

Article

Semantic Federated Clustering for Cross-Domain Recommendation with Interaction Graphs

Deling Huang ^{1,*} , Li Luo ¹  and Chang Su ²
¹ School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; s241231040@stu.cqupt.edu.cn

² Admissions Office, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; suchang@cqupt.edu.cn

* Correspondence: huangdl@cqupt.edu.cn

Abstract

In federated cross-domain recommendation, the existing research has widely focused on scenarios where at least some users or items overlap. How to enable cross-domain recommendation over non-overlapping users and items has emerged as a challenging research direction. However, current privacy-preserving cross-domain recommendation methods struggle to achieve effective semantic alignment across domains in such scenarios, hindering their ability to capture latent correlations between interaction graphs. Moreover, when refining noisy node representations and interaction sequences, these methods often overlook temporal noise implicitly embedded in semantic information, as well as the potential of jointly leveraging textual and temporal features for collaborative denoising. To address these challenges, this paper proposes a Semantic Federated Clustering-driven Cross-Domain Recommendation model (SFC-CDR). Specifically, we first design a clustering-anchored contrastive learning framework that performs clustering using weighted multi-layer node semantic representations. The weighted aggregated centroids serve as semantic anchors to dynamically update the pool of positive and negative samples in contrastive learning, thereby achieving geometric alignment in the shallow feature space and capturing fine-grained semantic correlations in the implicit concept space. Furthermore, we design a Joint Temporal–Frequency Filtering (JTFF) module. In the frequency domain, features are processed via an MoE adapter and FFT-based filtering, while in the temporal domain, a temporal convolutional network (TCN) is employed to model temporal patterns in user–item interaction sequences. The resulting temporal and frequency-domain representations are integrated through Temporal–Frequency Refinement Gating (TFRG), enabling complementary refinement of node semantic representations and user behavior sequences. Extensive experiments conducted on four real-world datasets demonstrate that SFC-CDR outperforms state-of-the-art baselines.

Keywords: cross-domain recommendation; interaction graph; semantic alignment; contrastive learning



Academic Editor: Jérôme Gensel

Received: 20 September 2026

Revised: 11 October 2026

Accepted: 16 October 2026

Published: 18 October 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The core objective of Cross-Domain Recommendation (CDR) is to address data sparsity and cold-start problems in multi-source heterogeneous data environments. Traditional CDR models [1–4] often rely on partially overlapping users or items to establish connections between source and target domains, thereby transferring preference-related knowledge to alleviate data

sparsity in the target domain. From a graph perspective, user–item interactions can be naturally represented as an interaction graph, where users and items serve as nodes and observed interactions form edges. However, in many real-world scenarios, the correspondences between cross-domain users or items are usually unknown. Identifying such correspondences not only poses considerable technical challenges but also raises privacy concerns, limiting the practical applicability of existing CDR models. Therefore, developing CDR techniques that do not rely on overlapping users or items has become an important research direction. Early studies mainly focused on leveraging partially overlapping users or items to construct models, which were then generalized to non-overlapping users and items [5–8]. Although these CDR methods can be applied to non-overlapping users and items, their training processes still depend on cross-domain overlaps. To overcome this limitation, recent studies have explored training methods that eliminate such dependencies [9,10]. In such settings, the interaction graphs of different domains are naturally constructed from non-overlapping users and items, making it difficult to establish explicit structural correspondence across domains. Therefore, the key challenge lies in learning transferable representations from graph-structured interactions and establishing effective cross-domain associations without relying on overlapping users or items.

Although these CDR methods are effective, they still face the following limitations in non-overlapping scenarios: (1) Bias in cross-domain graph semantic alignment. Since users and items are non-overlapping, most approaches adopt unsupervised strategies to align semantic representations across domains. While such methods can reduce the distribution discrepancy between semantic spaces, they often overlook the deep semantic associations among graph nodes, leading to alignment bias across domains. For example, in the movie domain, a user who prefers fantasy films may be incorrectly assigned to a semantically mismatched cluster in the book domain due to the alignment bias introduced by unsupervised methods, as illustrated in Figure 1. Such deviation makes it difficult to identify transferable semantic correlations between nodes across interaction graphs from different domains. (2) Multi-granularity denoising of graph node preferences. Eliminating irrelevant or harmful associations in node feature representations and interaction relations can effectively improve representation quality. However, existing methods primarily focus on noise in feature representations and interaction relations separately, while ignoring implicit temporal noise embedded in semantic information. Furthermore, they fail to jointly consider textual node attributes and temporal interaction signals for hybrid denoising. As a result, node representations and interaction sequences may remain vulnerable to temporal disturbances or incomplete noise removal, ultimately degrading cross-domain knowledge transfer and recommendation performance.

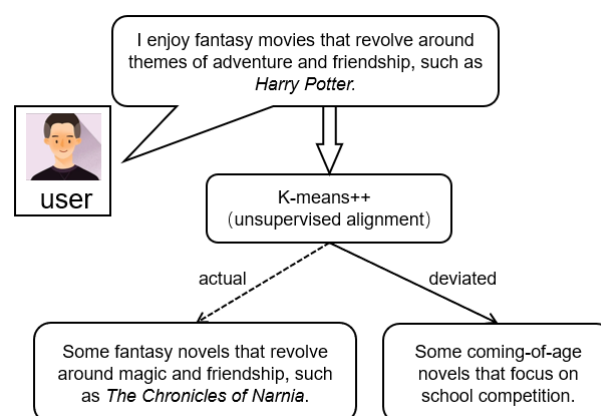


Figure 1. Examples of semantic alignment deviation in cross-domain scenarios involving non-overlapping users and items. Solid/dashed arrows indicate users’ deviated/actual alignment.

To address these challenges, we propose a Semantic Federated Clustering-driven Cross-Domain Recommendation model (SFC-CDR). First, we introduce a Gated Fusion Network (GFN), which leverages a gating mechanism to integrate semantic encoding features, ID embedding features, and nonlinear projection. By generating dynamic layer weights, GFN enables adaptive fusion of heterogeneous node information and constructs unified node semantic representations for subsequent cross-domain semantic alignment. To mitigate semantic alignment bias across interaction graphs from different domains, we design a Clustering-Anchored Contrastive Learning (CACL) module. Through a clustering-guided contrastive learning mechanism, CACL dynamically constructs cross-domain positive and negative sample pairs based on semantic anchor clusters and optimizes node semantic representations according to their semantic relevance. The semantic anchor clusters provide coarse-grained correspondence across domains, while contrastive learning further enables fine-grained semantic alignment. This process exploits deep semantic associations and facilitates transferable semantic knowledge across interaction graphs from different domains. To achieve joint denoising and refinement of node semantic representations and temporal interaction signals, we propose a Joint Temporal–Frequency Filtering (JTFF) framework. In the frequency domain, the cluster-enhanced node semantic representations are fed into a MoE adapter [11] to generate a cluster-aware embedding table, followed by an FFT-based filtering layer to suppress high-frequency noise. In the temporal domain, an ID embedding table is constructed via a standard lookup operation, and temporal patterns in user–item interaction sequences are captured using a Temporal Convolutional Network (TCN). The resulting temporal features are then coupled with the frequency-filtered node semantic representations through a Temporal–Frequency Refinement Gating (TFRG) mechanism, enabling complementary feature refinement and dynamic feature calibration. Through this joint process, heterogeneous noise is suppressed, yielding refined node representations for recommendation.

In summary, our proposed model makes the following contributions:

- We formulate non-overlapping cross-domain recommendation from an interaction-graph perspective, where each domain is modeled as a text-attributed user–item interaction graph. We then employ a Gated Fusion Network (GFN) to integrate semantic encoding features, ID embedding features, and nonlinear projection with dynamically generated layer weights, thereby constructing adaptive node semantic representations.
- We propose Clustering-Anchored Contrastive Learning (CACL), which employs semantic anchor clusters to establish coarse-grained correspondence and dynamically constructs positive and negative sample pairs for fine-grained semantic alignment across interaction graphs from different domains, without relying on explicit user or item correspondence.
- We design Joint Temporal–Frequency Filtering (JTFF), which jointly processes frequency-domain node semantic representations and temporal interaction signals. By combining FFT-based semantic filtering with TCN-based temporal modeling through TFRG, JTFF enables complementary feature refinement and suppresses heterogeneous noise in the representations used for recommendation.
- We conduct extensive experiments on four real-world datasets from the Amazon and OnlineRetail domains, and the results demonstrate the effectiveness of SFC-CDR compared with recent state-of-the-art methods.

2. Related Work

2.1. Cross-Domain Recommendation

Cross-domain recommendation (CDR) systems [12–15] are primarily designed to address data sparsity and cold-start problems in the target domain by enabling knowledge

sharing and collaborative optimization across different recommendation domains. From an interaction perspective, user–item interactions can be naturally represented as graph-structured data, where users and items constitute nodes and observed interactions form edges. Based on the availability of cross-domain entity correspondence, existing CDR methods can be broadly categorized into approaches based on partially overlapping users or items and approaches targeting non-overlapping users and items.

A. Partially overlapping users or items

Partially overlapping users or items refer to scenarios where different domains share a subset of common user or item entities, while the overlap is not complete. Such shared entities can serve as bridges for cross-domain knowledge transfer through federated learning [16–19], representation learning [20–23], and meta-learning [24–26]. For example, Yan et al. [27] enhanced privacy protection by adding Laplacian noise to gradients through local differential privacy (LDP) before uploading them while introducing a server-side embedding transformation mechanism. This approach leverages bridging-user embeddings to learn a cross-domain transformation model that addresses the cold-start problem. Wang et al. [28] enhanced the semantic understanding of user behaviors by incorporating encyclopedia-based knowledge graphs (KGs) and storing external KG information in a knowledge-enhanced memory network (KV-MN). Zhu et al. [29] introduced a multi-granularity transfer bridging mechanism to capture user preference features, injecting target-domain signals into the transfer process via target bridges, and further designed a target-aware contrastive learning layer to optimize user representation learning. However, matching information between partially overlapping users or items remains a key challenge in CDR. When the source and target domains belong to different platforms, identifying and exploiting shared entities may also introduce privacy risks, limiting the applicability of such approaches in practical cross-domain recommendation scenarios.

B. Non-overlapping users and items

Non-overlapping users and items indicate that no explicit entity correspondence exists between the source and target domains. Under this setting, CDR methods can generally be categorized into rating-pattern-based and content-based approaches. Rating-pattern-based approaches [30–33] primarily aim to uncover and transfer users' relatively stable rating behaviors or preference tendencies across domains. For instance, Zhang et al. [34] leverage auxiliary semantic signals to capture users' fine-grained rating patterns, using a self-attention mechanism to mitigate noise and subsequently infer target-domain preferences through a group-level preference fusion and reasoning framework. Similarly, Maneechote et al. [35] evaluate cross-domain similarities by comparing users' historical rating patterns across domains, enabling the dynamic adjustment of knowledge transfer strength.

Content-based approaches [36–40] focus on the intrinsic attributes and characteristics of items, seeking to establish cross-domain bridges by exploiting content-level similarities or correlations between items across domains. For example, Huai et al. [41] construct a heterogeneous information network to model semantic associations among users, items, and tags, while employing meta-paths to capture semantic information from different domains. They further adopt a hierarchical aggregation mechanism to filter irrelevant tags, thereby enhancing recommendation performance. Wang et al. [42], on the other hand, leverage large language models to generate domain-invariant and domain-specific textual representations, and align these representations between semantic and ID spaces through a contrastive learning strategy.

Although these methods can alleviate the lack of explicit entity correspondence, they may still suffer from transfer bias when knowledge is migrated between domains without overlapping users or items. In particular, the interaction structures of different domains can be naturally represented as interaction graphs, where directly establishing node-level

correspondence across domains is infeasible. Therefore, effective cross-domain transfer requires learning transferable semantic representations without relying on explicit shared entities. In this work, we address this challenge by dynamically generating weights from ID embedding features to construct layer-wise weighted semantic representations, followed by weighted clustering and dynamic contrastive learning to alleviate semantic transfer bias across interaction graphs from different domains.

2.2. Graph-Based Recommendation

Graph-based recommendation [43–46] has been widely studied to model relationships among users, items, and auxiliary entities. In particular, user–item interactions can be naturally represented as a bipartite graph, while heterogeneous information networks and knowledge graphs can further incorporate semantic and relational information. For example, Yu et al. [47] construct a cross-domain item–relation graph and learn item representations by means of graph neural networks. They further employ a bridge-transfer module to map source-domain user representations into the target-domain feature space, facilitating cross-domain knowledge transfer. Huang et al. [48] incorporate knowledge graphs as auxiliary information and adopt a pretraining–prompt–fine-tuning paradigm for graph prompt tuning. They design soft graph prompts and personalized graph prompts to efficiently exploit source-domain knowledge and boost cross-domain recommendation performance.

However, the existing graph-based recommendation methods typically leverage interaction structures to learn informative node representations, while cross-domain knowledge transfer remains challenging when the users and items are completely non-overlapping. In this setting, interaction data from different domains can be represented as interaction graphs constructed from non-overlapping users and items, where the absence of shared entities makes it difficult to establish transferable semantic relationships across domains. To address this challenge, our work adopts an interaction-graph perspective and focuses on semantic alignment and transfer through node representations.

2.3. Semantic Alignment

In cross-domain recommendation systems, semantic alignment [49–51] aims to reduce semantic discrepancies between heterogeneous domains and establish a shared representation space for effective knowledge transfer. In the context of graph-structured recommendation, semantic alignment can be viewed as establishing meaningful correspondences between node representations across interaction graphs from different domains, particularly when explicit user or item correspondence is unavailable. Among existing semantic alignment techniques, contrastive learning-based methods [52–54] construct cross-domain positive and negative sample pairs and employ contrastive objectives to learn a unified embedding space. Ni et al. [55] regarded cross-domain information as positive samples and marginally perturbed cross-domain information as negative samples, and employed a binary discriminator to maximize the mutual information between single-domain and cross-domain representations, thereby capturing users' correlated preferences across domains. Chang et al. [56] proposed a disentangled paradigm that separates inter-domain and intra-domain contrastive learning, and further introduced a stop-gradient operation to constrain excessive gradient propagation from the source domain to the target domain.

However, when users and items are non-overlapping, constructing reliable cross-domain positive and negative pairs remains challenging because explicit entity correspondence is unavailable. Existing contrastive learning-based methods may therefore rely heavily on distribution-level or manually constructed semantic associations, potentially introducing alignment bias and overlooking fine-grained semantic relationships among graph nodes. To address this limitation, our approach integrates weighted clustering with dynamic contrastive learning.

Specifically, clustering centers serve as semantic anchors to dynamically construct cross-domain homogeneous and heterogeneous pairs, thereby facilitating fine-grained semantic alignment between node representations across interaction graphs from different domains.

3. Methodology

In this section, we first clarify the symbol definitions in the text, and then outline the overall framework of the model.

3.1. Symbol Definition

In the setting, where users and items are non-overlapping across domain A and domain B , each domain can be represented as a text-attributed user-item interaction graph. For domain A , the graph is denoted as $\mathcal{G}^A = (\mathcal{V}^A, \mathcal{E}^A)$, where $\mathcal{V}^A = U^A \cup A$ denotes the node set consisting of user nodes and item nodes, and $\mathcal{E}^A$ denotes the observed user-item interaction edges. Similarly, domain B is represented as $\mathcal{G}^B = (\mathcal{V}^B, \mathcal{E}^B)$, where $\mathcal{V}^B = U^B \cup B$. Since users and items are non-overlapping across the two domains, the two interaction graphs have disjoint node sets, i.e., $\mathcal{V}^A \cap \mathcal{V}^B = \emptyset$.

The user sets of the two domains are represented as $U^A = \{u_1^A, u_2^A, \dots, u_i^A, \dots, u_m^A\}$ and $U^B = \{u_1^B, u_2^B, \dots, u_i^B, \dots, u_n^B\}$, where u_i^A and u_i^B denote the i -th users in domain A and domain B , respectively, and m and n represent the total numbers of users in the two domains. The item sets are represented as $A = \{a_1, a_2, \dots, a_j, \dots, a_k\}$ and $B = \{b_1, b_2, \dots, b_j, \dots, b_l\}$, where a_j and b_j denote the j -th items in domain A and domain B , respectively, and k and l represent the total numbers of items in the two domains. Each item is associated with an independent item ID and its corresponding textual description. From the graph perspective, users and items correspond to user nodes and item nodes, respectively, while each observed user-item interaction corresponds to an interaction edge. Taking domain A as an example, the textual description of item a_j is represented as $W_j^A = \{w_1, w_2, \dots, w_r, \dots, w_c\}$, where w_r denotes the r -th word in the item description and c denotes the number of words in the description. The textual description can be regarded as the textual attribute information of item node a_j . The interaction sequence of user $u_i^A \in U^A$ is denoted as $X_i^A = \{a_1, a_2, \dots, a_q, \dots, a_x\}$, where x represents the length of the interaction sequence X_i^A . From the graph perspective, X_i^A records the temporally ordered item interactions associated with user node u_i^A and reflects the temporal information of the corresponding interaction edges.

The multi-layer semantic encoding of item text extracted by the pretrained language model (PLM) is denoted as S^A , which can be viewed as the initial semantic representation of item nodes from the graph perspective. The layer-weighted semantic encoding obtained by weighting is denoted as S_w^A , representing the weighted semantic representations of item nodes. The cluster encoding obtained through cross-domain semantic clustering alignment is denoted as S_c^A , which can be viewed as the cluster-enhanced semantic representation of item nodes. We use text information and the user's interaction sequence to predict the probability of the user's next possible interaction item, which is denoted as $P(a_{x+1} | X_i^A)$. Table 1 summarizes the symbols used in this paper and their explanations.

Table 1. Symbol definition.

Symbol	Definition and Notation
U^A	user set of domain A
U^B	user set of domain B
u_i^A	user i from domain A, $u_i^A \in U^A$
u_i^B	user i from domain B, $u_i^B \in U^B$

Table 1. Cont.

Symbol	Definition and Notation
A	item set of domain A
B	item set of domain B
W_i^A	the description content of item a_i of domain A
X_i^A	interaction sequence of user u_i^A
S^A	multi-layer semantic encoding of domain A
S_w^A	layer-weighted semantic encoding of domain A
S_c^A	Cross-domain semantic embeddings of A
E^A	ID embedding table of domain A
E^B	ID embedding table of domain B
T^A	Cross-domain clustered embeddings of A

3.2. Overview of SFC-CDR

The proposed SFC-CDR consists of three modules: semantic extraction, cross-domain clustering-anchored contrastive learning, and temporal–frequency filtering, as illustrated in Figure 2. Each domain is modeled as a text-attributed user–item interaction graph, where users and items are graph nodes, user–item interactions form graph edges, and item texts provide semantic attributes for item nodes. The semantic extraction module employs a pre-trained language model (PLM) to obtain multi-layer semantic representations of item nodes in the source and target domains. These multi-layer representations are then used as the semantic representation basis in the Gated Fusion Network (GFN), where ID embedding features and nonlinear projection are introduced to generate dynamic layer weights. These weights are then applied to the corresponding semantic layers to obtain layer-weighted node semantic representations. The resulting representations from the source and target domains are uploaded to the server for subsequent cross-domain semantic alignment.

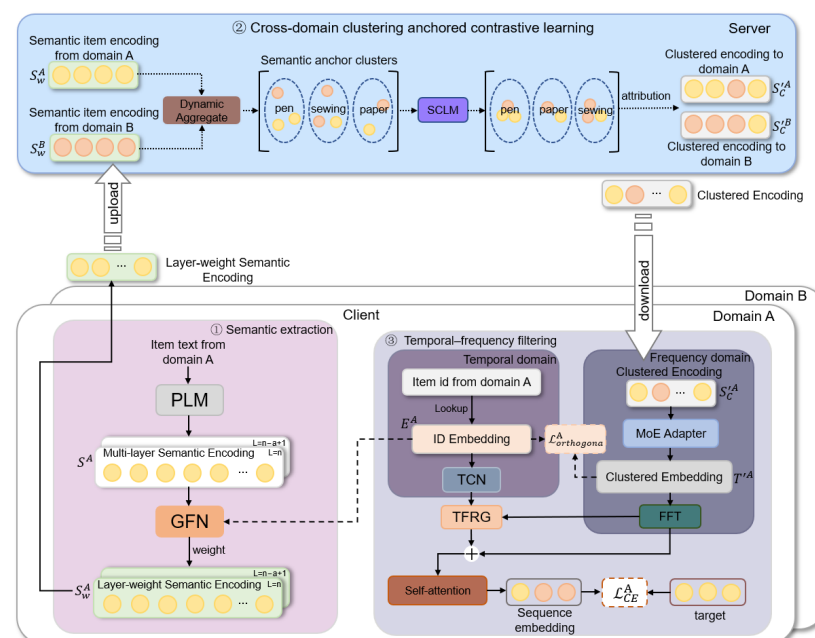


Figure 2. The overall architecture of SFC-CDR consists of three modules: (1) Semantic extraction mainly extracts semantic codes from project description texts. (2) The key of cross-domain clustering anchor contrastive learning lies in dynamically clustering semantics from different domains and optimizing clustering results through the Semantic Contrastive Learning Mechanism (SCLM). (3) Temporal–frequency filtering is designed to suppress implicit temporal noise in text information and ID embeddings.

In the cross-domain clustering-anchored contrastive learning module, the uploaded layer-weighted node semantic representations are dynamically clustered to construct semantic anchor clusters. These clusters provide coarse-grained semantic correspondence across interaction graphs from different domains without requiring explicit shared-node correspondence. The semantic anchors are subsequently used to dynamically construct positive and negative sample pairs through the Semantic Contrastive Learning Mechanism (SCLM), enabling fine-grained alignment based on the semantic relationships between node representations. After contrastive alignment, the resulting cluster-enhanced node semantic representations are returned to the clients for subsequent feature refinement and interaction-aware representation learning.

After receiving the cluster-enhanced representations, each client performs temporal-frequency filtering through two complementary branches: frequency-domain filtering and temporal-domain modeling. In the frequency domain, the cluster-enhanced representations are fed into the MoE adapter to generate a cluster-aware embedding table, followed by an FFT-based filtering layer to suppress high-frequency noise. In the temporal domain, an ID embedding table is obtained through a standard lookup operation based on item IDs, and the temporal patterns of user-item interaction sequences are modeled using a Temporal Convolutional Network (TCN). The Temporal-Frequency Refinement Gating (TFRG) module then couples the frequency-filtered node semantic representations with the temporally modeled interaction features, enabling noise suppression and complementary feature refinement for the final node representations used in recommendation.

Graph properties and their utilization. The constructed interaction graphs are user-item interaction graphs with text-attributed item nodes and ordered user interaction sequences. Unlike conventional graph learning methods that rely on homophily assumptions and neighborhood aggregation, SFC-CDR does not explicitly assume that adjacent nodes are homophilic. Instead, it exploits semantic relationships among item nodes by grouping semantically related items into semantic anchor clusters, which provide coarse-grained semantic correspondence across interaction graphs from different domains. This graph property is directly utilized by the proposed modules: GFN combines PLM-based semantic features with ID embeddings to obtain node semantic representations; SCLM uses semantic anchor clusters to construct semantically related positive pairs and unrelated or semantically confusing negative pairs, thereby enhancing cross-domain semantic alignment and node-level semantic discrimination. The resulting cluster-enhanced semantic representations are further refined by JTFF together with temporal interaction signals to capture temporal interaction patterns.

3.3. Semantic Extraction

In this section, we mainly construct a hierarchical semantic representation optimization framework through multi-layer PLM and GFN, aiming to adaptively integrate multi-granularity semantic features of text and provide enhanced semantic representations with both global consistency and local discriminability for downstream tasks. From the interaction-graph perspective, the extracted semantic features are regarded as node semantic representations associated with the corresponding item nodes, providing semantic information for subsequent cross-domain semantic alignment. To protect privacy, we only upload the layer-weighted semantic encoding to the server.

To extract deep semantic information from project texts, we adopt a multi-layer PLM to effectively preserve and enhance the core semantic information of the input text. Specifically, taking domain A as an example, we split the project text of domain A into individual words $W_i^A = \{w_1, w_2, \dots, w_c\}$ (where the number of words is c), then add the token [CLS] at the

beginning, and input it into the BERT [57] of the multi-layer PLM, thereby obtaining multi-layer semantic features for the corresponding item node in the interaction graph,

$$S_i^{A,L=j} = \text{BERT}^{L=j}([\text{CLS}]; w_1; \dots; w_c) \quad (1)$$

where “;” denotes the concatenation operation, $S_i^{A,L=j} \in \mathbb{R}^D$ represents the output feature of the i -th sample at the j -th layer in BERT, and D is the vector dimension. Here, $S_i^{A,L=j}$ can be viewed as the layer-specific semantic feature of the corresponding item node. We select the last a layers as its initial multi-layer semantic encoding, denoted as:

$$S_i^{A,L} = \{S_i^{A,L=n-a+1}, \dots, S_i^{A,L=n}\} \quad (2)$$

where n is the total number of layers of BERT, a is a hyperparameter, and the relevant evaluation is carried out in Section 4.4.

To enable the distinguishability of important information between similar semantic encodings and more accurate clustering across domains, we obtain adaptive filtering and weighted fusion of multi-layer semantic encodings through GFN. GFN is essentially an adaptive feature filter bank that dynamically adjusts the contribution of each layer of semantics through end-to-end learning. For each item node, the original multi-layer semantic encoding produced by the PLM is first projected into a unified d -dimensional hidden space through a linear transformation:

$$M_i^{A,L=j} = W_s S_i^{A,L=j} + b_s \quad (3)$$

where $S_i^{A,L=j} \in \mathbb{R}^D$ denotes the original semantic encoding of the i -th item node extracted from the j -th PLM layer, $W_s \in \mathbb{R}^{d \times D}$ is the projection matrix, and $b_s \in \mathbb{R}^d$ is the bias vector. Thus, $M_i^{A,L=j} \in \mathbb{R}^d$. For each item node, the multi-layer semantic encoding and ID embedding are then used to generate weights via the sigmoid function:

$$w_{i,j} = \text{Sigmoid}(W_g \cdot [M_i^{A,L=j} \oplus e_i] + b_g) \quad (4)$$

where $W_g \in \mathbb{R}^{d \times 2d}$ is the weight matrix of GFN, and $2d$ corresponds to the dimension of the concatenated d -dimensional semantic encoding and d -dimensional ID embedding. $b_g \in \mathbb{R}^d$ represents the bias vector of GFN. Accordingly, the resulting gate vector $w_{i,j} \in \mathbb{R}^d$. Then, the obtained weights are used to perform a weighted combination of the semantic encodings of each layer:

$$s_{w,i}^A = \sum_{j=n-a+1}^n w_{i,j} \cdot M_i^{A,L=j} \quad (5)$$

$s_{w,i}^A \in \mathbb{R}^d$ representing the layer semantic code of the i -th item node. Finally, the layer-weighted semantic coding representation is obtained as:

$$S_w^A = \{s_{w,1}^A, s_{w,2}^A, \dots, s_{w,i}^A, \dots, s_{w,k}^A\} \quad (6)$$

The resulting layer-weighted node semantic representations S_w^A are uploaded to the server to support the subsequent cross-domain clustering-anchored contrastive learning process.

3.4. Cross-Domain Clustering Anchor Contrastive Learning

In this part, the alignment training for non-overlapping users and items is mainly solved. In order to align and cluster domain A and domain B under the setting of non-overlapping users and items, through the centralized collaborative optimization of the

server, the leap from the feature association at the data layer to the concept mapping at the cognitive layer of the heterogeneous domain project semantics is realized. From the interaction-graph perspective, this process aims to establish semantic correspondence between item nodes across interaction graphs from different domains without requiring explicit node-level correspondence.

The K-means++ algorithm improves clustering stability through probabilistic initial center selection, but in high-dimensional sparse data, there still exists a deviation between the clustering structure and the real semantic topology. Therefore, we adopt an optimized k-means++ algorithm [10]. Specifically, we acquire the layer-weighted semantic encodings $S_{w,i}^A$ and $S_{w,i}^B$ of domains A and B from the client, which represent the semantic features of item nodes in the corresponding interaction graphs, and the server randomly selects cluster centers $C = \{c_1, c_2, \dots, c_k\}$ from the encodings with a predefined number of cluster centers (where k is the number of clusters). The clustering is determined by the distance from each point to the cluster center; the shorter the distance, the higher the contribution degree of the point to the cluster center, which is expressed as follows:

$$\ell_i = \arg \min_j \|s_{w,i} - c_j^{(t)}\|^2 \quad (7)$$

where $c_j^{(t)}$ denotes the centroid of the j -th cluster at the t -th iteration, $\|\cdot\|^2$ represents the squared Euclidean distance, and $s_{w,i}$ denotes the layer-weighted semantic representation derived from either domain A or domain B. The update of the next cluster center is carried out in a weighted aggregation manner:

$$c_j^{(t+1)} = \frac{\sum_{i:\ell_i^{(t)}=j} w_i^{(t)} \cdot s_{w,i}}{\sum_{i:\ell_i^{(t)}=j} w_i^{(t)}} \quad (8)$$

$$w_i^{(t)} = \frac{1}{\|s_{w,i} - c_j^{(t)}\| + \varepsilon} \quad (9)$$

where $c_j^{(t+1)}$ denotes the centroid vector of the j -th cluster at the $(t+1)$ -th iteration, ε is a small constant to prevent the denominator from being zero, w_i is the weight of the cluster code $s_{w,i}$, and the closer the distance, the greater the weight. Finally, the semantic anchor clusters $\hat{s} = \{r_1, r_2, \dots, r_x\}$ are obtained through clustering and serve as coarse-grained semantic anchors for the item nodes across the two interaction graphs.

Semantic anchor clusters achieve initial alignment across non-overlapping users and items but do not explicitly model the deep semantic correspondence between domain-specific semantic encodings, resulting in limited alignment accuracy across semantic spaces and insufficient association strength between similar concepts. Therefore, we design a Semantic Contrastive Learning Mechanism (SCLM). For each semantic anchor cluster r_j , the item-node representations belonging to the same cluster are separately aggregated within domains A and B to obtain domain-specific semantic anchor encodings r_j^A and r_j^B . The two encodings corresponding to the same semantic anchor cluster are treated as a semantically corresponding positive pair, while encodings from different semantic anchor clusters form negative pairs. Two strategies are adopted to construct negative samples: first, randomly selecting cross-domain samples from different clusters; second, screening semantically confusing sample pairs according to the inter-domain semantic-distance ranking, thereby forcing the model to distinguish closely related but heterogeneous concepts. Using the weighted aggregated cluster centers as semantic anchors, the positive and negative sample

pools are dynamically updated during contrastive learning. The contrastive objective is defined as:

$$\mathcal{L}_{\text{SCLM}} = -\frac{1}{|\mathcal{R}|} \sum_{j=1}^{|\mathcal{R}|} \log \frac{\exp(\text{sim}(r_j^A, r_j^B)/\tau)}{\exp(\text{sim}(r_j^A, r_j^B)/\tau) + \sum_{q \in \mathcal{N}_j} \exp(\text{sim}(r_j^A, q)/\tau)} \quad (10)$$

where $\mathcal{R}$ denotes the set of semantic anchor clusters, $\mathcal{N}_j$ denotes the dynamically constructed negative sample pool for the j -th anchor, $\text{sim}(\cdot, \cdot)$ denotes cosine similarity, and τ is the temperature coefficient. The objective pulls together the domain-specific encodings of the same semantic anchor cluster while pushing apart encodings belonging to different or semantically confusing clusters, thereby achieving fine-grained semantic alignment between the item nodes of the two interaction graphs. The obtained cluster codes are assigned back to domain A and domain B, denoted as $S_c'^A$ and $S_c'^B$, respectively. Then, $S_c'^A$ and $S_c'^B$ are sent back to the client.

3.5. Temporal–Frequency Filtering

The cluster coding received from the server side is an aggregated result of multi-domain features, which may carry noise and pseudo-semantics. From the interaction-graph perspective, the cluster coding provides refined semantic representations for the corresponding item nodes after cross-domain semantic alignment. Direct training would reduce the task relevance of sequence embeddings. Additionally, there exists implicit temporal noise in semantic information, causing sequence embeddings to be susceptible to sudden noise interference or incomplete noise filtering, which exerts a significant impact on the results. Therefore, we adopt Joint Temporal–Frequency Filtering (JTFF).

In the frequency domain, the cluster encoding is passed through the MoE adapter. The MoE adapter includes multiple expert modules and a routing mechanism (GR). Each expert module is essentially a parameter whitening transformation module (PW). The cluster encoding is passed through the expert modules:

$$\tilde{s}_i'^A = (\text{Dropout}(s_i'^A) - b^{\text{PW}}) \cdot W^{\text{PW}} \quad (11)$$

where $s_i'^A$ is the i -th cluster code, $W^{\text{PW}} \in \mathbb{R}^{d \times d}$ is the weight matrix of PW, and $b^{\text{PW}} \in \mathbb{R}^d$ is the bias term of PW. Using this method, semantic dimension features can be captured. These features correspond to the semantic representations of the aligned item nodes in the interaction graph. The routing mechanism can dynamically determine the contribution of each expert module according to the original semantics of the cluster code:

$$g_k = \text{Softmax}(s_i'^A \cdot W^{\text{MoE}} + \delta) \quad (12)$$

where $W^{\text{MoE}} \in \mathbb{R}^{d \times G}$ is the weight matrix that each expert of GR can learn, G is the number of experts, and δ is the Gaussian random noise for balancing expert load. The cluster embedding is obtained through the combination of the expert module and the routing mechanism:

$$t_i'^A = \sum_{k=1}^G g_k \cdot \tilde{s}_{i,k}'^A \quad (13)$$

where $\tilde{s}_{i,k}'^A$ is the output of the k -th PW layer, $t_i'^A \in \mathbb{R}^d$ is the i -th cluster embedding. We apply the MoE adapter to each cluster encoding to get each cluster embedding, which can be expressed as:

$$T'^A = \{t_1'^A, t_2'^A, \dots, t_i'^A, \dots, t_x'^A\} \quad (14)$$

To achieve spectrum-level refined regulation of the project's text interaction features, we adopt a learnable frequency domain filter layer based on FFT. Its core lies in using a differentiable frequency domain weight matrix to selectively enhance or suppress different frequency components of the features, thereby realizing adaptive refined processing for noise components in the spectrum space. First, we map the cluster embedding table from temporal interaction features to the frequency domain:

$$Y_i'^A = \mathcal{F}(T_i'^A) \in \mathbb{C}^{x \times d} \quad (15)$$

where x denotes the number of cluster embeddings in the cluster embedding table, d denotes the embedding dimension, $\mathcal{F}(\cdot)$ is the FFT that converts the cluster embedding to the frequency domain space, and $\mathbb{C}$ is the complex number field. Then, the feature $Y_i'^A$ of each dimension is processed by a learnable filtering parameter W_c to achieve adaptive filtering:

$$\tilde{Y}_i'^A = W_c \odot Y_i'^A \quad (16)$$

where $\odot$ is element-wise multiplication. Finally, convert the filtered frequency domain feature $\tilde{Y}_i'^A$ back to the original domain feature $T_i'^A$:

$$T_i'^A \leftarrow \mathcal{F}^{-1}(\tilde{Y}_i'^A) \in \mathbb{R}^{x \times d} \quad (17)$$

where $\mathcal{F}^{-1}(\cdot)$ is the inverse FFT, which functions to convert the filtered frequency-domain signal back to the original domain. The filtered cluster embeddings are:

$$\tilde{T}^A = \{T_1'^A, T_2'^A, \dots, T_i'^A, \dots, T_x'^A\} \quad (18)$$

In the temporal domain, we perform a lookup operation through a learnable embedding matrix to obtain ID embeddings $E^A = \{e_1^A, e_2^A, \dots, e_i^A, \dots, e_x^A\}$ from the item texts in domain A and then capture the temporal dependencies of user interactions using TCN with its causal convolution characteristics:

$$h_i^{L=j} = \text{ReLU}\left(\sum_{k=0}^{K-1} W_k^{L=j} \cdot h_{i-k}^{L=j-1} + b^{L=j}\right) \quad (19)$$

where $h_i^{L=0} = e_i^A$, where $h_i^{L=j} \in \mathbb{R}^d$ represents the temporal feature extracted at temporal position i in the j -th convolutional layer. K is the convolution kernel size, $W_k^{L=j} \in \mathbb{R}^{d \times d}$ is the weight matrix of the k -th convolution kernel in the j -th layer, and $b^{L=j} \in \mathbb{R}^d$ is the bias term of the TCN.

Then, the text features processed in the frequency domain and the user behavior features modeled in the temporal domain are fused through the Temporal-Frequency Refinement Gating (TFRG):

$$g_i = \text{Sigmoid}\left(W_e \cdot h_i^{L=j} + W_t \cdot T_i'^A + b_{eg}\right) \quad (20)$$

where $W_e \in \mathbb{R}^{d \times d}$ is the linear transformation weight matrix for the temporal feature $h_i^{L=j}$, and $W_t \in \mathbb{R}^{d \times d}$ is the linear transformation weight matrix for the text feature $T_i'^A$. $b_{eg} \in \mathbb{R}^d$ is the bias term of TFRG. The ID feature after TFRG is:

$$\hat{e}_i^{A'} = g_i \odot h_i^{L=j} \quad (21)$$

The filtered ID embedding is:

$$\hat{E}^A = \{\hat{e}_1^{A'}, \hat{e}_2^{A'}, \dots, \hat{e}_i^{A'}, \dots, \hat{e}_x^{A'}\} \quad (22)$$

Finally, we add the i -th cluster embedding $T_i'^A$ after temporal–frequency filtering to the ID embedding $\hat{e}_i'^A$ to obtain the fused embedding vector $v_i'^A$ of multiple source features:

$$v_i'^A = \hat{e}_i'^A + T_i'^A \quad (23)$$

All sets of fused embedded vectors are:

$$\tilde{V}'^A = [v_0'^A, v_1'^A, \dots, v_i'^A, \dots, v_x'^A] \quad (24)$$

Then, $v_i'^A$ is passed through a commonly used transformer structure to dynamically capture key dependencies within the sequence, providing high-quality user interest representations for the final recommendation prediction:

$$V'^{A,l} = \text{LayerNorm}(\text{MHAtt}(\tilde{V}'^A)), \quad l \in \{1, 2, \dots, L\} \quad (25)$$

where MHAtt is the multi-head attention mechanism, $V'^{A,l}$ is the sequence embedding of the l -th Transformer layer, and there are a total of L layers. We select $V'^{A,L}$ as the final sequence embedding, that is $V'^{A,i} = V'^{A,L}$.

3.6. Model Optimization

We use the user's interaction sequence in domain A, together with item representations, to predict the probability of the user's next interaction. The item representations are obtained by fusing the deep semantic encodings of item description texts with ID embeddings. The prediction probabilities are as follows:

$$P(a_{x+1} | X_i^A) = \text{Softmax}(V'^{A,i} \cdot E^{A^\top}) \quad (26)$$

where E^A is the ID embedding in domain A. To optimize the matching degree between the predicted probability of the model for the next interaction and the true label, we construct it using cross-entropy loss:

$$\mathcal{L}_{\text{CE}}^A = -\frac{1}{|X_A|} \sum_{X_i^A \in X_A} \log P(a_{i+1} | X_i^A) \quad (27)$$

where X_A represents the training set in domain A.

Considering that semantic embeddings and ID embeddings respectively capture the differentiated feature dimensions of items in the model, to enhance the diversity and specificity of the feature space, a forced orthogonality loss is introduced to constrain the two modalities:

$$\mathcal{L}_{\text{orthogonal}}^A = \|t_1'^{A^\top} e_1^A\|_2 + \dots + \|t_x'^{A^\top} e_x^A\|_2 \quad (28)$$

where $\|\cdot\|_2$ denotes the Euclidean norm of the vector. The total loss is the superposition of the cross-entropy loss and the orthogonality loss:

$$\mathcal{L}^A = \mathcal{L}_{\text{CE}}^A + \mathcal{L}_{\text{orthogonal}}^A \quad (29)$$

4. Experiment

In this chapter, to verify the effectiveness of the proposed SFC-CDR model, we conduct a large number of experiments on two pairs of domains in the Amazon public dataset and answer the following key research questions through a multi-dimensional evaluation framework.

- RQ1: Does our method have a significant performance improvement compared with the current state-of-the-art methods?
- RQ2: What are the impacts of different components such as clustering encoding, temporal-domain filtering, frequency-domain filtering, and hybrid filtering on our model?
- RQ3: How does the performance of SFC-CDR change with the transformation of hyperparameters?
- RQ4: What impact will using different coarse-grained alignment methods have on SFC-CDR?
- RQ5: How semantically consistent are the cross-domain items assigned to the same semantic cluster?

4.1. Experimental Setup

4.1.1. Dataset

To verify the cross-domain recommendation performance of the SFC-CDR model, we selected two groups of domain data pairs from different data sources for evaluation, which were constructed into the “OnlineRetail-Pantry” and “Office-Arts” datasets, respectively. From the interaction-graph perspective, each domain is represented as an interaction graph, where users and items correspond to nodes and user–item interactions form the edges. The data for the “Office”, “Arts”, and “Pantry” domains are all derived from the Amazon (<https://nijianmo.github.io/amazon/index.html>, accessed on 20 September 2026) e-commerce platform, whose dataset contains real user reviews and meta-information about products and is a commonly used benchmark dataset in recommendation system research. The “OnlineRetail” domain data comes from a UK (<https://www.kaggle.com/carrie1/ecommerce-data>, accessed on 20 September 2026) online retail enterprise, covering user purchase records and item attribute descriptions. In terms of domain division, the Office and Arts data from the Amazon platform are divided into one group of cross-domain experimental objects, and the OnlineRetail and Pantry data (belonging to different platforms) are divided into another group of comparative experimental objects. Under the non-overlapping setting, each domain is represented as an interaction graph, with users and items not shared across domains. Considering the temporal characteristics of the sequential recommendation task, all user interaction records are sorted in ascending order of timestamps. Under the non-overlapping setting, unique IDs are reassigned to users within each domain (starting from 0). Regarding the processing of item text features: for the Amazon dataset, the title, category, and brand fields of products are concatenated as descriptive text (following the standard preprocessing method in similar studies); for the OnlineRetail dataset, the original item description text is used directly. These item descriptions serve as textual attributes associated with the corresponding item nodes in the interaction graphs. To alleviate the problem of data sparsity, we filter out the user and the item with fewer than five interactions. The statistical information of the finally processed datasets is shown in Table 2.

Table 2. The statistics of our datasets, where “Avg.n” denotes the average length of the user interaction sequence.

Datasets	#Users	#Items	#Inters	Avg.n
OnlineRetail	16,521	3470	503,386	30.47
Pantry	13,102	4899	113,861	8.69
Office	87,347	25,987	597,491	6.84
Arts	45,487	21,020	349,664	7.69

4.1.2. Evaluation Index

In our experiments, we adopt the leave-one-out strategy to evaluate the next-item prediction performance on sequential interaction data. From the interaction-graph perspective, each user's interaction sequence corresponds to an ordered sequence of user-item edges in the domain-specific interaction graph. For each user in the target domain, the interaction sequence is partitioned such that the last interaction is used for testing, the second-to-last interaction for validation, and all remaining interactions for training. Model training is conducted based on both target items and historical interactions beyond the target items. During validation and testing, the final performance is computed by averaging over all users' interaction sequences with respect to the candidate items. The evaluation metrics include Recall@N and NDCG@N, where N is set to 10 and 50 in our experiments.

4.1.3. Baseline

To verify the effectiveness of the SFC-CDR model, we compare SFC-CDR with five other cross-domain recommendation methods.

- CCDD [58]: The proposed method employs intra-domain contrastive learning with subgraph augmentation to learn robust representations, while inter-domain tasks are designed from the perspectives of users, categories, and neighbors, and optimized jointly through a multi-task learning framework.
- RecGURU [59]: This method includes a general user representation unit (GURU) and a cross-domain sequence recommendation (CDSRes) unit. GURU generates user embeddings using a self-attention autoencoder and integrates them into a unified general user representation (GUR) through adversarial training. The CDSRec unit combines GUR with an attention mechanism to improve recommendation performance.
- PFCR [40]: This method adopts the federated learning architecture, solves the privacy problem through user interaction with the local client and gradient encryption, constructs a unified feature space by using item description texts, builds an inter-domain bridge by relying on natural language commonalities, and designs two prompt fine-tuning strategies to adapt to the target domain.
- FedDCSR [16]: This method uses inter-domain-intra-domain sequence representation decoupling (SRD) to decompose user sequence features into domain-shared and exclusive features, and designs an intra-domain contrastive information maximization (CIM) strategy to learn richer exclusive features through data augmentation.
- FFMSR [10]: This method obtains semantic encoding through a pre-trained model, server clustering promotes inter-domain semantic transfer, fuses ID and text modalities, and processes noise with a fast Fourier transform (FFT) filtering layer and a gating mechanism.

4.1.4. Parameter Setting

The implementation of the SFC-CDR model is based on Python 3.8.13 and the Pytorch 1.11.0 framework, and all experimental processes are completed in the hardware environment of NVIDIA RTX 3090 GPU. In terms of model structure parameters, the semantic extraction module uses a pre-trained language model (PLM) to generate original semantic vectors with a dimension of 768, and performs feature fusion by intercepting the output of its last 2 layers, while configuring the overall hidden layer dimension of the model to 256; these semantic representations are associated with the corresponding item nodes in the domain-specific interaction graphs. In the cross-domain semantic fusion module, considering the difference in data scale between the Office-Arts and OnlineRetail-Pantry datasets, the number of server-side clustering centroids is set to 120 to adapt to the data distribution characteristics of different domains; The semantic filtering module sets a 2-

layer filtering network for the sequence encoder, and only a 1-layer structure is used for clustering embedding filtering to balance the feature refinement effect and computational complexity. In terms of training strategy, the model optimization adopts the Adam algorithm (initial learning rate of 0.0005), and the batch sample size is uniformly set to 512; the joint pre-training stage iterates for 5 epochs, and an early stopping mechanism (with a patience value of 15) is introduced during the local client fine-tuning process after federated training, and finally the model with the optimal Recall@10 metric on the validation set is selected for evaluation in the testing stage.

To ensure a fair comparison, all methods are evaluated using the same dataset preprocessing, domain-pair construction, data splitting, and leave-one-out sequential evaluation protocol. The same evaluation metrics, including Recall@10/50 and NDCG@10/50, are applied to all methods. For each baseline, we follow the parameter settings and implementation details reported in the original paper or publicly available implementation, while keeping the common experimental conditions consistent across methods. The reported results are obtained under these consistent experimental settings, ensuring a fair and comparable evaluation of all methods.

4.2. Experimental Results (RQ1)

Experimental results compared with other baseline models are shown in Table 3. Our proposed model achieves the best performance on two metrics (Recall@N and NDCG@N), which verifies the usefulness of our method in achieving fine-grained semantic alignment across interaction graphs from different domains and filtering implicit temporal noise in graph node semantic representations. Specifically, we can draw the following observations:

Table 3. Comparison results on *OnlineRetail-Pantry* and *Office-Arts*. The optimal result is marked in bold, and the suboptimal result is underlined.

Dataset	Metric	CCDR	RecGURU	PFCR	FedDCSR	FFMSR	SFC-CDR (Ours)	Improv
OR	Recall@10	0.1345	0.1465	0.1558	0.1407	<u>0.1653</u>	0.1721	+4.11%
	NDCG@10	0.0655	0.0533	0.0737	0.0688	<u>0.0761</u>	0.0777	+2.10%
	Recall@50	0.3586	0.3884	0.398	0.3367	<u>0.413</u>	0.428	+3.63%
	NDCG@50	0.1106	0.1187	0.1265	0.1114	<u>0.1302</u>	0.1339	+2.84%
Pantry	Recall@10	0.0405	0.0467	0.0615	0.0292	<u>0.0698</u>	0.073	+4.58%
	NDCG@10	0.0201	0.0206	0.0291	0.0141	<u>0.0325</u>	0.0339	+4.31%
	Recall@50	0.126	0.1272	0.159	0.0875	<u>0.1816</u>	0.1871	+3.03%
	NDCG@50	0.0384	0.0381	0.0501	0.0265	<u>0.0566</u>	0.0587	+3.71%
Office	Recall@10	0.0546	0.1142	0.1203	0.1022	<u>0.131</u>	0.1337	+2.06%
	NDCG@10	0.0289	0.0766	0.078	0.0787	<u>0.0889</u>	0.0912	+2.59%
	Recall@50	0.1092	0.1755	0.1936	0.1525	<u>0.2071</u>	0.2129	+2.80%
	NDCG@50	0.0399	0.0899	0.094	0.0896	<u>0.1056</u>	0.1096	+3.79%
Arts	Recall@10	0.0669	0.1086	0.1151	0.0852	<u>0.1259</u>	0.1315	+4.45%
	NDCG@10	0.0345	0.0649	0.0676	0.0591	<u>0.0712</u>	0.0742	+4.21%
	Recall@50	0.1477	0.1977	0.2198	0.1505	<u>0.2375</u>	0.2474	+4.17%
	NDCG@50	0.0521	0.0846	0.0902	0.0733	<u>0.0953</u>	0.0995	+4.41%

The optimal and suboptimal results are marked in bold and underlined, respectively. Significant improvements over the best baseline results are shown in bold (t -test, $p < 0.05$).

- Our model SFC-CDR outperforms other baselines. It achieves a 5.04% improvement over the best baseline FFMSR in Recall@10 and a 4.36% improvement in NDCG@10, realizing the best performance. In addition, our method captures and filters temporal noise, eliminating the impact of noise more comprehensively on the semantic representations associated with graph nodes.

- CDR methods based on partially overlapping users or items, such as CCDR, RecGURU, FedDCSR, and PFCR, demonstrate that project-related gradients sent to the server can improve model accuracy. From the interaction-graph perspective, these methods mainly exploit partially shared entities to facilitate information transfer between domain-specific interaction graphs.
- FFMSR has been studied in scenarios with non-overlapping users and items. It achieves cross-domain graph semantic alignment through an optimized k-means++ algorithm and refines cross-domain post-training node semantic representations using FFT-based filtering layers. Although FFMSR outperforms the other baseline methods, its performance remains below that of SFC-CDR, highlighting the potential benefits of our model design. SFC-CDR achieves fine-grained cross-domain semantic alignment through a cluster-anchored contrastive learning mechanism and further reduces noise in node semantic representations using Joint Temporal–Frequency Filtering (JTFF).

4.3. Ablation Study (RQ2)

To verify the effectiveness of each core component of the SFC-CDR model, we set up several variants.

- w/o Semantic Contrastive Learning Mechanism (SCLM): In cross-domain semantic alignment, SCLM is removed, and the model only performs clustering alignment between the semantic representations of nodes across interaction graphs from different domains.
- w/o Temporal Convolutional Network (TCN): In temporal filtering, TCN is removed, and ID embedding does not extract the temporal dependencies of user interactions along the corresponding graph-structured interaction sequences.
- w/o frequency-domain filtering (FF): In frequency-domain filtering, the MoE adapter and FFT-based module are removed, and no filtering operation is performed on cluster encoding associated with the graph nodes.
- w/o Temporal–Frequency Refinement Gating (TFRG) filtering: In mixed filtering modalities, this filtering operation is removed, and the temporal dependencies extracted in the temporal domain are directly fused with the text features filtered in the frequency domain to form node representations associated with the corresponding interactions.

The results of the ablation study are presented in Table 4.

Table 4. Ablation analysis on *OnlineRetail-Pantry* and *Office-Arts*.

Variants	OR				Pantry				Office				Arts			
	Recall		NDCG		Recall		NDCG		Recall		NDCG		Recall		NDCG	
	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50
w/o SCLM	0.1709	0.4138	0.0757	0.1311	0.0715	0.1866	0.0329	0.0573	0.1310	0.2112	0.0906	0.1084	0.1303	0.2458	0.0728	0.0981
w/o TCN	0.1714	0.4185	0.0762	0.1325	0.0721	0.1859	0.0330	0.0573	0.1326	0.2119	0.0908	0.1088	0.1308	0.2466	0.0736	0.0986
w/o FF	0.1716	0.4215	0.0766	0.1334	0.0728	0.1867	0.0334	0.0582	0.1341	0.2121	0.0921	0.1089	0.1311	0.2464	0.0739	0.0989
w/o TFRG	0.1719	0.4204	0.0773	0.1326	0.0726	0.1864	0.0342	0.0579	0.1328	0.2124	0.0909	0.1091	0.1332	0.2480	0.0732	0.0988
SFC-CDR	0.1721	0.428	0.0777	0.1339	0.073	0.1871	0.0339	0.0587	0.1337	0.2129	0.0912	0.1096	0.1315	0.2474	0.0742	0.0995

Based on the above results, the constituent modules of the SFC-CDR model contribute to its overall recommendation performance.

- Without SCLM, “w/o SCLM” shows a performance decrease of 2.1% and 3.04% in terms of Recall@10 and NDCG@10 compared to SFC-CDR. This suggests that SCLM contributes to semantic alignment across interaction graphs from different domains.
- Pruning the module that captures the time series dependencies of user interactions (“w/o TCN”) leads to a decline in Recall@N and NDCG@N, suggesting that capturing

the dynamic behaviors of user interactions contributes to recommendation performance. By efficiently utilizing the sparse interaction data in the target domain, TCN can capture users' evolving interest preferences, thereby improving the recommendation accuracy of the model from the perspective of graph-structured user-item interactions.

- The performance degradation in “w/o Frequency-domain filtering” suggests the presence of noise and pseudo-semantics in the cluster encoding received from the client-side, indicating that frequency-domain filtering can help refine cluster encodings associated with graph node semantics.
- By comparing SFC-CDR with the “w/o TFRG” module, we can see that the fusion of temporal dependencies in user interactions and semantic features contributes to forecasting users' next interaction and jointly refining graph node semantic representations with temporal interaction information.

4.4. Impact of Hyperparameters (RQ3)

On the OnlineRetail-Pantry platform, we conduct experiments on three key hyperparameters: the TCN convolution kernel size s , the number of clustering centroids K , and the semantic encoding layer a . We aim to explore the impact of these parameters on SFC-CDR. The experimental results are presented in Figures 3 and 4 for s , Figures 5 and 6 for K , and Figures 7 and 8 for a .

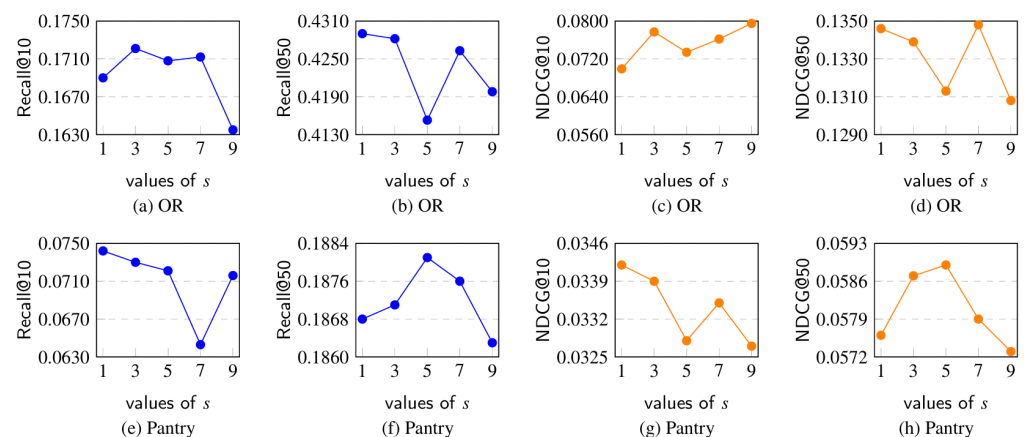


Figure 3. Impact of the TCN convolution kernel size s on OnlineRetail-Pantry.

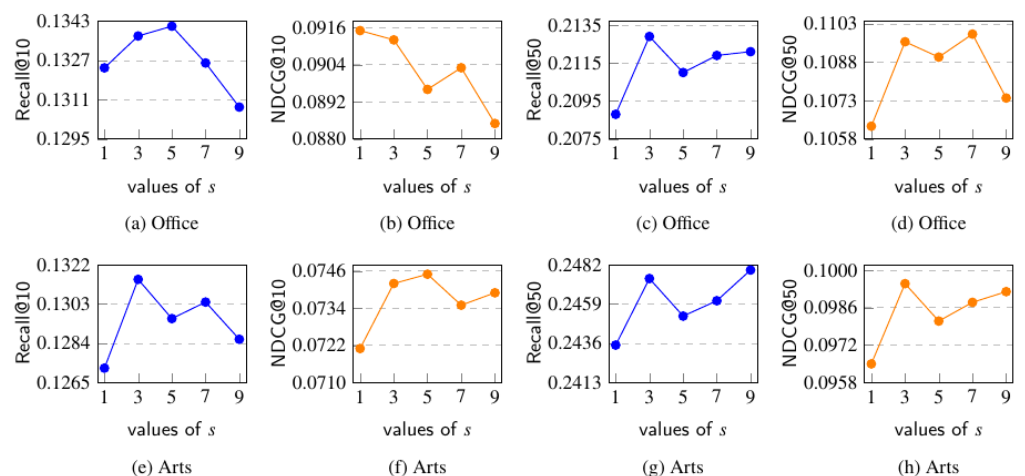


Figure 4. Impact of the TCN convolution kernel size s on Office-Arts.

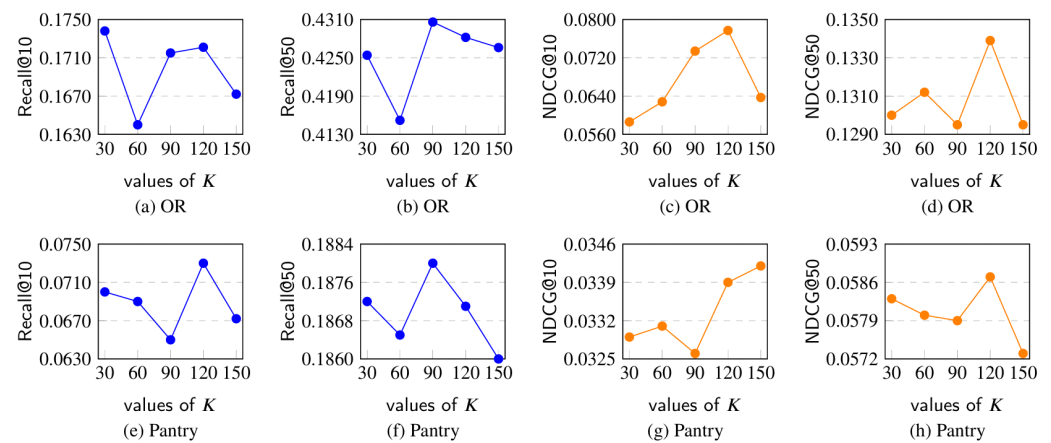


Figure 5. Impact of the number of clustering centroids K on OnlineRetail-Pantry.

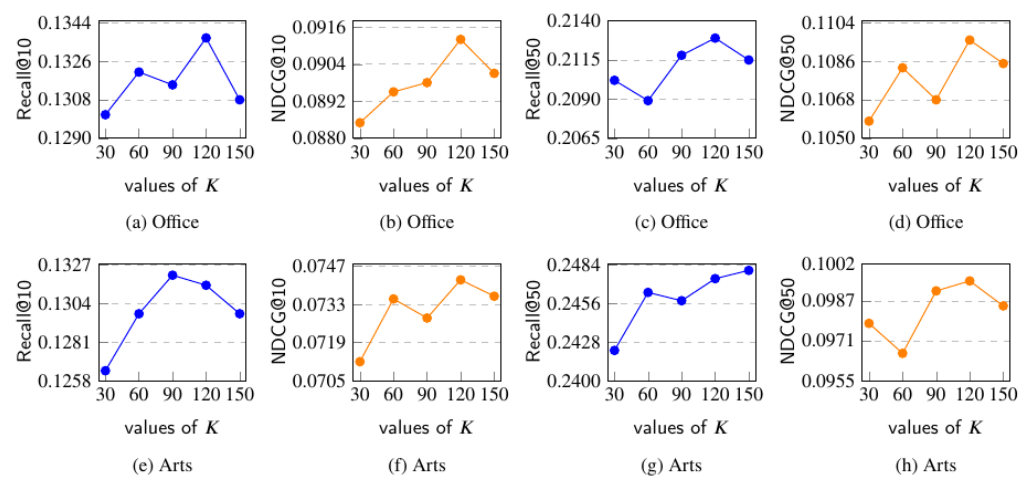


Figure 6. Impact of the number of clustering centroids K on Office-Arts.

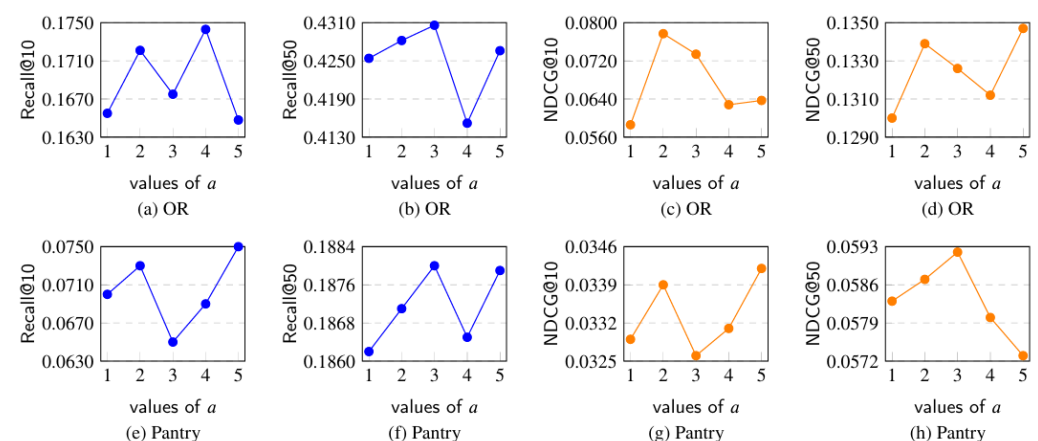


Figure 7. Impact of the semantic encoding layer a in PLM on OnlineRetail-Pantry.

4.4.1. Impact of s

In Figures 3 and 4, when s is set to 3, SFC-CDR can achieve the best effect. This shows that the convolution kernel size in the model for modeling temporal dependencies of graph-structured user–item interaction sequences effectively balances performance, noise robustness, and model complexity.

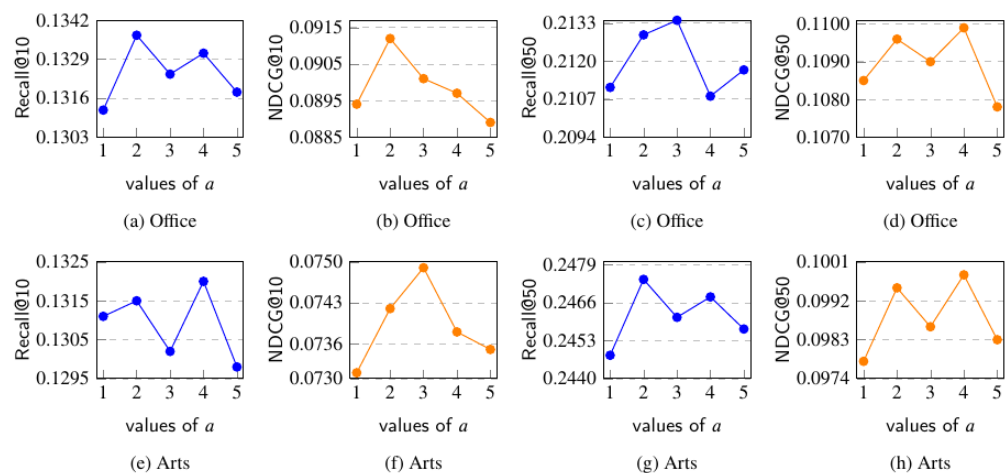


Figure 8. Impact of the semantic encoding layer a in PLM on Office-Arts.

4.4.2. Impact of K

From Figures 5 and 6, we can observe that, when the number of clustering centroids is 120, the best performance can be obtained. If the clustering centroids are smaller, they will lead to insufficient data in each category; if the clustering centroids are larger, it will lead to the clustering results not being able to reflect the true structure of the data and the semantic organization of node representations across interaction graphs from different domains, making the model results deviate from the true values.

4.4.3. Impact of a

The variation in the number of semantic encoding tiers in PLM has a non-negligible impact on SFC-CDR. Figures 7 and 8 intuitively show that in the dimensionality of feature fetching, $a = 2$ can form a good balance between holding key semantic information of graph nodes and suppressing redundant noise.

4.5. Different Coarse-Grained Alignment Methods (RQ4)

Since the setting involves non-overlapping users and items and must satisfy privacy-preserving constraints, we adopt an unsupervised clustering approach to achieve alignment of node semantic representations across interaction graphs from different domains. We select three mainstream clustering methods: DEC, K-means, and K-means++, and compare them with SFC-CDR under the same experimental setting with a fixed number of clustering centers. The experimental outcomes are shown in Table 5. It can be observed that K-means++ slightly outperforms DEC on the OnlineRetail dataset, while SFC-CDR achieves the best performance on most metrics across all datasets. This indicates that SFC-CDR enables better alignment of semantic features associated with graph nodes, thereby improving recommendation performance.

Table 5. Results of clustering methods on *OnlineRetail-Pantry* and *Office-Arts*.

Variants	OR				Pantry				Office				Arts			
	Recall		NDCG		Recall		NDCG		Recall		NDCG		Recall		NDCG	
	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50	@10	@50
DEC	0.1719	0.4312	0.0773	0.1337	0.0729	0.1866	0.0335	0.0573	0.1329	0.2112	0.0906	0.1084	0.1303	0.2458	0.0740	0.0989
K-means	0.1708	0.4185	0.0762	0.1325	0.0721	0.1859	0.0330	0.0573	0.1326	0.2119	0.0908	0.1088	0.1308	0.2466	0.0736	0.0986
K-means++	0.1717	0.4236	0.0771	0.1334	0.0728	0.1867	0.0334	0.0582	0.1341	0.2121	0.0910	0.1089	0.1311	0.2464	0.0749	0.0989
SFC-CDR	0.1721	0.428	0.0777	0.1339	0.073	0.1871	0.0339	0.0587	0.1337	0.2129	0.0912	0.1096	0.1315	0.2474	0.0742	0.0995

4.6. Cross-Domain Semantic Consistency (RQ5)

To further examine the semantic consistency of the proposed cross-domain clustering, we compare the semantic similarity of cross-domain item pairs within the same cluster (intra-cluster) and across different clusters (inter-cluster). As shown in Tables 6 and 7, both the original PLM-based evaluation and the cross-PLM evaluation using Sentence-BERT show a consistent pattern: the Intra-cluster semantic similarity is higher than the inter-cluster similarity for both domain pairs. These results suggest that the proposed clustering captures meaningful semantic relationships among cross-domain items.

Table 6. Cross-domain semantic consistency between intra-cluster and inter-cluster item pairs.

Domain Pair	Intra-Cluster	Inter-Cluster	Δ_{semantic}
OR–Pantry	0.8335	0.7674	+0.0661
Office–Arts	0.8876	0.8513	+0.0363

Table 7. Cross-domain semantic consistency between intra-cluster and inter-cluster item pairs (cross-PLM validation via Sentence-BERT).

Domain Pair	Intra-Cluster	Inter-Cluster	Δ_{semantic}
OR–Pantry	0.1714	0.1565	+0.0149
Office–Arts	0.2610	0.2251	+0.0358

5. Conclusions

In this paper, we have proposed a Semantic Federated Clustering-driven Cross-Domain Recommendation model (SFC-CDR) for cross-domain recommendation scenarios where entities in the source and target domains are non-overlapping. From the interaction-graph perspective, these scenarios correspond to interaction graphs from different domains with no explicit shared-node correspondence. Specifically, a clustering anchor contrastive learning module captures deep semantic correlations among node text representations to achieve more precise semantic space alignment, enabling knowledge transfer across non-overlapping users and items. To thoroughly filter redundant semantic noise, we introduce Joint Temporal–Frequency Filtering (JTFF) that operates on both temporal dependencies of user behaviors in graph-structured user–item interactions and dependency relationships in text semantic features associated with graph nodes, thereby enhancing the model’s robustness and accuracy against heterogeneous noise. Experimental results on four public datasets demonstrate that the proposed model significantly outperforms all baseline models.

In the future, we plan to explore time-series noise processing to better extract long-term user behavior preferences from graph-structured interaction sequences. We also realize that our model is highly dependent on item text descriptions as semantic attributes of graph nodes, and its applicability can be significantly limited if the text information in the dataset is missing or insufficient. The evaluation is conducted on two domain pairs from e-commerce domains; generalization to more heterogeneous domains remains to be explored.

Author Contributions: Conceptualization, D.H.; methodology, D.H. and L.L.; data curation, L.L.; formal analysis, D.H.; investigation, L.L.; project administration, D.H.; software, L.L.; supervision, D.H. and C.S.; funding acquisition, C.S.; resources, C.S.; validation, L.L.; visualization, L.L.; writing—review and editing, D.H.; writing—original draft preparation, L.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Chongqing Municipal Education Commission Key Project of Higher Education in Chongqing (Grant No. CQZSKS2025012).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data supporting the findings of this study are available from the corresponding author upon reasonable request. The code is not publicly available due to ongoing research.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analysis, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Jin, K.; Yu, W.; Long, Y.; Lai, N.; Huang, J. A diffusion multi-interest framework for cross-domain recommendation. *Expert Syst. Appl.* **2025**, *294*, 128738.
2. Duan, L.; Zhu, L.; Ren, P. A Dynamic Cross-Domain Recommendation Model with Target-Aware Complementary Preference Transfer and Information Fusion. *Eng. Appl. Artif. Intell.* **2025**, *148*, 110404.
3. Xiao, S.; Chen, R.; Song, H.; Han, Q. Cross-domain recommendation via interest-aware pseudo-overlapping user alignment. *Expert Syst. Appl.* **2025**, *292*, 128638.
4. Chang, J.; Yu, F. Knowledge-guided clustering based algorithm for cross-domain recommendation problems with partially overlapping users and multiple source-domains. *Expert Syst. Appl.* **2025**, *284*, 127936.
5. Man, T.; Shen, H.; Jin, X.; Cheng, X. Cross-domain recommendation: an embedding and mapping approach. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August 2017; pp. 2464–2470.
6. Kang, S.; Hwang, J.; Lee, D.; Yu, H. Semi-Supervised Learning for Cross-Domain Recommendation to Cold-Start Users. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 1563–1572.
7. Hu, G.; Zhang, Y.; Yang, Q. CoNet: Collaborative Cross Networks for Cross-Domain Recommendation. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Turin, Italy, 22–26 October 2018; pp. 667–676.
8. Chen, C.; Yang, Y.; Zhou, J.; Li, X.; Bao, F. Cross-Domain Review Helpfulness Prediction Based on Convolutional Neural Networks with Auxiliary Domain Discriminators. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, New Orleans, LO, USA, 2–4 June 2018; Volume 2 (Short Papers), pp. 602–607.
9. Gao, X.; Ding, L.; Chen, J.; Yang, Y.; Xiang, Y. User group-enhanced user feature distribution transfer framework for non-overlapping cross-domain recommendations. *Knowl.-Based Syst.* **2025**, *314*, 113186.
10. Lu, Z.; Guo, L.; Yu, X.; Cheng, Z.; Han, X.; Zhu, L. Federated Semantic Learning for Privacy-preserving Cross-domain Recommendation. *ACM Trans. Inf. Syst.* **2025**, *43*, 1–27.
11. Hou, Y.; Mu, S.; Zhao, W.X.; Li, Y.; Ding, B.; Wen, J. Towards Universal Sequence Representation Learning for Recommender Systems. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '22), Washington, DC, USA, 14–18 August 2022; pp. 585–593.
12. Kong, M.; Hou, M.; Zhao, S.; Liu, F.; Su, R.; Chen, Y. DADIN: Domain Adversarial Deep Interest Network for cross domain recommender systems. *Expert Syst. Appl.* **2024**, *243*, 122880.
13. Xu, J.; Gan, M.; Zhang, H.; Zhang, S. IDC-CDR: Cross-domain Recommendation based on Intent Disentanglement and Contrast Learning. *Inf. Process. Manag.* **2024**, *61*, 103871.
14. Wang, L.; Sang, L.; Zhang, Q.; Wu, Q.; Xu, M. A privacy-preserving framework with multi-modal data for cross-domain recommendation. *Knowl. Based Syst.* **2024**, *304*, 112529.
15. Jin, C.; Duan, Y.; Zhou, L.; Li, F. Cross-domain recommender system with embedding- and mapping-based knowledge correlation. *Knowl. Based Syst.* **2024**, *304*, 112514.
16. Zhang, H.; Zheng, D.; Yang, X.; Feng, J.; Liao, Q. FedDCSR: Federated Cross-domain Sequential Recommendation via Disentangled Representation Learning. In Proceedings of the 2024 SIAM International Conference on Data Mining (SDM), Houston, TX, USA, 18–20 April 2024; pp. 535–543.
17. Zhao, W.; Xiao, Y.; Li, T.; Wang, R.; Li, Q.; Wang, G. A Cross-Domain Recommendation Model Based on Asymmetric Vertical Federated Learning and Heterogeneous Representation. *IEEE Trans. Emerg. Top. Comput. Intell.* **2025**, *9*, 2344–2358.
18. Zhu, G.; Liu, X.; Tang, S.; Niu, J. Aligning Before Aggregating: Enabling Communication Efficient Cross-Domain Federated Learning via Consistent Feature Extraction. *IEEE Trans. Mob. Comput.* **2024**, *23*, 5880–5896.

19. Liang, K.; Li, S.; Ding, M.; Tian, F.; Wu, Y. Privacy-Preserving Coded Schemes for Multi-Server Federated Learning with Straggling Links. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 1222–1236.
20. Liu, M.; Li, J.; Guo, Z.; Li, G. Extracting latently overlapping users by graph neural network for non-overlapping cross-domain recommendation. *Knowl.-Based Syst.* **2024**, *290*, 111508.
21. Tian, C.; Xie, Y.; Chen, X.; Li, Y.; Zhao, X. Privacy-preserving Cross-domain Recommendation with Federated Graph Learning. *ACM Trans. Inf. Syst.* **2024**, *42*, 1–29.
22. Xie, Y.; Yu, C.; Jin, X.; Cheng, L.; Hu, B.; Li, Z. Heterogeneous graph contrastive learning for cold start cross-domain recommendation. *Knowl. Based Syst.* **2024**, *299*, 112054.
23. Yang, Z.; Peng, Z.; Wang, Z.; Qi, J.; Chen, C.; Pan, W.; Wen, C.; Wang, C.; Fan, X. Federated Graph Learning for Cross-Domain Recommendation. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 64865–64888.
24. Do, P.; Zhang, Q.; Zhang, G.; Lu, J. meta-GRS: A Graph Neural Network for Cross-Domain Recommender System via Meta-Learning. *Procedia Comput. Sci.* **2023**, *225*, 2536–2545.
25. Zhu, X.; Yin, Y.; Wang, L. MIMNet: Multi-interest Meta Network with Multi-granularity Target-guided Attention for cross-domain recommendation. *Neurocomputing* **2025**, *620*, 129208.
26. Guan, R.; Pang, H.; Giunchiglia, F.; Liang, Y.; Feng, X. Cross-Domain Meta-Learner for Cold-Start Recommendation. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 7829–7843.
27. Yan, D.; Zhao, Y.; Yang, Z.; Jin, Y.; Zhang, Y. FedCDR: Privacy-preserving federated cross-domain recommendation. *Digit. Commun. Netw.* **2022**, *8*, 552–560.
28. Wang, Y.; Xie, Q.; Tang, M.; Bao, Z.; Li, L.; Liu, Y. Knowledge Memory Graph convolution network for cross-domain recommendation. *Knowl. Based Syst.* **2025**, *317*, 113415.
29. Zhu, X.; Duan, L.; Dietze, S.; Yu, R. TJMN: Target-enhanced joint meta network with contrastive learning for cross-domain recommendation. *Knowl.-Based Syst.* **2025**, *310*, 112919.
30. Liu, W.; Zheng, X.; Hu, M.; Chen, C. Collaborative Filtering with Attribution Alignment for Review-based Non-overlapped Cross Domain Recommendation. In Proceedings of the ACM Web Conference 2022, Lyon, France, 25–29 April 2022; pp. 1181–1190.
31. Liu, S.; Hao, X.; Chen, X.; Ma, W.; Gu, F. Personalized cross-domain recommendation with meta networks and contrastive learning. *Appl. Soft Comput.* **2025**, *184*, 113827.
32. Liu, W.; Chen, C.; Liao, X.; Hu, M.; Su, J.; Tan, Y.; Wang, F. User Distribution Mapping Modelling with Collaborative Filtering for Cross Domain Recommendation. In Proceedings of the ACM Web Conference 2024, Singapore, 13–17 May 2024; pp. 334–343.
33. Zhao, G.; Zhang, X.; Tang, H.; Shen, J.; Qian, X. Domain-Oriented Knowledge Transfer for Cross-Domain Recommendation. *IEEE Trans. Multimed.* **2024**, *26*, 9539–9550.
34. Zhang, F.; Xiong, Y.; Shi, P.; Ding, L. Deep User Rating Pattern Mining and Fusion Inference Method for Cross-Domain Recommendation. *Expert Syst. Appl.* **2025**, *269*, 126374.
35. Maneechote, N.; Maneeroj, S.; Takasu, A. Cross-Domain Recommendation with Personalized Rating Pattern Compatibility as Transfer Rate Between Domains. *IEEE Access* **2025**, *13*, 119978–119998.
36. Wang, Y.; Xie, Q.; Tang, M.; Li, L.; Yuan, J.; Liu, Y. A Dual Perspective Framework of Knowledge-correlation for Cross-domain Recommendation. *ACM Trans. Knowl. Discov. Data* **2024**, *18*, 28.
37. Tian, X. Content-based Filtering for Improving Movie Recommender System. In Proceedings of the 2023 International Conference on Data Science, Advanced Algorithm and Intelligent Computing (DAI 2023), Shanghai, China, 24–26 November 2024; pp. 598–609.
38. Guan, J.; Yin, Z.; Zhang, T.; Chen, L.; Zhang, Y.; Huang, F.; Han, S.; Chen, J. Effective Two-Stage Knowledge Transfer for Multi-Entity Cross-Domain Recommendation. In Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2, 5, Toronto, ON, Canada, 3–7 August 2025; pp. 4467–4475.
39. Gan, C.; Huang, B.; Hu, B.; Ma, J.; Liu, Z.; Zhang, Z.; Zhou, J.; Zhang, G.; Zhong, W. PEACE: Prototype Learning Augmented transferable framework for Cross-domain Recommendation. In Proceedings of the 17th ACM International Conference on Web Search and Data Mining, Mérida, Mexico, 4–8 March 2024; pp. 228–237.
40. Guo, L.; Lu, Z.; Yu, J.; Nguyen, Q.; Yin, H. Prompt-enhanced Federated Content Representation Learning for Cross-domain Recommendation. In Proceedings of the ACM on Web Conference 2024, Singapore, 13–17 May 2024; pp. 3139–3149.
41. Huai, Z.; Yang, Y.; Zhang, M.; Zhang, Z.; Li, Y.; Wu, W. M2GNN: Metapath and Multi-interest Aggregated Graph Neural Network for Tag-based Cross-domain Recommendation. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Taipei, Taiwan, 23–27 June 2023; pp. 1468–1477.
42. Wang, K.; Zhang, J.; Liu, K. Enhancing Cross-Domain Recommendation with Plug-In Contrastive Representations from Large Language Models. In Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval, Padova, Italy, 13–18 July 2025; pp. 1799–1809.
43. Kim, K.; Sohn, M.; Kim, J. Cross-Domain Explainable Recommendation Using Graph Convolutional Networks and a Topic Model. *IEEE Access* **2025**, *13*, 118310–118323.

44. Gautam, C.; Sinha, B.B. LLM-Boosted Graph Contrastive Learning for Cold-Start Cross-Domain Recommendation. In 2026 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI), Virtual Event, 12–14 March 2026; pp. 1–6.
45. Soundariya, S.; Ramakrishnan, S.; Devadharshini, S.K.; Madhumita, S.; Promodini, L. Cross Domain Recommendation System Using Graph Embedding and Attention Techniques. In Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Mandi, India, 24–28 June 2024; pp. 1–8.
46. Li, Y.; Hou, L.; Li, J. Towards Knowledge-Aware and Deep Reinforced Cross-Domain Recommendation Over Collaborative Knowledge Graph. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 7171–7187.
47. Yu, X.; Sheng, B. GBTCDR: Cross-Domain Recommendation Based on Graph Representation Learning and Bridge Transfer. In Proceedings of the 2026 International Conference on Artificial Intelligence and Data Mining (AIDM), Chengdu, China, 22–24 May 2026; pp. 234–237.
48. Huang, X.; Huang, L.; Gao, Y.; Li, Z.; Lai, P.; Wang, C.; Yu, P. Knowledge-Aware Graph Prompt Tuning for Cross-Domain Recommendation. *IEEE Trans. Multimed.* **2025**, *12*, 3923–3937.
49. Ni, X.; Nie, J.; Zuo, Z.; Xie, H.; Liang, X.; Jiang, M.; Xu, J.; Yu, S.; Liu, M. DITN: User’s indirect side-information involved domain-invariant feature transfer network for cross-domain recommendation. *Inf. Process. Manag.* **2023**, *60*, 103494.
50. Wang, L.; Qi, Y.; Bai, Y.; Sun, Z.; Li, D.; Li, X. MuKGB-CRS: Guarantee privacy and authenticity of cross-domain recommendation via multi-feature knowledge graph integrated blockchain. *Inf. Sci.* **2023**, *638*, 118915.
51. Yuan, X.; Qi, A.; Wu, H.; Wang, J.; Guo, Y.; Li, S.; Zhao, L. Cross-modal feature alignment and fusion with contrastive learning in multimodal recommendation. *Knowl. Based Syst.* **2025**, *326*, 114020.
52. Bai, Z.; Zheng, Y.; Yang, P.; Liu, S.; Zhang, Y.; Chang, Y. DCRLRec: Dual-domain contrastive reinforcement large language model for recommendation. *Inf. Process. Manag.* **2025**, *62*, 104140.
53. Lai, Y.; Huang, S.; Chen, C.; Wu, C. Multi-interest transfer using contrastive learning for cross-domain recommendation. *Decis. Support Syst.* **2025**, *195*, 114473.
54. Won, H.; Oh, B.; Yang, H.; Lee, K. Cross-modal contrastive learning for aspect-based recommendation. *Inf. Fusion* **2023**, *99*, 101858.
55. Ni, R.; Cai, W.; Jiang, Y. Contrastive cross-domain sequential recommendation via emphasized intention features. *Neural Netw.* **2024**, *179*, 106488.
56. Chang, H.; Gu, L.; Hu, C.; Zhang, Z.; Zhu, H.; Xu, Y.; Fang, Y.; Chen, Z. Separated Contrastive Learning for Matching in Cross-domain Recommendation with Curriculum Scheduling. In Proceedings of the Companion Proceedings of the ACM on Web Conference 2025, New York, NY, USA, 28 April–2 May 2025; pp. 114–123.
57. Devlin, J.; Chang, M.; Lee, K.; Toutanova, K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Minneapolis, MN, USA, 2–7 June 2019; Volume 1, pp. 4171–4186.
58. Xie, R.; Liu, Q.; Wang, L.; Liu, S.; Zhang, B.; Lin, L. Contrastive cross-domain recommendation in matching. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, DC, USA, 14–18 August 2022; pp. 4226–4236.
59. Li, C.; Zhao, M.; Zhang, H.; Yu, C.; Cheng, L.; Shu, G.; Kong, B.; Niu, D. RecGURU: Adversarial learning of generalized user representations for cross-domain recommendation. In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, Virtual Event, 21–25 February 2022; pp. 571–581.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.