

Graph Neural Networks for User Preference Prediction in Multi-Domain Recommendation Systems

Uday D. Yadav

School of Computing, Clemson University, Clemson, SC, USA.
contactuday@clemson.edu

Olivier Terry

Department of Computer Science, University of New Hampshire, Durham, NH, USA.
olivierterry@unh.edu

Harish Tandon

Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS, USA.
harisht@ku.edu

Abstract

Multi-domain recommendation systems must reconcile heterogeneous user interaction traces, item catalogs, contextual signals, and business constraints within a coherent infrastructure. This paper presents a system-level examination of graph neural networks for user preference prediction across domains. Rather than focusing on isolated model accuracy, the discussion addresses how relational inductive biases enable unified representations of users, items, attributes, contexts, and cross-domain transitions. The paper examines architectural choices for message passing over large-scale heterogeneous graphs, the incorporation of categorical and context-aware query signals, and the implications of multi-domain data sharing for privacy, governance, and operational stability. It further analyzes structural trade-offs in balancing domain-specific specialization with shared preference knowledge, the allocation of computational resources in continuous training pipelines, and the management of concept drift across catalogs with different temporal dynamics. Fairness and robustness receive particular attention because multi-domain graphs may propagate bias through cross-domain edges and create feedback loops that affect downstream user exposure. The paper also considers sustainability through sampling strategies, sparsified propagation, and lifecycle-aware system design. The conclusion identifies open challenges for future research, including cross-domain causal identification, federated graph learning, and accountable personalization governance. Overall, the paper argues that graph neural networks offer a promising systemic foundation for multi-domain recommendation but require careful architectural, infrastructural, and policy-level design to realize their benefits responsibly.

Keywords

graph neural networks; multi-domain recommendation; user preference modeling; recommendation infrastructure; fairness; data governance.

1. Introduction

The increasing fragmentation of digital services has created a persistent challenge for personalization infrastructures. Users interact with streaming catalogs, e-commerce inventories, local service directories, social feeds, and news platforms, often under different identities or through shared account contexts. A recommendation system that treats each domain independently may miss transferable preference signals such as durable interest in a topic, price sensitivity, aesthetic style, or geographic routine. At the same time, directly merging domain-specific datasets introduces heterogeneities in item metadata schema, interaction density, temporal granularity, and user intent. These conditions make user preference prediction less a problem of fitting a single matrix factorization and more a problem of constructing a relational representation that can adapt to multiple data regimes. The need for systemic reasoning has led researchers to consider graph-based abstractions in which users, items, contexts, and domain markers are nodes connected by observed interactions and semantic relationships.

Graph neural networks have emerged as a robust class of methods for learning on such relational structures. Rather than assuming independence between records, graph models propagate information along edges to produce context-sensitive node representations. Early formulations of graph convolutional matrix completion demonstrated that collaborative filtering can be expressed as link prediction on bipartite user-item graphs, where the embedding of a node is refined by its neighborhood [1]. Inductive learning on large graphs further enabled the generation of representations for previously unseen nodes, an essential property when new users or items arrive frequently in production systems [2]. Semi-supervised graph convolution provided the basic mechanism of localized feature aggregation and nonlinear transformation [3], while attention-based graph models introduced the ability to weight neighboring nodes differently according to learned relevance [4]. These foundations underpin many subsequent recommendation architectures.

A comprehensive survey of graph neural network methods highlights their versatility in capturing relational patterns beyond traditional dyadic interactions [5]. Industrial deployment at web scale has shown that traversal-based graph convolution can support recommendation over heterogeneous item graphs containing billions of nodes and edges [6]. Simplified graph convolution, as exemplified by LightGCN, demonstrates that careful removal of nonlinear feature transformations can improve collaborative filtering performance while reducing computational overhead [7]. Knowledge graph attention networks integrate side information from entity relations, allowing preference predictions to reflect semantic associations among items [8]. Social recommendation models leverage trust and friendship graphs to regularize user representations [9]. Multi-behavior recommendation systems use distinct interaction types such as view, click, cart, and purchase to enrich preference learning [10]. Cross-domain recommendation has been surveyed as a distinct problem with challenges including domain alignment, privacy, and transferability [11]. Together, these lines of work provide the technical foundation for a system-level treatment of multi-domain preference prediction.

2. Graph-Based Preference Modeling Across Heterogeneous Domains

A multi-domain graph representation unifies otherwise fragmented behavioral signals. In such a representation, users are connected to items across domains through domain-specific interaction edges. Items are connected to categorical nodes that describe type, genre, location, brand, or price band. Contextual nodes may encode session time, device, query category, or external event. The graph can also include domain nodes connected to the catalog elements they govern. Message passing operates over these heterogeneous edges, allowing a user

representation to be influenced not only by direct interactions but also by second-order relations such as users who interacted with similar local services or items that share a category with previously consumed content. The benefit of this design is that transferable preferences can emerge without requiring a unified feature schema across all domains. Each domain can retain its own metadata structure while contributing to shared user nodes and contextual nodes.

In practice, domain heterogeneity requires careful treatment of edge semantics. A click in a short-video domain may be impulsive and have limited preference durability, whereas a repeated purchase in an e-commerce domain may carry stronger signal. Treating all edges as equivalent can inject noise and dilute domain-specific structure. Attention mechanisms can learn domain-aware edge weighting, but the system must supply adequate supervision for such weights. Contextual retrieval remains important for local categorical query understanding, where user intent is expressed through short queries that combine categorical constraints and geographic or temporal context [12]. This type of signal is valuable because it grounds cross-domain recommendation in concrete information needs rather than only in passive behavioral similarity.

The representational capacity of a graph neural network depends on its ability to distinguish non-isomorphic local structures and to aggregate features from relevant higher-order neighborhoods [13]. In multi-domain graphs, some of the most predictive structures involve motifs such as user-item-category-item-user, which connect different domains through shared categorical nodes. Methods that explicitly incorporate latent factor models with graph propagation can improve recommendation accuracy by combining global low-rank patterns with local graph smoothing [14]. Multi-component graph convolutional collaborative filtering further suggests that separating the modeling of user-user, item-item, and user-item relations can reduce interference among heterogeneous signals [15]. These architectural insights matter in multi-domain settings where interference is a central risk.

3. System Architecture and Infrastructure Design

Deploying graph neural networks for multi-domain recommendation requires an infrastructure that supports graph construction, feature storage, sampling, training, inference, and monitoring. The architecture usually separates offline graph processing from online prediction. Offline pipelines ingest interaction logs, domain catalogs, and context tables; resolve identities; build edges; and compute node features. Online systems retrieve precomputed embeddings or perform lightweight propagation in a serving graph. The boundary between offline and online should be governed by latency requirements and freshness. For example, user interest signals in local categorical query understanding may change rapidly, while durable item metadata changes slowly. A single training cadence can therefore be suboptimal across domains.

A distributed graph store must support heterogeneous node and edge types, versioned snapshots, and low-latency neighborhood access. The graph is often sharded by entity type and geography, with edge sampling being the main control point for computational cost. Sampling strategies introduce trade-offs between estimate bias, variance, and infrastructure load. Neighborhood sampling can make training scalable but may underrepresent high-degree nodes or rare cross-domain edges. The platform must log sampling decisions to support reproducibility and auditability. In multi-domain systems, cross-domain edges may be relatively sparse, so naïve uniform sampling can systematically suppress transfer learning. Weighted sampling or stratified sampling by domain and edge type is therefore an important

operational mechanism, though it introduces additional governance complexity. When data cannot be centralized because of privacy or regulatory constraints, federated learning provides an infrastructure pattern in which local graph models are trained on decentralized clients and aggregated centrally [16]. This pattern is especially relevant when domains correspond to separate organizational units or jurisdictions.

Training pipelines for graph models are not isolated model-fitting exercises; they are production systems with hidden technical debt that accumulates through data dependencies, configuration complexity, and entangled feature pipelines [17]. In a multi-domain graph, changes to one domain catalog can alter the semantics of shared categorical nodes and thus affect model behavior in other domains. The infrastructure must therefore include contract tests for schema evolution, lineage tracking for edge construction, and staged rollouts for model updates. Operational monitoring should distinguish between domain-specific performance degradation and cross-domain contamination. Without such monitoring, a model may appear stable at the global level while silently failing for a minority domain or local context.

4. Governance, Fairness, and Robustness

Multi-domain recommendation systems create fairness risks that are more complex than those in single-domain systems. When a graph connects users and items across domains, bias can propagate along cross-domain edges, making it difficult to isolate the source of a discriminatory pattern. Algorithmic fairness research has documented many ways in which machine learning systems can inherit or amplify social biases through data, representation, and optimization choices [18]. In legal and policy scholarship, disparate impact analysis has shown that neutral system design can still produce materially unequal outcomes for protected groups when historical data encode structural disadvantage [19]. These concerns are not hypothetical in recommendation systems because exposure decisions shape access to information, economic opportunity, and local services.

Graph-based models may amplify such effects because homophily and message passing can concentrate disadvantage within certain neighborhoods. If a user cluster from one domain is underrepresented, cross-domain edges may spread that underrepresentation into other domains. This propagation mechanism complicates accountability because a disparate outcome in one domain may originate from another domain's data collection practices or edge construction choices. From a systemic perspective, fairness cannot be guaranteed solely by removing sensitive attributes from user features. Because graph message passing aggregates information from neighbors, sensitive signals may be indirectly encoded in item categories, interaction frequency, or social connections. A model may therefore reproduce discriminatory outcomes even when explicit demographic variables are excluded. Counterfactual fairness provides a framework for reasoning about whether a decision would change under a controlled intervention on a protected attribute, but evaluating such counterfactuals in a multi-domain graph requires strong causal assumptions [20]. Meanwhile, fair graph neural network research highlights that representation learning on biased graphs can amplify disparity through homophily and degree imbalance, necessitating fairness constraints during training or post-hoc audit [21].

Robustness concerns in multi-domain graphs extend beyond fairness. Recommendation models are exposed to noisy interactions, injected spam, and adversarial manipulation from users or sellers. Adversarial training has been proposed to improve the stability of personalized ranking models under perturbation, but such techniques add training cost and

may reduce accuracy on clean data [22]. In production, concept drift occurs when user preferences, catalogs, or external events change the distribution of interactions. A multi-domain graph may exhibit heterogeneous drift rates: a local services domain may respond rapidly to seasonality, while a media domain evolves more gradually. Adaptation mechanisms must therefore be domain-aware, with different refresh intervals and replay buffers per domain [23]. A unified global model that ignores these differences risks overfitting to fast-changing domains while underfitting stable ones, or vice versa.

5. System Governance, Explainability, and Sustainability

Multi-domain graph models operate within a broader governance context that includes regulatory compliance, internal audit, and stakeholder accountability. The complexity of message passing over heterogeneous graphs makes it difficult to trace why a particular recommendation was produced, especially when cross-domain edges contribute indirectly to a user embedding. Explainable recommendation techniques attempt to provide post hoc rationales by identifying influential paths or neighboring items, but the fidelity of such explanations in multi-domain settings remains an open problem [24]. Without transparent provenance for edge construction and feature aggregation, auditors cannot verify whether a system is compliant with data protection principles or whether a recommendation was influenced by a protected attribute. Governance frameworks should therefore require logging of graph snapshots, training configurations, and sampling decisions alongside model artifacts. This practice aligns with software engineering principles for maintainable machine learning systems and supports reproducibility in audits.

Sustainability is also a structural concern. Large graph neural networks require repeated neighborhood sampling, feature propagation, and gradient computation across billions of edges. Training and inference costs can grow quickly when multi-domain graphs contain dense cross-domain connections. Efficient training algorithms such as cluster-based graph convolution reduce memory and communication overhead by partitioning the graph into dense subgraphs, but they may introduce additional bias if domains are unevenly represented in cluster partitions [25]. From a system design perspective, sustainability requires not only algorithmic efficiency but also a lifecycle-aware approach: models should be retired or compressed when their marginal value declines, and graph snapshots should be archived according to retention policies. A multi-domain platform that continuously retrains without evaluating the marginal utility of new domains may incur unnecessary energy consumption and infrastructure cost, undermining both operational and environmental sustainability.

6. Conclusion

This paper has examined multi-domain recommendation from a system-level perspective, emphasizing the structural, governance, and operational implications of graph neural network architectures. Graph representations provide a natural mechanism for unifying heterogeneous user-item interactions, contextual signals, and cross-domain transfer, but they also introduce new risks related to bias propagation, opacity, and computational overhead. The analysis suggests that the value of graph neural networks lies not merely in their predictive accuracy but in their capacity to make relational dependencies explicit and auditable. Future research should prioritize causal identification of cross-domain effects, federated graph learning that preserves local data control, and standardized benchmarks for fairness and robustness in multi-domain settings. Practical deployments will require collaboration between system architects, fairness practitioners, and domain experts to ensure that the infrastructure supports responsible personalization at scale. As recommendation systems continue to expand across

domains, the ability to govern their graph-based representations will become a defining characteristic of trustworthy and sustainable AI.

References

1. van den Berg, R., Kipf, T. N., & Welling, M. (2017). Graph convolutional matrix completion. arXiv preprint arXiv:1706.02263.
2. Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems* (pp. 1024–1034).
3. Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In *Proceedings of the International Conference on Learning Representations*.
4. Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., & Bengio, Y. (2018). Graph attention networks. In *Proceedings of the International Conference on Learning Representations*.
5. Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2021). A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1), 4–24.
6. Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W. L., & Leskovec, J. (2018). 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).
7. He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., & Wang, M. (2020). 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* (pp. 639–648).
8. Wang, X., He, X., Cao, Y., Liu, M., & Chua, T.-S. (2019). KGAT: Knowledge graph attention network for recommendation. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 950–958).
9. Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J., & Yin, D. (2019). Graph neural networks for social recommendation. In *Proceedings of the World Wide Web Conference* (pp. 417–426).
10. Jin, B., Gao, C., He, X., Jin, D., & Li, Y. (2020). Multi-behavior recommendation with graph convolutional networks. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 659–668).
11. Zhu, F., Wang, Y., Chen, C., Zhou, J., Li, L., & Liu, G. (2021). Cross-domain recommendation: challenges, progress and prospects. arXiv preprint arXiv:2103.01696.
12. Yu, X. (2026, January). AI-Driven Personalization across Domains for Local Categorical Query Understanding and Context-Aware Retrieval. In *Proceedings of the 2nd International Conference on Artificial Intelligence, Digital Media Technology and Social Computing* (pp. 77-82).
13. Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2019). How powerful are graph neural networks? In *Proceedings of the International Conference on Learning Representations*.

14. Wang, X., He, X., Wang, M., Feng, F., & Chua, T.-S. (2019). Neural graph collaborative filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 165–174).
15. Wang, X., Jin, H., Zhang, A., He, X., Xu, T., & Chua, T.-S. (2020). Disentangled graph collaborative filtering. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 1001–1010).
16. Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2), 1–19.
17. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J.-F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. In *Advances in Neural Information Processing Systems* (pp. 2503–2511).
18. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6), 1–35.
19. Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671–732.
20. Kusner, M. J., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. In *Advances in Neural Information Processing Systems* (pp. 4066–4076).
21. Dai, E., & Wang, S. (2021). Say no to the discrimination: Learning fair graph neural networks with limited sensitive attribute information. In *Proceedings of the 14th ACM International Conference on Web Search and Data Mining* (pp. 680–688).
22. He, X., He, Z., Du, X., & Chua, T.-S. (2018). Adversarial personalized ranking for recommendation. In *Proceedings of the 41st International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 355–364).
23. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), 1–37.
24. Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. *Foundations and Trends in Information Retrieval*, 14(1), 1–101.
25. Chiang, W.-L., Liu, X., Si, S., Li, Y., Bengio, S., & Hsieh, C.-J. (2019). Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 257–266).