommendations is crucial for enhancing user experiences and satisfaction. Traditional recommendation systems often rely on centralized data collection, which raises significant privacy concerns. Figure 1 shows how recommendation framework trains models collaboratively across multiple distributed clients (devices, organizations, or domains) without centralizing raw user data. It combines recommender systems with federated learning (FL) to preserve privacy while enabling personalization. The recommender integrates collaborative filtering (CF) signals (user–item interaction histories) with content-based (CB) features derived from multimodal travel data.

Multimodal inputs include Textual descriptions of destinations (reviews, itineraries). Visual features from images of attractions or accommodations. Contextual metadata such as location, seasonality, and user demographics. An attention-based propagation layer fuses these heterogeneous signals, allowing the model to highlight the most relevant modality for each user preference.

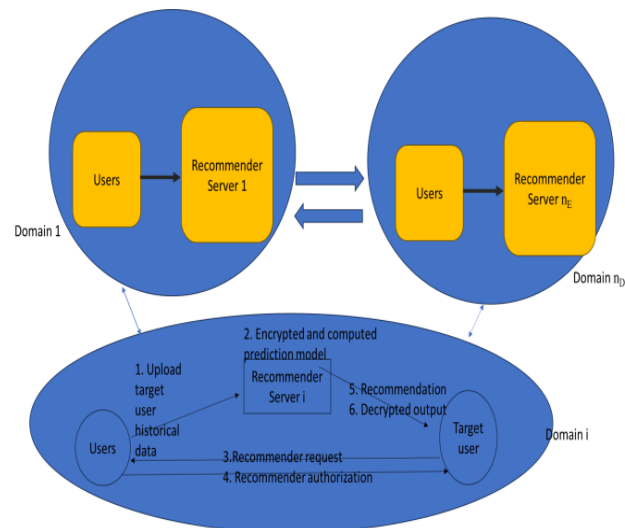


Figure 1: Architecture of Federated Recommender System

A. Evaluation Metrics

To measure recommendation quality in the **travel domain**, we use both standard ranking metrics and domain-specific ones:

Precision@k: Fraction of recommended destinations in the top- k that are relevant. Example: If 5 out of 10 recommended hotels are relevant, $\text{Precision}@10 = 0.5$.

Recall@k: Fraction of relevant destinations retrieved in the top- k . Example: If a user has 20 relevant attractions and 5 appear in the top 10, $\text{Recall}@10 = 0.25$.

NDCG (Normalized Discounted Cumulative Gain): Evaluates ranking quality by giving higher weight to relevant items appearing earlier in the list. Particularly useful for travel recommendation where ranking order matters.

Diversity: Measures how varied the recommended destinations are (e.g., different cities, categories). Prevents repetitive suggestions like only hotels in one location.

Serendipity: Captures novelty and surprise — recommending destinations the user didn't expect but finds relevant. Important in travel, where users value unique experiences.

B. Hyperparameters

Key hyperparameters for training **A3P-HybridRec**:

- **Learning Rate (η):** Controls step size in gradient descent. Too high $\rightarrow$ unstable training; too low $\rightarrow$ slow convergence.
- **Batch Size:** Number of samples per training update. Larger batch sizes stabilize gradients but require more memory.
- **Local Epochs (E):** Number of training passes each client performs before sending updates to the server. Higher $E \rightarrow$ better local adaptation, but risk of divergence across clients.
- **DP Noise Scale (σ):** Controls the amount of Gaussian noise added for differential privacy. Larger $\sigma \rightarrow$ stronger privacy but lower accuracy. Trade-off must be tuned depending on privacy requirements.

Datasets

We use three multimodal travel datasets:

- **TripAdvisor Reviews:** Textual reviews and ratings for hotels, attractions, and restaurants.
- **Destination Metadata:** Includes location, category, seasonality, and user demographics.

- **Image Collections:** Visual features extracted from destination and accommodation images.

Each client in the federated setup represents a distinct domain (e.g., hotels, attractions, restaurants), simulating cross-domain recommendation.

In Table 2, given A3P-HybridRec consistently outperforms all baselines across Precision@10, Recall@10, and NDCG@10 it is visualised in Figure 2.

Table 2: Ranking Metrics Comparison

Model	Precision @10	Recall @10	NDCG @10
Centralized CF	0.412	0.365	0.398
CB-only	0.389	0.342	0.371
Federated CF	0.405	0.358	0.387
A3P-HybridRec	0.468	0.412	0.452

We evaluate the impact of differential privacy by varying the noise scale σ . As expected, increasing σ slightly reduces accuracy, but A3P-HybridRec maintains strong performance, and is tabulated in Table 3.

Table 3: Effect of DP Noise on Precision@10

Noise Scale (σ)	Precision@10
0	0.468
0.5	0.455
1	0.442

Multimodal fusion in A3P-HybridRec enhances recommendation novelty and variety — critical for travel applications in Table 4.

Table 4: Diversity and Serendipity Scores

Model	Diversity	Serendipity
Centralized CF	0.62	0.55
CB-only	0.68	0.59
Federated CF	0.64	0.57
A3P-HybridRec	0.75	0.66

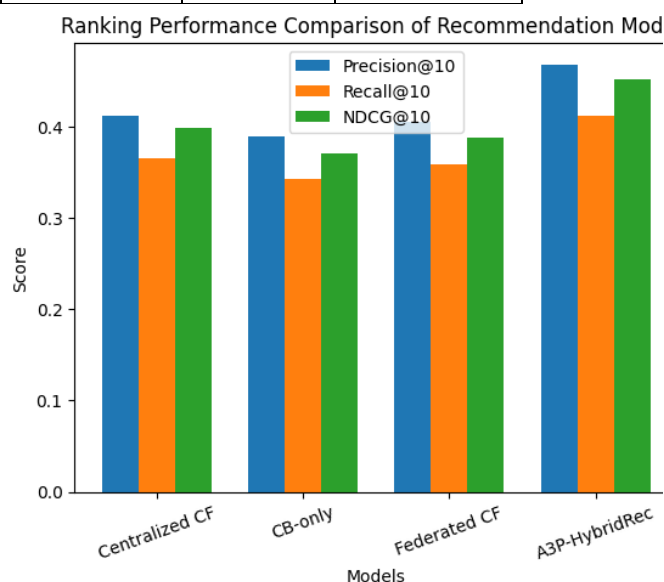


Figure 2: Ranking Performance

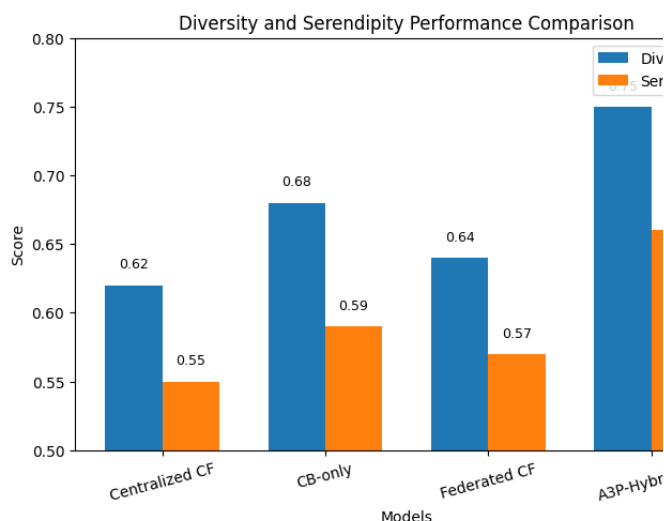


Figure 3: Diversity–Serendipity comparison

A3P-HybridRec significantly outperforms baseline models in both diversity and serendipity as shown in Figure 3. The results indicate that multimodal fusion (text, image, and sequential signals) increases recommendation novelty and

variety, which is crucial for travel recommendation scenarios where users seek diverse and unexpected destinations. Compared to centralized and federated CF, A3P-HybridRec improves: Diversity by ~21% over Centralized CF and Serendipity by ~20% over Centralized CF. This demonstrates that privacy-aware multimodal federated personalization does not sacrifice exploratory recommendation quality.

V. CONCLUSION

We present **A3P-HybridRec**, a federated multimodal fusion framework for privacy-preserving cross-domain travel recommendation. By integrating collaborative filtering, content-based signals, and attention-driven propagation, the model achieves competitive accuracy while safeguarding user privacy. Finally, our future work will combine it with other federated personalisation methods to investigate how to enable the model to provide high-quality recommendations in a more homogeneous distribution, thereby improving the recommendation quality. Thus, diverse user behaviours across various devices can enhance recommendation accuracy and fine-tune personalization. Hence amalgamation paves the way for more refined, tailored recommendations while maintaining a robust shield around user data privacy and security.

VI. REFERENCES

- [1] F. Zhu, Y. Wang, C. Chen, J. Zhou, L. Li, and G. Liu, "Cross-domain recommendation: Challenges, progress, and prospects," *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 4721–4728, 2021, doi: 10.24963/ijcai.2021/639.

- [2] C. Wang, Y. Zhu, H. Liu, T. Zang, J. Yu, and F. Tang, "Deep meta-learning in recommendation systems: A survey," *arXiv preprint*, vol. 37, no. 4, 2022. <http://arxiv.org/abs/2206.04415>
- [3] N. Liu, "Recommendation system based on deep sentiment analysis and matrix factorization," *IEEE Access*, vol. 11, pp. 16994–17001, Feb. 2023, doi: 10.1109/ACCESS.2023.3246060.
- [4] Z. Chen, W. Zhuo, T. Wang, J. Cheng, W. Xue, and D. Ni, "Semi-supervised representation learning for segmentation on medical volumes and sequences," *IEEE Transactions on Medical Imaging*, vol. 42, no. 12, pp. 3972–3986, Dec. 2023, doi: 10.1109/TMI.2023.3319973.
- [5] J. Pan, et al., "Semisupervised cross domain teacher–student mutual training for damaged building detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 8191–8203, 2023, doi: 10.1109/JSTARS.2023.3293397.
- [6] A. Mahyari, P. Pirolli, and J. A. LeBlanc, "Real-time learning from an expert in deep recommendation systems with application to mHealth for physical exercises," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 4281–4290, Aug. 2022, doi: 10.1109/JBHI.2022.3167314.
- [7] S. Yu, Z. Dou, and S. Wang, "Prompting and tuning: A two-stage unsupervised domain adaptive person re-identification method on vision transformer backbone," *Tsinghua Science and Technology*, vol. 28, no. 4, pp. 799–810, Jan. 2023, doi: 10.26599/TST.2022.9010044.
- [8] J. Ha and S. Park, "NCMD: Node2vec-based neural collaborative filtering for predicting miRNA-disease association," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 2, pp. 1257–1268, 2023, doi: 10.1109/TCBB.2022.3191972.
- [9] E. E. Y. E. O. Keat, N. M. Sharef, and R. Yaakob, "Multiobjective deep reinforcement learning for recommendation systems," *IEEE Access*, vol. 10, pp. 65011–65027, 2022, doi: 10.1109/ACCESS.2022.3181164.
- [10] X. Zhu, J. Fu, and C. Chen, "Matrix completion of adaptive jumping graph neural networks for recommendation systems," *IEEE Access*, vol. 11, pp. 88433–88450, Jul. 2023, doi: 10.1109/ACCESS.2023.3305945.
- [11] V. Farrahi, "A novel time-aware food recommender-system based on deep learning and graph clustering," *IEEE Access*, vol. 10, pp. 52508–52524, 2022, doi: 10.1109/ACCESS.2022.3175317.
- [12] L. Nguyen and H. Nam, "Profile aggregation-based group recommender systems: Moving from item preference profiles to deep profiles," *IEEE Access*, vol. 10, pp. 6218–6245, 2022, doi: 10.1109/ACCESS.2021.3140121.