information

MDPI

Article

MGGCL: Motif-Guided Graph Contrastive Learning for Recommendation

Li Pang [1], Yuqi Zhang [1], Nancy Wang [1], Jian Yu [1,*] and Deling Huang [2]

[1] School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, Auckland 1010, New Zealand; li.pang@autuni.ac.nz (L.P.); yuqi.zhang@aut.ac.nz (Y.Z.); nancy.wang@autuni.ac.nz (N.W.)
[2] School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; huangdl@cqupt.edu.cn
* Correspondence: jian.yu@aut.ac.nz

Abstract

Graph motifs capture crucial structural patterns within user–item interaction graphs and offer meaningful semantics that are typically underutilised in conventional graph-based collaborative filtering. Existing contrastive learning methods in recommender systems rely primarily on simple perturbations, which limits their ability to leverage deeper motif-based structural information. To address this gap, we propose a novel motif-guided contrastive learning framework for recommender systems. To exploit complementary structural biases inherent to user–item interactions, our approach explicitly incorporates three distinct motifs into the construction of the contrastive view. By contrasting views that capture the structural differences among different motifs, our model learns meaningful group-level relationships and potentially suppresses noise arising from sparse or isolated interactions. Extensive experiments on four real-world recommendation datasets validate that our motif-based contrastive approach achieves the best overall performance, outperforming state-of-the-art baselines on three of the four benchmarks with statistically significant margins on the more skewed datasets, while remaining competitive on the fourth, which demonstrates notable robustness and improved accuracy.

Keywords: recommender systems; collaborative filtering; graph contrastive learning; network motifs

1. Introduction

Graph-based collaborative filtering has gained considerable attention within recommender systems due to its effectiveness in modelling complex user–item interactions through graph neural networks (GNNs) [1–3]. The core strength of graph-based collaborative filtering lies in its capacity to represent users and items as interconnected nodes, enabling efficient exploitation of relational data and structural dependencies. This approach has notably improved predictive accuracy and facilitated richer interaction modelling compared to traditional recommendation methods.

Recent advancements [4,5] in recommender systems have introduced graph contrastive learning (GCL), a self-supervised framework designed to enhance robustness and representation quality. GCL operates by generating multiple augmented graph views through perturbation strategies and maximising agreement between embeddings derived from these diverse views. This process has shown significant potential in addressing prominent challenges such as data sparsity, susceptibility to noisy interactions, and incomplete

Academic Editors: Costas Vassilakis and Dionisis Margaris

Received: 28 May 2026
Revised: 24 June 2026
Accepted: 30 June 2026
Published: 1 July 2026

information. Consequently, GCL techniques have shown improved generalisation and robustness in various recommendation scenarios [6,7]. Additionally, GCL have also been successfully applied in broader applications such as computer vision and natural language processing [8], legal precedent analysis [9], and bioinformatics [10].

However, conventional GCL methods typically employ simple and uniformly random augmentation techniques, such as node dropout or edge perturbations, which can inadvertently distort critical structural information and remove meaningful interactions. Such random augmentations have been shown to limit the quality and interpretability of learnt graph representations [11–15]. In response to these limitations, recent approaches have started exploring more structured and semantically meaningful augmentation methods, aiming to preserve valuable relational and structural context. For example, some methods integrate socially-informed augmentations to retain homophily signals, which is crucial for social recommendations [16,17], while others leverage hierarchical clustering techniques to maintain multi-resolution structural semantics [18,19]. Additionally, explicit neighbour-aware methods have been proposed to reduce the randomness inherent in conventional sampling by considering structural neighbours as positive pairs [6,20]. Some other strategies formalise a variety of graph operations through principled message-passing frameworks or global statistical transformations to retain important local and global interaction patterns [21,22].

Despite the progress these structured methods represent, they commonly rely on implicit assumptions or heuristic augmentation strategies, limiting explicit interpretability and failing to differentiate effectively between strong and weak user–item interactions. To bridge this gap, we propose a novel contrastive learning framework called motif-guided graph contrastive learning (MGGCL). MGGCL integrates the structural semantics into graph-based collaborative filtering by constructing contrastive views with motifs, which are frequently recurrent subgraph patterns that capture multi-hop local interactions. To be specific, we select two star motifs and a rectangle motif that represent different structural semantics for recommendation graphs. After that, we prove that the rectangle-motif-induced adjacency structure is contained within the union of the structures captured by the two star motifs. This provides justification for selecting the three motifs to construct contrastive views, as our analysis demonstrates that similarities between these views capture robust correlations among users and items, while their differences emphasise sparse or isolated user preferences. We then construct contrastive views by augmenting the adjacency matrix with motif adjacency matrices so that the view embeddings are generated with motif-guided neighbourhood aggregation.

By strategically constructing and contrasting motif-specific views, MGGCL emphasises meaningful structural signals and suppresses noise from isolated or sparse interactions, substantially improving the robustness and performance of recommendations while making the view-construction process structurally transparent. Thus, MGGCL systematically addresses the limited transparency of existing random and heuristic approaches by explicitly incorporating higher-order structural semantics into contrastive view construction.

In many real-world recommendation graphs, interactions are heavily skewed. Popular users and items create dense star-shaped motifs that can dominate learnt representations, while semantically richer co-interaction rectangle motifs remain under-represented. Our motif-guided contrastive learning framework explicitly constructs two graph views: one preserves all star motifs, and the other retains rectangle motifs. We refer to star-exclusive edges that are not part of any retained rectangle motif as hub-only edges. Since rectangle-engaged edges are reinforced in both views, whereas hub-only edges receive motif weight only in the star view, the contrastive gradient weakens hub-only links and amplifies structurally meaningful co-interaction patterns, as derived in Section 4.4. Empirical results

on four benchmarks confirm this mechanism in practice. MGGCL yields the largest gains on highly skewed Douban-Book and iFashion datasets, where the improvements in Recall and NDCG over state-of-the-art baselines are substantial and consistent across trials.

Our main contributions are as follows.

- We design a novel motif-guided graph contrastive learning framework that explicitly integrates structural semantics into graph view construction through motif-based augmentations. Unlike conventional stochastic perturbations, our motif-based augmentation strategy provides clear structural transparency: because each view is built from explicitly defined motifs, the construction lays the groundwork for future interpretability and attribution of model behaviour to specific local graph patterns.
- We formally establish a hierarchical relationship between the motif adjacency matrices constructed from star motifs and rectangle motifs, where edges forming rectangles represent higher-order, structurally robust user–item interactions. By explicitly contrasting the rectangle-motif view against the star-motifs view, the model is optimised to strengthen node representations derived from these rectangle patterns where nodes have stronger correlation. Concurrently, the model naturally reduces the influence of weaker interactions that appear only in the star-motifs view, leading to improved robustness of the learnt representations.
- We demonstrate that incorporating motif-aware adjacency matrices into contrastive views enhances neighbourhood aggregation efficiency by explicitly capturing meaningful multi-hop structural relationships. This structural richness allows the model to achieve improved node embeddings through contrastive objectives, leading to better discriminative power and representation robustness compared to standard stochastic augmentation strategies.
- We conduct extensive experiments on four real-world recommendation benchmarks and achieve consistent accuracy improvements. In particular, skewed datasets benefit effectively from MGGCL's ability to reduce popularity bias from overly popular nodes and to amplify under-represented structurally meaningful motifs. Systematic parameter sensitivity analysis further demonstrates stable performance without exhaustive hyperparameter tuning.

The rest of the paper is organised as follows. In Section 2, we briefly review and discuss existing research on graph-based contrastive learning approaches and motif-based approaches for recommender systems. Section 3 provides the background knowledge to understand our proposed method. Section 4 presents our proposed MGGCL framework. Section 5 shows our experimental results and performance analysis. Finally, Section 6 concludes this paper.

2. Related Work

This section discusses three key areas relevant to our study: (1) GCL recommendation models with simple random augmentations; (2) GCL recommendation models that generate contrastive views using structural and semantic augmentations; and (3) the use of motifs in recommender systems.

2.1. GCL Recommendation Models with Simple Random Augmentations

Graph contrastive learning has emerged as a powerful self-supervised learning paradigm for graph-based recommendation. It focuses on learning robust user and item representations by contrasting multiple views of the interaction graph. This approach addresses data sparsity and improves generalisation without relying on explicit supervision. The foundational work by Wu et al. [4] introduced SGL, which creates contrastive views through random node and edge dropout. SGL encourages agreement between embeddings

derived from these views, thereby mitigating sparsity and enhancing robustness. However, this stochastic augmentation strategy can distort meaningful graph structures, which leads to inconsistencies in semantics and reduced interpretability.

To address these limitations, later models explored simplified or more efficient augmentation strategies. SimGCL [7] avoids explicit structural perturbations by injecting Gaussian noise directly into node embeddings, preserving graph connectivity while maintaining competitive performance. SelfCF [23] shifts the focus to augmentation in the output space by applying contrastive learning directly to the final embeddings. By avoiding alteration on the input graph and reliance on negative sampling, this method offers a more lightweight and efficient training process.

Several models aim to refine the contrastive signal or improve its robustness. BDCL [24] applies directional perturbations, preserving sparse behavioural data without random corruption. RecDCL [25] introduces both batch-wise and feature-wise contrastive objectives to improve embedding orthogonality. DCCR [26] uses a self-adaptive noise mechanism and constructs dual collaborative views to extract high-value samples and enhance recommendation diversity. GGDHSCL [27] introduces a generative diffusion mechanism coupled with hard negative sampling, which strengthens the contrastive signal rather than modifying view structures.

Recent work has also emphasised improved handling of noisy or uninformative views. SymmetricGCL [15] introduces a denoising strategy using symmetric contrastive pairs to mitigate degradation from noisy augmentations. WeightedGCL [28] enhances SGL with dynamic view-level weighting that allows more informative contrastive views to contribute more significantly to learning.

To explore more stable and informative view construction, NLGCL [20] generates contrastive views from naturally existing neighbour layers instead of random perturbations, reducing semantic drift while maintaining graph topology. DimCL [29] augments representations in a dimension-aware fashion to encourage diverse embedding components for contrastive learning.

To support privacy and personalisation, personalised federated contrastive learning [30] adapts GCL to the federated setting by aggregating user-device interactions without centralising user data. This design introduces personalisation and scalability but requires careful calibration to maintain embedding consistency across clients.

Collectively, these works illustrate a trajectory from random augmentations toward more principled and robust contrastive learning techniques. Yet, many still operate with implicit structure and offer limited topological semantics, motivating structurally grounded alternatives.

2.2. Methods with Structured and Semantically Guided Augmentations

Recognising the limitations of uniform random perturbations, several researchers have explored structured or semantically informed augmentations. SG-CLR [16] enhances the semantic integrity of contrastive views by guiding edge additions based on homophily. HGCL [18] takes a hierarchical perspective by clustering items into multiple resolution layers to preserve coarse-to-fine structural semantics. NCL [6] advances the idea of local structure by treating topological neighbours as natural positives rather than sampling corrupted ones. This strategy captures contextual consistency and reduces semantic drift during augmentation.

JCG [31] combines structural and semantic contrastive objectives to address data imbalance and reduce the influence of high-degree nodes. This improves generalisation across heterogeneous graph distributions. MD-GCCF [32] advances multi-view learning by integrating deep collaborative signals and designing both local and global contrastive

objectives. By mitigating over-smoothing and reducing redundancy, this architecture yields more informative representations for recommendation.

Several models introduce learnable or deterministic transformation schemes. CL-KDM [22] bypasses traditional dropout strategies entirely by maximising statistical dependence between views, which preserves global latent geometries without edge removal. LMACL [33] combines GCN and GAT to learn augmentation strategies to generate multi-hop and neighbour-aware contrastive views. These methods present promising steps toward augmentations that encode high-quality semantics or structure. However, many rely on heuristics or predefined rules and rarely integrate clear higher-order graph patterns such as motifs into the view generation process. Recent contrastive methods have also begun to address noise and bias more explicitly: for instance, counterfactual graph contrastive learning generates counterfactual structural augmentations and maximises agreement between them to suppress spurious, bias-inducing correlations [34]. This is complementary to our motif-guided contrast, which suppresses noise structurally by contrasting rectangle-engaged co-interactions against sparse hub-only edges.

2.3. Motif-Based Approaches for Graph Recommendation and Contrastive Learning

Motifs are recurrent subgraph structures such as stars and rectangles that offer a principled way to capture higher-order relationships and encode meaningful local semantics in graphs. Although extensively studied in biological and social networks, motifs have only recently been adopted in recommender systems. Their ability to reflect structured behavioural patterns, such as co-engagement, item co-consumption, or user hubs, makes them suitable for addressing sparsity and noise.

Recent studies have explored motif-based enhancements in recommender systems. Zhang et al. [35] propose a framework that employs star and rectangle motifs to define richer neighbourhoods in collaborative filtering. Wang et al. [36] extend Graph Attention Networks by integrating motif structures to capture high-order service relationships. Motifs-Res [37] formulates motif co-occurrence as hyperedges and applies hypergraph convolution with contrastive loss to exploit local motif semantics. MotifSR [38] incorporates motif-enhanced temporal subgraphs to better capture session-level user behaviours in sequential recommendation. These models demonstrate the growing interest in exploiting motif semantics to strengthen signal quality and model generalisation in graph-based recommendation.

Complementing this, motif-based contrastive learning has been studied in other domains. Motif-driven contrastive learning of graph representations [39] introduces motif-induced views to enforce global and local consistency in general graph representation learning, where motif co-occurrence defines semantically consistent neighbourhoods. MotifCC [40] proposes a framework that builds subgraphs from triangle motifs and applies contrastive learning between lower-order and higher-order views to jointly preserve structural uniqueness and encourage clustering separability. MHGCL [41] jointly leverages motif structures and node homogeneity to refine the selection of contrastive pairs, thereby suppressing noise and improving representation discrimination in general graph tasks. Although effective, these works do not target recommender systems and typically operate in node classification or community detection settings.

Our work departs from the above by being the first to integrate motif-based structures into contrastive view generation specifically for collaborative filtering. Instead of treating motifs as auxiliary enhancements or post-processing filters, we construct motif-induced adjacency matrices and use them in structurally meaningful view construction. Rectangle motifs capture higher-order co-engagement signals, while the combination of star motifs additionally includes sparse preferences or transient behaviours. By contrasting these motif-guided views, our approach prioritises robust local structures and reduces potential

noise arisen from sparse or isolated patterns. This approach unifies structural transparency, semantic augmentation, and robustness within the contrastive learning framework for graph-based recommendation, and provides a clear basis for future interpretability analysis. More broadly, the motif-based structural priors we employ can be situated within the wider paradigm of graph self-supervised and representation learning [42,43], which encompasses the contrastive techniques surveyed above.

3. Preliminaries

3.1. Motifs in Recommendation Graphs

In recommender system research, user–item interactions are typically represented using bipartite graphs. Bipartite graphs naturally fit collaborative filtering tasks since they capture the relationships between two distinct node sets, i.e., users and items. The formal definition of bipartite graphs is provided below.

Definition 1 (Bipartite Graphs [44]). *A bipartite graph $G = (U, V, E)$ is a graph where all nodes in G are partitioned into two disjoint classes U and V, i.e., $U \cap V = \emptyset$, such that every edge has its terminal node in different classes from its initial node: nodes in the same partition class must not be adjacent.*

Motifs are significant patterns that frequently recur within graphs, appearing more frequently than expected in randomised graphs [45]. Since motifs typically include more than two nodes, they offer valuable insight into the local structural semantics of the graph while simultaneously influencing its global characteristics. Figure 1 illustrates bipartite motifs composed of up to four nodes. Users are represented by circles, and items by squares. Following the naming system proposed in [35], these motifs are named M_b^a where a indicates the number of user nodes in the motif while b indicates the number of item nodes in the motif. For ease of understanding, we refer to these motifs according to their topological structures: M_2^1, M_3^1, M_1^2, and M_1^3 are termed star motifs, M_2^{2-} is termed a path motif, and M_2^2 is termed a rectangle motif.

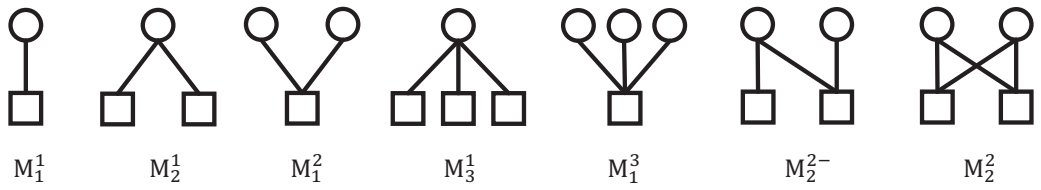

Figure 1. All motifs consisting of two to four nodes in bipartite graphs. The superscript indicates the number of user nodes in the motif, whereas the subscript denotes the number of item nodes. The hyphen in M_2^{2-} indicates one less edge in the motif compared to M_2^2.

A common approach to incorporate network motifs into graph learning frameworks is to construct motif adjacency matrices.

Definition 2 (Motif Adjacency Matrix [46]). *Given a graph $G = (V, E)$ and a motif m, the motif adjacency matrix $\mathbf{A}_m \in \mathbb{N}^{|V|\times|V|}$ is defined such that $(\mathbf{A}_m)_{ij}$ equals the number of instances of motif m in which nodes i and j co-occur.*

Generating $\mathbf{A}_m$ typically involves subgraph isomorphism countings between the motif m and all subgraphs in G. In this paper, we use the algorithms proposed in [35] to efficiently generate motif adjacency matrices. The paper provides dedicated algorithms with their time-complexity analysis for all bipartite motifs of up to four nodes. This generation is a one-time offline pre-computation. The resulting matrices are stored in sparse format

and reused across all training runs and hyperparameter settings, so the motif-counting cost is amortised during preprocessing and does not contribute to the per-run training or inference complexity.

3.2. Contrastive Learning for Recommender Systems

In the context of graph-based collaborative filtering, contrastive learning (CL) operates by generating multiple views of the user–item interaction graph through augmentations such as edge or node dropout. These augmented views are processed by graph encoders to produce representations, and the model is trained to maximise the agreement between representations of the same node across different views while distinguishing between different nodes.

The joint learning scheme used in CL methods is a combination of the recommendation loss $\mathcal{L}_{\text{rec}}$ and the contrastive loss $\mathcal{L}_{\text{cl}}$:

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{cl}}, \tag{1}$$

where the CL task serves as an auxiliary objective, with its influence controlled by a hyperparameter λ. For the recommendation loss, a common choice is the Bayesian personalised rank (BPR) loss:

$$\mathcal{L}_{\text{rec}} = -\sum_{u,i \in \mathcal{B}} \log(\sigma(\mathbf{z}_u^{\top}\mathbf{z}_i - \mathbf{z}_u^{\top}\mathbf{z}_j)), \tag{2}$$

where $\mathcal{B}$ is the sampled batch, σ is the sigmoid function, $\mathbf{z}_u$ is the embedding vector for user u, $\mathbf{z}_i$ and $\mathbf{z}_j$ are embedding vectors for positive item i and negative item j, respectively.

The common choice for contrastive loss is the InfoNCE loss:

$$\mathcal{L}_{\text{cl}} = -\sum_{i \in \mathcal{B}} \log \frac{\exp\left(\phi(\mathbf{z}_i^{(1)}, \mathbf{z}_i^{(2)})/\tau\right)}{\sum_{j \in \mathcal{B}} \exp\left(\phi(\mathbf{z}_i^{(1)}, \mathbf{z}_j^{(2)})/\tau\right)}, \tag{3}$$

where $\phi(\cdot,\cdot)$ denotes a similarity function, typically the cosine similarity; $\mathbf{z}_i^{(1)}$ and $\mathbf{z}_i^{(2)}$ are the representations of the same sample i under two different views or augmentations; $\mathbf{z}_j^{(2)}$ is the second-view representation of another sample j in the batch $\mathcal{B}$; and τ is the temperature parameter that controls the sharpness of the distribution.

4. Methodology

In this section, we present our MGGCL framework. We first explain our motivation for using motifs in contrastive learning to learn structural semantics, and then introduce the design of the model architecture.

4.1. Motif-Guided Structural Semantics

Traditional contrastive learning in recommendation systems often relies on random edge dropout or uniform augmentation to generate views, which risks discarding meaningful interaction patterns or amplifying noise. Our approach instead leverages motif-guided structural semantics to create semantically distinct views. Specifically, we choose the three motifs with different structural semantics: M_2^1 implies potential similarity or complementarity between items, M_1^2 reflects possible shared interests between users, and M_2^2 indicates stronger relevance among users and items. We do not choose M_3^1, M_1^3 and M_2^{2-} for the following reasons: M_3^1 and M_1^3 are expanded versions of M_2^1 and M_1^2 respectively, sharing similar structural semantics but excluding isolated or peripheral sparse subgraphs present in M_2^1 and M_1^2; although M_2^{2-} contains structures of M_2^1 and M_1^2, it blurs individual-level signals by mixing two users via a high-degree item or two items by a "shopaholic" user.

By contrasting views constructed with M_2^2 and the combination of M_2^1 and M_1^2, we explicitly decouple two fundamental aspects of user–item interactions: (1) star motifs, i.e., M_2^1 and M_1^2, which reflect preferences or transient behaviours, and (2) rectangle motif, i.e., M_2^2, which encodes robust group behaviours such as item communities or shared user tastes. The contrast arises from the observation that edges appearing in M_2^1 and M_1^2 but absent in M_2^2 often correspond to low-consistency signals, as their interactions are not strengthened by broader user–item coalitions. These differences are inherently meaningful as they capture less common behaviours or noises that lack reinforcement from similar user–item interactions. By contrasting these views, the model learns to prioritise interactions validated by both local and global structural contexts while retaining sensitivity to sparse but informative signals. This approach avoids the opacity of random augmentation, since the constructed views are directly aligned with observable interaction semantics in real-world bipartite systems.

4.2. Structural Hierarchy Between Motif Adjacency Matrices

We briefly formalise the structural hierarchy between motif adjacency matrices used to define the two contrastive views:

Proposition 1. *Let $\mathbf{A}_2^1$ and $\mathbf{A}_1^2$ denote the motif adjacency matrices derived from star motifs M_2^1 and M_1^2 respectively, and let $\mathbf{A}_2^2$ denote the motif adjacency matrix corresponding to the rectangle motif M_2^2. Then*

$$supp(\mathbf{A}_2^2) \subseteq supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2)$$

where $supp(\mathbf{A})$ is the set of index pairs (i, j) for which $\mathbf{A}_{ij} \neq 0$.

Proof. Consider a rectangle motif instance with users u_1, u_2 and items i_1, i_2. It contains the four edges (u_1, i_1), (u_1, i_2), (u_2, i_1), (u_2, i_2). Each of these edges belongs to at least one of M_2^1 or M_1^2. Edge (u_1, i_1) appears in an M_2^1 instance (u_1, i_1, i_2) and an M_1^2 instance (i_1, u_1, u_2). Analogous reasoning holds for the other three edges.

If $(p, q) \in supp(\mathbf{A}_2^2)$, then some instances of $\mathbf{A}_2^2$ contain (p, q), and the above guarantees that (p, q) appears in at least one instance of M_2^1 or M_1^2. Hence $(p, q) \in supp(\mathbf{A}_2^1)$ or $(p, q) \in supp(\mathbf{A}_1^2)$, which implies

$$(p, q) \in supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2).$$

Since every edge in $supp(\mathbf{A}_2^2)$ is also in $supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2)$, then

$$supp(\mathbf{A}_2^2) \subset supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2).$$

When every instance of M_2^1 and M_1^2 is "closed" by at least one instance of M_2^2, $supp(\mathbf{A}_2^2) = supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2)$.

Hence we have

$$supp(\mathbf{A}_2^2) \subseteq supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2).$$

□

Remark 1. *The inclusion in Proposition 1 is* strict *for the majority of all real-world graphs. Equality can occur only in the degenerate case where* every *edge that participates in an instance of a M_2^1 or M_1^2 also belongs to at least one instance of M_2^2. Equivalently, every two-hop user–item path ($u \to i \to u$ or $i \to u \to i$) is "closed" by an additional user or item, forming a M_2^2. Such a closure is highly unlikely in the sparse interaction graphs typical of recommender systems, so we treat the strict inclusion case $supp(\mathbf{A}_2^2) \subset supp(\mathbf{A}_2^1 \cup \mathbf{A}_1^2)$ as the practical default.*

By definition, any two users connected to two common items inherently contain the edge connections constituting star motifs as substructures. Specifically, the rectangle motif can always be decomposed into combinations of star motifs, guaranteeing hierarchical inclusion. Consequently, this structural nesting explicitly isolates the incremental higher-order interactions captured by rectangles, clearly delineating which local structural signals each contrastive view encodes. This deterministic hierarchical structure significantly narrows the potential explanation space. Although direct interpretability experiments are beyond the current scope, the transparent construction mechanism provides a clear path for future analyses, such as motif attribution and saliency enrichment, which can systematically reveal the influence of these motifs.

4.3. Motif-Guided Graph Contrastive Learning

In this section, we present the architecture of our framework motif-guided graph contrastive learning (MGGCL), as illustrated in Figure 2. Our approach leverages motifs to construct contrasting views for self-supervised learning in recommendation graphs. Let $\mathcal{G} = (\mathcal{U}, \mathcal{V}, \mathcal{E})$ denote a bipartite graph with users $\mathcal{U}$, items $\mathcal{V}$, and edges $\mathcal{E}$ representing interactions. We select the following three motifs to capture distinct interaction structures:

- **Motif M_2^1**: User-centric co-interactions, represented by pairs $\{(u, i_1), (u, i_2)\}$ where a user u interacts with two items i_1, i_2.
- **Motif M_1^2**: Item-centric co-interactions, represented by pairs $\{(u_1, i), (u_2, i)\}$ where two users u_1, u_2 interact with the same item i.
- **Motif M_2^2**: Dense bipartite hubs, represented by a 2×2 complete bipartite graph $\{(u_1, i_1), (u_1, i_2), (u_2, i_1), (u_2, i_2)\}$. These motifs are stronger collaborative signals since they reflect group behaviour or shared preferences.

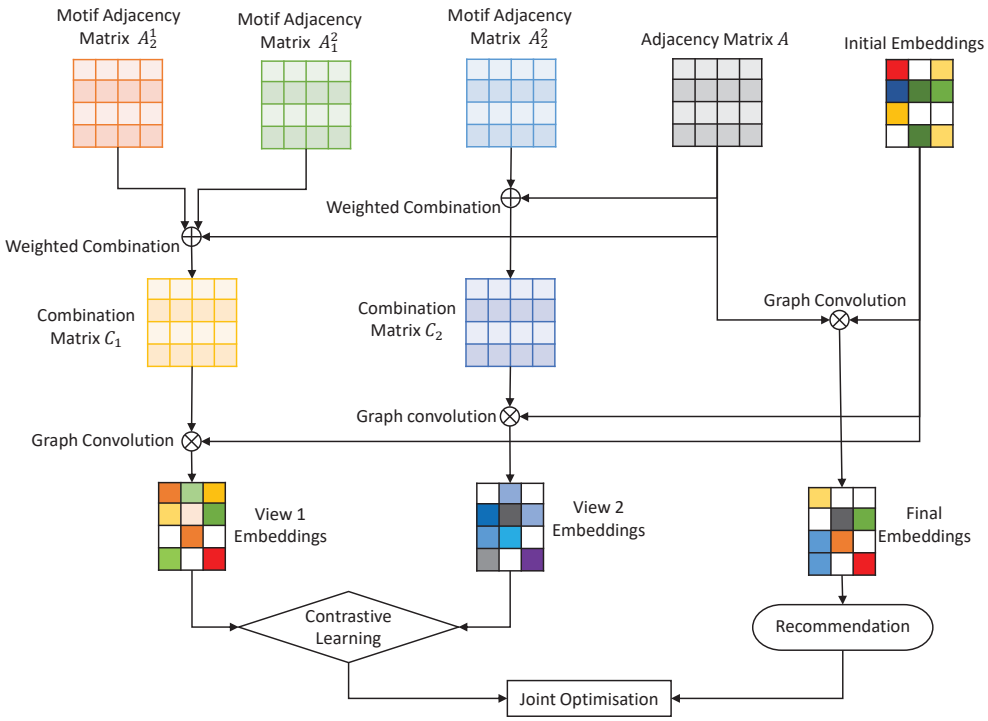

Figure 2. Architecture of the motif-guided graph contrastive learning framework

Two contrastive views are generated by combining motif adjacency matrices with the original adjacency matrix $\mathbf{A}$:

$$\mathbf{C}_1 = (1-\theta)\cdot\widetilde{\mathbf{A}} + \theta\cdot\frac{\widetilde{\mathbf{A}}_2^1 + \widetilde{\mathbf{A}}_1^2}{2}, \tag{4}$$

$$\mathbf{C}_2 = (1-\theta)\cdot\widetilde{\mathbf{A}} + \theta\cdot\widetilde{\mathbf{A}}_2^2, \tag{5}$$

where $\widetilde{\mathbf{A}}$ is normalised adjacency matrix; $\widetilde{\mathbf{A}}_2^1$, $\widetilde{\mathbf{A}}_1^2$, and $\widetilde{\mathbf{A}}_2^2$ are normalised motif adjacency matrices of the three corresponding motifs; $\theta \in (0,1]$ is the motif ratio that balances the contribution between the original adjacency matrix and the motif adjacency matrices, which enables the model to adjust to graphs with varying levels of motif density. $\mathbf{C}_1$ emphasises sparse pairwise interactions, while $\mathbf{C}_2$ amplifies dense hub structures. Both matrices are row-normalised to stabilise training.

Node embeddings are generated using a shared graph encoder $f_\theta(\cdot)$ based on LightGCN [2], which avoids nonlinearities and normalisation for efficiency. We adopt LightGCN because it is the standard backbone of the contrastive-learning recommenders we compare against, so the comparison isolates the effect of the motif-guided view construction rather than that of a more expressive encoder. This view construction is encoder-agnostic and could easily be combined with more expressive backbones. The encoder propagates embeddings as

$$\mathbf{H}^{(l)} = \mathbf{C}\cdot\mathbf{H}^{(l-1)}, \tag{6}$$

where $\mathbf{C} \in \{\mathbf{C}_1, \mathbf{C}_2\}$ and $\mathbf{H}^{(l)}$ is the l-th layer embedding. The final embedding $\mathbf{z}_u$ for user u (and analogously for items) is obtained by averaging all layer outputs:

$$\mathbf{z}_u = \frac{1}{L}\sum_{l=1}^{L}\mathbf{H}_u^{(l)}. \tag{7}$$

By augmenting the adjacency matrix with motif-induced adjacency matrices, the graph convolution layer is guided by higher-order structural patterns and aggregates information more selectively. Unlike simply stacking additional convolution layers, which would enlarge the receptive field uniformly and indiscriminately amplify high-degree hubs, the motif augmentation extends the receptive field in a structure-selective manner. It emphasises aggregation through co-interaction structures while down-weighting isolated hub-only links. The benefit is therefore not simply an enlarged receptive field, but a structurally selective one achieved without adding graph convolution layers. As a result, the contrastive learning process better reflects meaningful local and global graph structures rather than popularity alone, thereby improving recommendation performance. In terms of cost, both $\mathbf{C}_1$ and $\mathbf{C}_2$ are sparse matrices, so their memory cost scales with the number of non-zero motif-induced entries, and each propagation step $\mathbf{C}\cdot\mathbf{H}^{(l-1)}$ is a sparse-dense matrix product with the same per-layer cost structure as LightGCN.

We employ a noise-tolerant contrastive loss to maximise mutual information between cross-view user and item embeddings while accounting for the view-specific motif adjacency matrices used in neighbourhood aggregation. For a user u, embeddings $\mathbf{z}_u^{C_1}$ and $\mathbf{z}_u^{C_2}$ from $\mathbf{C}_1$ and $\mathbf{C}_2$ form a positive pair. Negative pairs are sampled from other users/items in the batch. The loss is defined as

$$\mathcal{L}_{\text{cl}} = -\sum_{u\in\mathcal{U}}\log\frac{\exp\left(\phi(\mathbf{z}_u^{C_1}, \mathbf{z}_u^{C_2})/\tau\right)}{\sum_{v\in\mathcal{U}}\exp\left(\phi(\mathbf{z}_u^{C_1}, \mathbf{z}_v^{C_2})/\tau\right)}, \tag{8}$$

where $\phi(\cdot,\cdot)$ is the cosine similarity function introduced in Equation (3) and τ is the temperature hyperparameter. The full objective combines $\mathcal{L}_{\text{cl}}$ with a BPR loss $\mathcal{L}_{\text{rec}}$:

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{cl}}, \tag{9}$$

where λ controls the contribution of contrastive learning signals. The model retains all edges but implicitly reduces the weights of hub-only edges exclusive to M_2^1 and M_1^2, treating them as structured noise relative to the dominant hub structures in M_2^2. We make this down-weighting effect explicit through a worked example below.

4.4. A Worked Example of Down-Weighting Hub-Only Edges

The reduced weighting of hub-only edges described above is explained in the following example. We illustrate it on the minimal bipartite graph in Figure 3 (left), where users u_1 and u_2 both interact with items i_1 and i_2, forming a rectangle, while u_3 interacts only with the popular item i_1. The edge (u_3, i_1) is therefore a hub-only edge that participates in star motifs but in no rectangle.

Figure 3 presents the motif adjacency matrices, with nonzero cells shaded. The boxed u_3 row across the three motif matrices illustrates Proposition 1 directly. The support of $\mathbf{A}_2^2$ is contained in the support of $\mathbf{A}_2^1 \cup \mathbf{A}_1^2$. In particular, the hub-only edge (u_3, i_1) appears in $\mathbf{A}_1^2$, where it is supported by two users sharing i_1, but is absent from $\mathbf{A}_2^2$ because it is not closed by any rectangle motif. It is also absent from $\mathbf{A}_2^1$ as u_3 has no second item, so among the motifs, the hub-only edge is retained only by the item-centric star $\mathbf{A}_1^2$. Consequently, in the two contrastive views at $\theta = 1$, where $\mathbf{C}_1 = (\widetilde{\mathbf{A}}_2^1 + \widetilde{\mathbf{A}}_1^2)/2$ and $\mathbf{C}_2 = \widetilde{\mathbf{A}_2^2}$, the entire u_3 row of $\mathbf{C}_2$ is zero. Under the rectangle view u_3 is an isolated node that aggregates no neighbourhood signal, whereas in $\mathbf{C}_1$ it retains non-zero weights.

This structural asymmetry is what drives the gradient behaviour. The contrastive loss $\mathcal{L}_{\text{cl}}$ in Equation (8) maximises the agreement $\phi(\mathbf{z}_u^{C_1}, \mathbf{z}_u^{C_2})$ between a node's two view embeddings. Considering a single propagation step for clarity, $\mathbf{z}_u^{C_k} = \sum_j (\mathbf{C}_k)_{uj} \mathbf{h}_j$, the cross-view discrepancy for node u is governed by the difference between the u-th rows of $\mathbf{C}_1$ and $\mathbf{C}_2$. By Proposition 1, this row-wise difference is supported exactly on the hub-only edges incident to u. For a hub node such as u_3, this discrepancy becomes particularly pronounced because its $\mathbf{C}_2$ row is empty. Consequently, the alignment gradient concentrates on the view-specific hub-only contribution, which is discounted during contrastive optimisation. The same conclusion holds for the full L-layer averaged embedding $\mathbf{z}_u = \frac{1}{L}\sum_{l=1}^{L} \mathbf{H}_u^{(l)}$. Since u_3 has an empty row in $\mathbf{C}_2$, the corresponding row remains empty in every power $\mathbf{C}_2^l$. Therefore, u_3 remains disconnected in the rectangle view at all propagation depths. Rectangle-engaged co-interactions are reinforced consistently in both views and therefore incur no such penalty; instead, they are amplified. Equivalently, the alignment term $\phi(\mathbf{z}_u^{C_1}, \mathbf{z}_u^{C_2})$ in the InfoNCE objective is satisfied most readily for nodes whose two views agree. Rectangle edges therefore contribute more strongly to mutual-information maximisation, whereas hub-only edges behave as structured noise that the objective suppresses. More generally, for $\theta < 1$, a hub-only edge retains a weight proportional to $(1-\theta)$ in $\mathbf{C}_2$, whereas its weight in $\mathbf{C}_1$ includes both the $(1-\theta)$ term and motif reinforcement. Thus, the edge is relatively down-weighted rather than fully removed. The case $\theta = 1$ shown here represents the sharpest instance and is also the optimal setting on three of our four datasets.

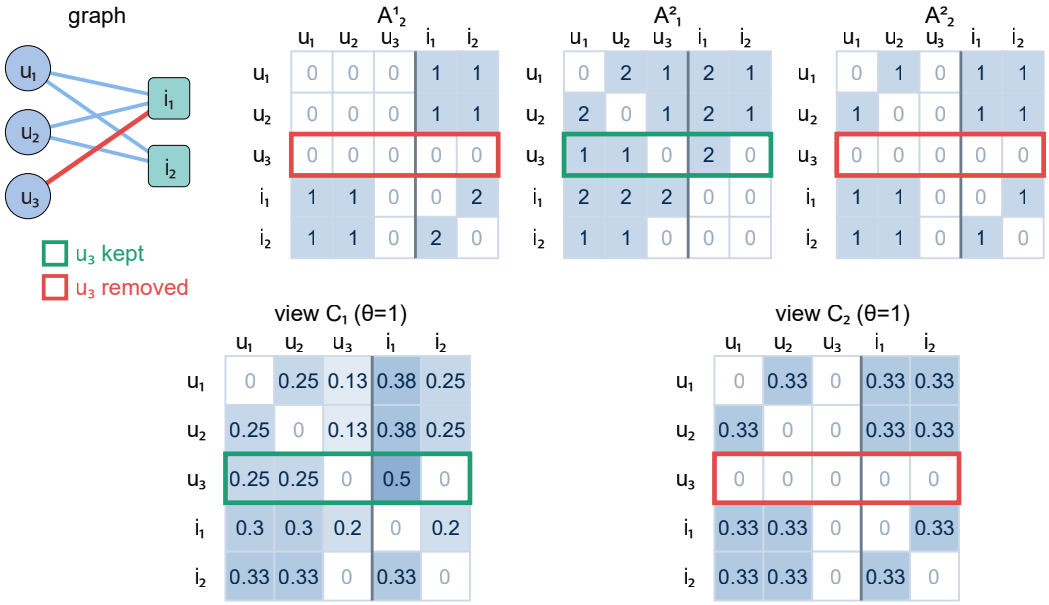

Figure 3. Worked example of how the motif-guided views down-weight a hub-only edge. **Left**: a bipartite graph in which u_1 and u_2 co-interact with i_1 and i_2 (a rectangle), while u_3 links only the hub item i_1. **Top**: the motif adjacency matrices $\mathbf{A}_2^1$, $\mathbf{A}_1^2$, $\mathbf{A}_2^2$ in square form, with nonzero (support) cells shaded; the boxed u_3 row is retained only by $\mathbf{A}_1^2$ (green) and is absent from $\mathbf{A}_2^1$ and $\mathbf{A}_2^2$ (red). **Bottom**: the row-normalised contrastive views at $\theta = 1$; the u_3 row is nonzero in $\mathbf{C}_1$ but entirely zero in $\mathbf{C}_2$, so the hub-only edge (u_3, i_1) is isolated under the rectangle view.

5. Experiments

5.1. Experimental Settings

We conduct experiments on four public benchmark datasets: Douban-Book [7], iFashion [4], Amazon-Kindle [47] and Amazon-Book [2]. We follow prior work in using the Gini coefficient to measure inequality in item distribution [48] and extend it to quantify both user and item popularity inequality in our datasets. The statistics of these datasets are presented in Table 1. All records are randomly split into training, validation and test sets at a ratio of 7:1:2. We adopt two widely used evaluation metrics, Recall@K and NDCG@K, with $K = 5, 10, 20, 40$. All results are based on the average of five experimental trials.

Table 1. Statistics of the datasets.

Dataset	# of Users	# of Items	# of Edges	Density	User-Gini	Item-Gini
Douban-Book	12,859	22,294	598,420	0.00209	0.624	0.6455
iFashion	300,000	81,614	1,607,813	0.00007	0.3569	0.7666
Amazon-Kindle	138,333	77,801	1,909,965	0.00014	0.5426	0.5422
Amazon-Book	52,643	91,599	2,984,108	0.00062	0.4468	0.4617

5.2. Baselines

We select the following baseline models for performance comparison.

- LightGCN [2] is a simplified GCN model for recommender systems. It keeps the linear propagation of node embeddings while removing all feature transformations and non-linear activation functions in the model architecture.
- NCL [6] is a model that boosts graph collaborative filtering by using contrastive learning enriched with neighbourhood information.
- SGL [4] is a self-supervised CF model that constructs contrasting views by randomly dropping edges or nodes from the graph.

- SelfCF [23] is a self-supervised framework that improves collaborative filtering by augmenting output embeddings without negative sampling.
- SimGCL [7] is a self-supervised CF model that constructs contrasting views by simply adding noise to the node embeddings.
- GDMSR [49] is a recent diffusion-based graph recommendation model that corrupts and denoises the user–item bipartite graph to recover higher-order collaborative signals.
- MGCN [35] is a motif-based collaborative filtering model that propagates over a combination of the normalised adjacency and motif adjacency matrices, where the motif component may use an individual motif or a combination of motifs such as star and rectangle motifs. As a motif-augmented graph collaborative filtering framework without a contrastive objective, it is the most directly comparable non-contrastive motif-based baseline to MGGCL.

5.3. Hyperparameter Settings

The hyperparameters of all baseline methods are tuned according to their original papers. For a fair comparison, the following settings are shared by all models except GDMSR, which specifically requires an embedding size of 1000 and batch size of 400: the embedding size is 32, the batch size is 4096, and the number of MLPs or graph convolution layers is fixed at two.

For MGGCL, we search the learning rate over $\{1e^{-4}, 1e^{-3}, 1e^{-2}\}$ and the L_2 regularisation coefficient over $\{1e^{-5}, 1e^{-4}, 1e^{-3}\}$. The contrastive learning coefficient λ is searched over $\{0.01, 0.05, 0.1, 0.2, 0.5, 1, 2, 5\}$ and the motif ratio θ over $\{0.01, 0.05, 0.1, 0.2, 0.5, 1\}$, whose sensitivity is analysed in Figures 4 and 5. The temperature τ in the contrastive loss (Equations (3) and (8)) is fixed at 0.2 for all datasets, following SGL [4] and SimGCL [7], where $\tau = 0.2$ is reported as empirically optimal.

5.4. Performance Comparison

We present the performance comparison in Table 2. Our proposed MGGCL specifically leverages motif-guided structural semantics, distinguishing it from traditional contrastive learning methods that rely on random edge dropout or uniform augmentation. Evaluating the results across four real-world datasets, MGGCL achieves the best overall performance, attaining the best results on three of the four datasets.

Douban-Book: Among the four benchmark datasets, Douban-Book exhibits the highest interaction density, with users engaging in approximately 46.5 rating events on average and each item receiving 26.8 ratings. Nonetheless, the item-popularity Gini coefficient of around 0.65 indicates a long-tail skewed distribution. This indicates that some popular books receive disproportionately close attention compared to the majority of other niche items. The user-side distribution is of a similar nature, as the user-popularity Gini coefficient is 0.62, which shows that a small proportion of power users contribute a particularly large number of ratings. The inequalities for user and item sides show that interaction patterns are not uniformly distributed. Instead, hubs and clusters are displayed.

Considering the moderate connectivity and long-tailed popularity displayed in the dataset, star-shaped motifs around bestsellers and rectangle motifs among reader niches are both well-represented in the interaction graph. Hence, MGGCL can use the motif abundance in the dataset to refine neighbourhood aggregation, which then effectively leads to enhanced recommendation performance. As shown in Table 2, the Recall@5 and NDCG@5 are 27.8% and 32.1% higher than the strongest baseline (NCL and MGCN, respectively). The model's explicit rectangle-vs-star contrast particularly boosts NDCG. This shows that even in datasets with already abundant user–item interactions, the model achieves additional performance gains by focusing on higher-order and structurally closed patterns.

Table 2. Performance comparison.

Dataset	Model	Recall@5	NDCG@5	Recall@10	NDCG@10	Recall@20	NDCG@20	Recall@40	NDCG@40
Douban-Book	LightGCN	0.05131 ± 0.00040	0.09851 ± 0.00040	0.08345 ± 0.00060	0.09868 ± 0.00050	0.12720 ± 0.00080	0.10662 ± 0.00060	0.18881 ± 0.00110	0.12462 ± 0.00080
	NCL	0.06189 ± 0.00078	0.10724 ± 0.00087	0.09416 ± 0.00089	0.10696 ± 0.00088	0.13840 ± 0.00179	0.11534 ± 0.00099	0.19833 ± 0.00111	0.13326 ± 0.00078
	SGL	0.02765 ± 0.00100	0.04562 ± 0.00080	0.04267 ± 0.00140	0.04653 ± 0.00110	0.06066 ± 0.00200	0.04962 ± 0.00140	0.08224 ± 0.00260	0.05543 ± 0.00180
	SelfCF	0.03269 ± 0.00080	0.06171 ± 0.00060	0.05577 ± 0.00110	0.06306 ± 0.00080	0.08698 ± 0.00160	0.06954 ± 0.00110	0.13724 ± 0.00210	0.08474 ± 0.00140
	SimGCL	0.04641 ± 0.00093	0.07628 ± 0.00120	0.06822 ± 0.00124	0.07634 ± 0.00111	0.09862 ± 0.00169	0.08245 ± 0.00125	0.13711 ± 0.00214	0.09403 ± 0.00125
	GDMSR	0.02950 ± 0.00090	0.05620 ± 0.00070	0.04700 ± 0.00120	0.05560 ± 0.00090	0.07410 ± 0.00170	0.06070 ± 0.00120	0.11110 ± 0.00220	0.07130 ± 0.00150
	MGCN	0.05589 ± 0.00038	0.10853 ± 0.00060	0.09024 ± 0.00052	0.10802 ± 0.00058	0.13620 ± 0.00066	0.11600 ± 0.00064	0.19830 ± 0.00082	0.13355 ± 0.00074
	MGGCL	**0.07910** ± 0.00066	**0.14342** ± 0.00065	**0.11548** ± 0.00011	**0.13882** ± 0.00042	**0.16124** ± 0.00030	**0.14424** ± 0.00025	**0.21507** ± 0.00035	**0.15844** ± 0.00034
iFashion	LightGCN	0.03396 ± 0.00040	0.02338 ± 0.00040	0.05338 ± 0.00060	0.02985 ± 0.00050	0.08137 ± 0.00080	0.03726 ± 0.00060	0.11999 ± 0.00110	0.04557 ± 0.00080
	NCL	0.02428 ± 0.00070	0.01685 ± 0.00060	0.03831 ± 0.00100	0.02153 ± 0.00080	0.05923 ± 0.00140	0.02707 ± 0.00100	0.08953 ± 0.00180	0.03358 ± 0.00130
	SGL	0.03956 ± 0.00018	0.02763 ± 0.00012	0.06028 ± 0.00045	0.03455 ± 0.00020	0.08959 ± 0.00064	0.04230 ± 0.00026	0.12953 ± 0.00066	0.05090 ± 0.00026
	SelfCF	0.02489 ± 0.00080	0.01695 ± 0.00060	0.04006 ± 0.00110	0.02200 ± 0.00080	0.06362 ± 0.00160	0.02819 ± 0.00110	0.09945 ± 0.00210	0.03585 ± 0.00140
	SimGCL	0.03748 ± 0.00026	0.02644 ± 0.00021	0.05628 ± 0.00028	0.03273 ± 0.00021	0.08175 ± 0.00027	0.03948 ± 0.00019	0.11541 ± 0.00031	0.04676 ± 0.00021
	GDMSR	0.01310 ± 0.00090	0.00910 ± 0.00070	0.02010 ± 0.00120	0.01150 ± 0.00090	0.02980 ± 0.00170	0.01410 ± 0.00120	0.04460 ± 0.00220	0.01730 ± 0.00150
	MGCN	0.03852 ± 0.00062	0.02686 ± 0.00042	0.05977 ± 0.00078	0.03400 ± 0.00050	0.08852 ± 0.00096	0.04165 ± 0.00060	0.12655 ± 0.00118	0.04989 ± 0.00072
	MGGCL	**0.04929** ± 0.00018	**0.03455** ± 0.00018	**0.07479** ± 0.00033	**0.04307** ± 0.00027	**0.10945** ± 0.00030	**0.05224** ± 0.00027	**0.15567** ± 0.00086	**0.06221** ± 0.00037

https://doi.org/10.3390/info17070645

Table 2. *Cont.*

Dataset	Model	Recall@5	NDCG@5	Recall@10	NDCG@10	Recall@20	NDCG@20	Recall@40	NDCG@40
Amazon-Kindle	LightGCN	0.07224 ± 0.00040	0.06478 ± 0.00040	0.10752 ± 0.00060	0.07673 ± 0.00050	0.15187 ± 0.00080	0.09029 ± 0.00060	0.20721 ± 0.00110	0.10493 ± 0.00080
	NCL	0.05060 ± 0.00070	0.04338 ± 0.00060	0.08057 ± 0.00100	0.05382 ± 0.00080	0.12128 ± 0.00140	0.06624 ± 0.00100	0.17040 ± 0.00180	0.07934 ± 0.00130
	SGL	0.06228 ± 0.00100	0.05520 ± 0.00080	0.09737 ± 0.00140	0.06728 ± 0.00110	0.14247 ± 0.00200	0.08097 ± 0.00140	0.19977 ± 0.00260	0.09612 ± 0.00180
	SelfCF	0.01919 ± 0.00080	0.01630 ± 0.00060	0.03084 ± 0.00110	0.02040 ± 0.00080	0.05053 ± 0.00160	0.02637 ± 0.00110	0.07936 ± 0.00210	0.03383 ± 0.00140
	SimGCL	0.06732 ± 0.00047	0.05973 ± 0.00053	0.10175 ± 0.00042	0.07142 ± 0.00040	0.14468 ± 0.00051	0.08445 ± 0.00038	0.19721 ± 0.00029	0.09834 ± 0.00030
	GDMSR	0.03400 ± 0.00090	0.03620 ± 0.00070	0.04730 ± 0.00120	0.03940 ± 0.00090	0.06430 ± 0.00170	0.04400 ± 0.00120	0.08630 ± 0.00220	0.04990 ± 0.00150
	MGCN	**0.08127** ± 0.00050	**0.07415** ± 0.00045	**0.11701** ± 0.00066	**0.08590** ± 0.00052	**0.16274** ± 0.00082	**0.09963** ± 0.00062	**0.21764** ± 0.00102	**0.11422** ± 0.00074
	MGGCL	**0.07250** ± 0.00018	**0.06572** ± 0.00034	**0.10900** ± 0.00005	**0.07793** ± 0.00021	**0.15591** ± 0.00049	**0.09209** ± 0.00026	**0.21230** ± 0.00073	**0.10701** ± 0.00034
Amazon-Book	LightGCN	0.01096 ± 0.00040	0.01884 ± 0.00040	0.01955 ± 0.00060	0.02096 ± 0.00050	0.03407 ± 0.00080	0.02664 ± 0.00060	0.05754 ± 0.00110	0.03548 ± 0.00080
	NCL	0.01115 ± 0.00070	0.01911 ± 0.00060	0.02015 ± 0.00100	0.02158 ± 0.00080	0.03506 ± 0.00140	0.02743 ± 0.00100	0.05871 ± 0.00180	0.03636 ± 0.00130
	SGL	0.01252 ± 0.00100	0.02133 ± 0.00080	0.02226 ± 0.00140	0.02384 ± 0.00110	0.03779 ± 0.00200	0.02985 ± 0.00140	0.06284 ± 0.00260	0.03925 ± 0.00180
	SelfCF	0.00724 ± 0.00080	0.01344 ± 0.00060	0.01305 ± 0.00110	0.01465 ± 0.00080	0.02228 ± 0.00160	0.01809 ± 0.00110	0.03718 ± 0.00210	0.02365 ± 0.00140
	SimGCL	0.01404 ± 0.00008	0.02352 ± 0.00015	0.02436 ± 0.00006	0.02593 ± 0.00008	0.04062 ± 0.00015	0.03223 ± 0.00010	0.06454 ± 0.00010	0.04125 ± 0.00008
	GDMSR	0.01370 ± 0.00090	0.02400 ± 0.00070	0.02340 ± 0.00120	0.02600 ± 0.00090	0.03850 ± 0.00170	0.03160 ± 0.00120	0.06060 ± 0.00220	0.03990 ± 0.00150
	MGCN	0.01103 ± 0.00022	0.01938 ± 0.00030	0.01988 ± 0.00034	0.02146 ± 0.00036	0.03427 ± 0.00050	0.02701 ± 0.00046	0.05747 ± 0.00072	0.03575 ± 0.00058
	MGGCL	**0.01458** ± 0.00010	**0.02502** ± 0.00007	**0.02570** ± 0.00020	**0.02767** ± 0.00011	**0.04301** ± 0.00024	**0.03431** ± 0.00015	**0.06981** ± 0.00025	**0.04439** ± 0.00016

iFashion: iFashion is the sparsest dataset in our study. It combines a significantly large user base of 300,000 with only around 1.6 million interactions. With a user popularity Gini coefficient of 0.36, the user activity is relatively evenly spread. In contrast, a high item popularity Gini coefficient of 0.77 reveals extreme popularity skewness. Only a small fraction of fashion items dominate the total amount of purchases, while the vast majority of items see very little demand. Traditional methods struggle to separate signals from noise in such sparse interactions, whereas MGGCL's motif-guided view construction manages to surface local dense clusters, e.g., users co-purchasing the same style of products. Consequently, MGGCL outperforms the strongest baseline by 20.2% on Recall@40, which confirms its ability to amplify rare yet semantically rich co-interaction patterns in ultra-sparse settings.

Amazon-Kindle: Amazon-Kindle has the second lowest density among the four datasets, with moderate user and item popularity Gini coefficients around 0.54. This reflects a balance between devoted e-book enthusiasts and the broad distribution of interest across genres and series. In terms of experimental results, MGGCL outperforms all contrastive-learning baselines but is only marginally better than LightGCN, and the motif-augmented MGCN attains the best scores on this dataset, surpassing MGGCL across all metrics. A possible reason is that, when popularity skewness is moderate, the raw first-order adjacency used by LightGCN already captures a large share of the meaningful signal without being strongly affected by hub bias, so the benefit of hub-suppressing motif contrast is limited. In this regime, simply augmenting the adjacency with motif structure, as MGCN does, is already sufficient to capture the available signal, and the additional dual-view contrast in MGGCL provides little further benefit. MGGCL nonetheless remains the strongest contrastive-learning method on this dataset and stays close to MGCN, which is consistent with our broader finding that the contrastive mechanism is most valuable under strong popularity skew.

Amazon-Book: The Amazon-Book dataset consists of a large item catalogue of around 92,000 but fewer users, which produces a moderate sparsity level. Among the four datasets, the sparsity level of Amazon-Book is between Douban-Book and Amazon-Kindle. With relatively low Gini coefficients for both user and item popularities, the dataset shows the most uniform interaction patterns of the four. Despite this, previous research [50,51] shows that the dataset still exhibits considerable popularity bias. As a result, MGGCL shows smaller improvements than Douban-Book and iFashion but still outperforms the strongest baseline by 8% on Recall@40 and 7.6% on NDCG@40.

Summary: Across datasets of varying density and from heterogeneous domains, MGGCL delivers the best overall performance, leading on three of the four benchmarks and remaining competitive with the motif-augmented MGCN on the fourth. The comparison with MGCN, a motif-enhanced message-passing model without contrastive learning, indicates the value of the dual-view contrast: MGGCL improves clearly over motif-enhanced message passing on the skewed datasets, while on the near-uniform Amazon-Kindle motif augmentation alone already suffices. Its motif-guided view construction (i) exploits abundant higher-order structures in dense graphs, (ii) uncovers latent micro-clusters in ultra-sparse graphs, and (iii) alleviates popularity bias where star motifs dominate. These results underscore the general efficacy of explicitly modelling higher-order interaction semantics for robust recommendation.

5.5. Parameter Sensitivity Analysis

In this section, we investigate the impact of the contrastive learning coefficient λ and the motif ratio θ on MGGCL's performance.

5.5.1. Impact of Contrastive Learning Coefficient λ

MGGCL achieves the best performance on the four datasets with the following combinations of λ and θ: $\lambda = 0.2$ and $\theta = 1$ for Douban-Book, $\lambda = 0.05$ and $\theta = 1$ for iFashion, $\lambda = 0.1$ and $\theta = 1$ for Amazon-Kindle, and $\lambda = 1$ and $\theta = 0.5$ for Amazon-Book. By fixing θ at the corresponding best values for the datasets, we obtain the results illustrated in Figure 4.

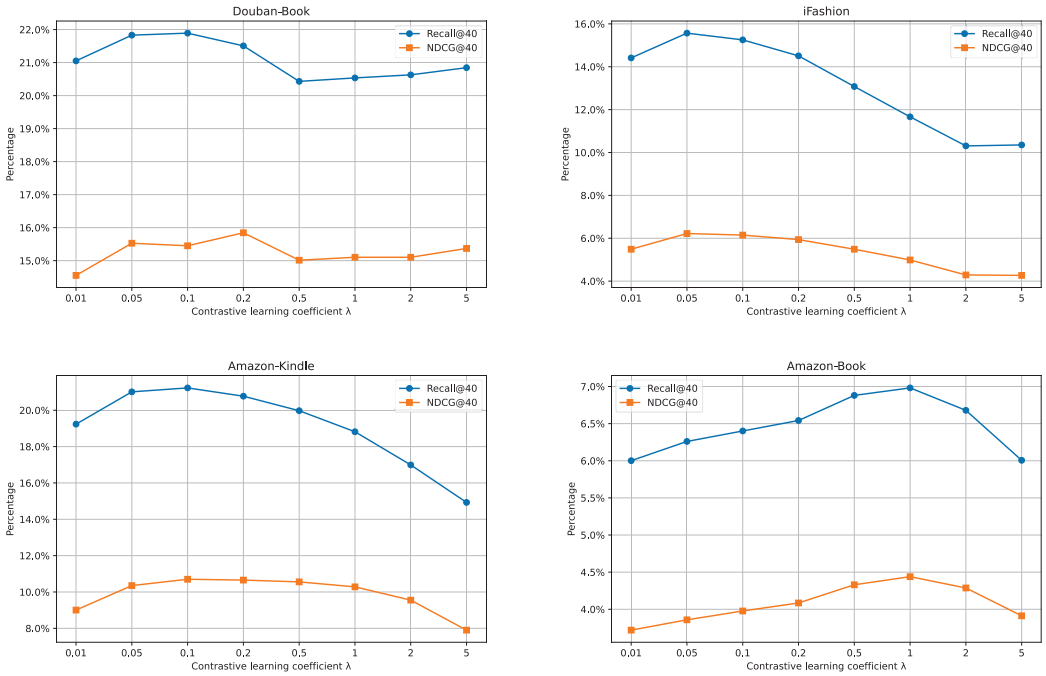

Figure 4. Impact of the contrastive learning coefficient λ.

We can see that the choice of contrastive learning coefficient λ significantly affects the model's performance. Specifically, for each dataset, we observe a clear peak at the optimal λ value, beyond which performance typically declines. For Douban-Book, iFashion, and Amazon-Kindle datasets, smaller λ values (0.05 to 0.2) yield better performance, suggesting that overly emphasising contrastive loss may reduce the model's capacity to leverage primary recommendation signals effectively. In contrast, for Amazon-Book, the best performance at a relatively higher $\lambda = 1$ indicates that stronger contrastive signals help to distinguish informative motifs from the noise inherent in sparse datasets.

5.5.2. Impact of Motif Ratio θ

Similarly, by fixing λ at the corresponding best values for the datasets, we obtain the results illustrated in Figure 5.

Figure 5 demonstrates the impact of the motif ratio θ, which balances the contribution between the original adjacency matrix and the motif adjacency matrices. The results show that setting $\theta = 1$ generally produces optimal outcomes for Douban-Book, iFashion, and Amazon-Kindle datasets, confirming that entirely motif-guided representations effectively capture meaningful structural semantics for these scenarios. For Amazon-Book, however, optimal performance at $\theta = 0.5$ indicates the importance of balancing motif-based information and the original graph structure, confirming the flexibility of MGGCL to adapt to varying levels of motif density across different datasets.

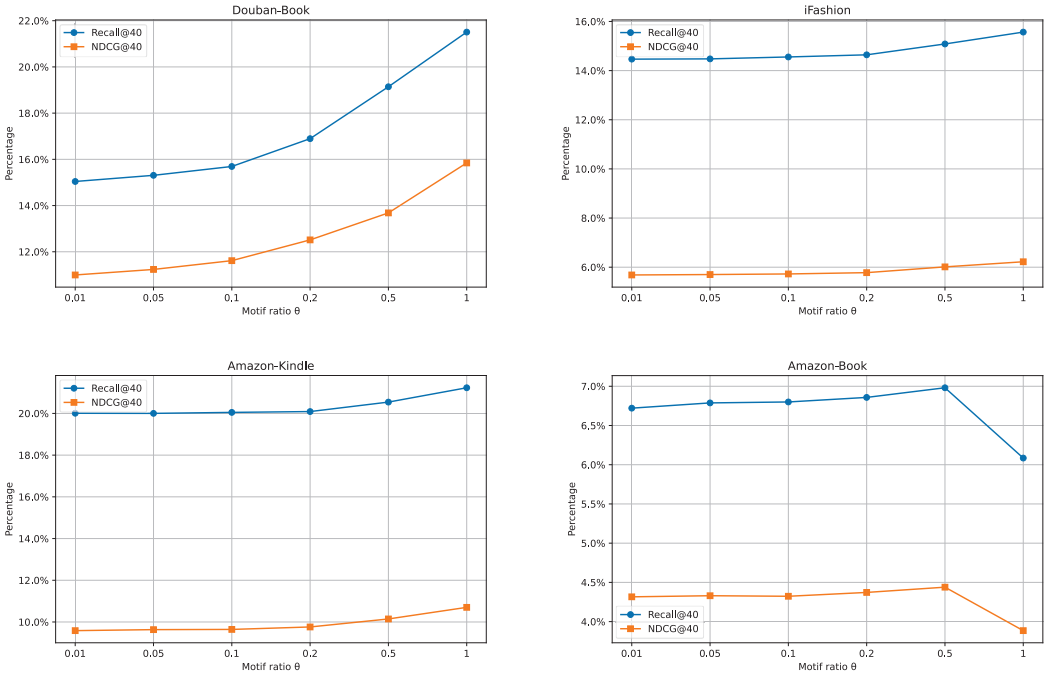

Figure 5. Impact of the motif ratio θ.

6. Conclusions and Future Work

In this paper, we introduced a novel motif-guided graph contrastive learning framework to explicitly integrate structural semantics into the construction of contrastive views for graph-based collaborative filtering. Unlike conventional contrastive learning techniques that rely on random perturbations or uniform edge-drop strategies, MGGCL strategically contrasts views dominated by dense rectangle motifs against sparse star motifs. This motif-aware approach allows our model to explicitly capture and differentiate meaningful and robust group behaviours from sparse and potentially noisy user–item interactions.

Extensive experiments conducted on four real-world datasets demonstrated that MGGCL achieves the best overall performance, outperforming state-of-the-art baselines on three of the four benchmarks and remaining competitive on the fourth. These results validate our analysis that explicitly modelling motifs helps distinguish critical interaction patterns and improve recommendation quality.

Our method and analysis are framed within the spatial message-passing GNN paradigm, where motifs define higher-order neighbourhoods and the mechanism is characterised in the vertex domain, as shown in Proposition 1 and Section 4.4. A natural next step is to complement this perspective with spectral and learning-theoretic analysis, characterising how motif-augmented operators affect the graph spectrum, connectivity, stability, conditioning, and generalisation behaviour of the contrastive objective. Such analysis would further clarify why the selected motifs balance local invariance and global discrimination. Since the motif-guided view construction is encoder-independent, combining it with more expressive backbones and extending it to larger-scale or cross-domain recommendation scenarios are also promising directions.

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

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are openly available in MGGCL-code at https://github.com/burghdun/MGGCL (accessed on 27 May 2026).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, X.; He, X.; Wang, M.; Feng, F.; Chua, T.S. Neural Graph Collaborative Filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 165–174. [CrossRef]
2. He, X.; Deng, K.; Wang, X.; Li, Y.; Zhang, Y.; Wang, M. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*; SIGIR '20; Association for Computing Machinery: New York, NY, USA, 2020; pp. 639–648. [CrossRef]
3. Yu, W.; Zhang, Z.; Qin, Z. Low-Pass Graph Convolutional Network for Recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*; AAAI Press: Washington, DC, USA, 2022.
4. Wu, J.; Wang, X.; Feng, F.; He, X.; Chen, L.; Lian, J.; Xie, X. Self-supervised Graph Learning for Recommendation. In P*roceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*; SIGIR '21; Association for Computing Machinery: New York, NY, USA, 2021; pp. 726–735. [CrossRef]
5. Zhou, K.; Wang, H.; Zhao, W.X.; Zhu, Y.; Wang, S.; Zhang, F.; Wang, Z.; Wen, J.R. S3-Rec: Self-Supervised Learning for Sequential Recommendation with Mutual Information Maximization. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*; CIKM '20; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1893–1902. [CrossRef]
6. Lin, Z.; Tian, C.; Hou, Y.; Zhao, W.X. Improving Graph Collaborative Filtering with Neighborhood-enriched Contrastive Learning. In *Proceedings of the ACM Web Conference 2022*; WWW '22; Association for Computing Machinery: New York, NY, USA, 2022; pp. 2320–2329. [CrossRef]
7. Yu, J.; Yin, H.; Xia, X.; Chen, T.; Cui, L.; Nguyen, Q.V.H. Are Graph Augmentations Necessary? Simple Graph Contrastive Learning for Recommendation. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*; SIGIR '22; Association for Computing Machinery: New York, NY, USA, 2022; pp. 1294–1303. [CrossRef]
8. Radford, A.; Kim, J.W.; Hallacy, C.; Ramesh, A.; Goh, G.; Agarwal, S.; Sastry, G.; Askell, A.; Mishkin, P.; Clark, J.; et al. Learning Transferable Visual Models From Natural Language Supervision. In *Proceedings of the 38th International Conference on Machine Learning*; PMLR; JMLR: Cambridge, MA, USA, 2021; pp. 8748–8763.
9. De Martino, G.; Marra, P.; D'Aversa, A.; Pulito, L.; Pellicani, A.; Pio, G.; Ceci, M. Leveraging textual content, citational aspects and dissenting opinions through a multi-view contrastive learning methodology for legal precedent analysis. *Comput. Law Secur. Rev.* **2026**, *60*, 106257. [CrossRef]
10. Peng, Y.; Wu, J.; Sun, Y.; Zhang, Y.; Wang, Q.; Shao, S. Contrastive-learning of language embedding and biological features for cross modality encoding and effector prediction. *Nat. Commun.* **2025**, *16*, 1299. [CrossRef] [PubMed]
11. Liang, H.; Du, X.; Zhu, B.; Ma, Z.; Chen, K.; Gao, J. Graph contrastive learning with implicit augmentations. *Neural Netw.* **2023**, *163*, 156–164. [CrossRef] [PubMed]
12. Trivedi, P.; Lubana, E.S.; Yan, Y.; Yang, Y.; Koutra, D. Augmentations in Graph Contrastive Learning: Current Methodological Flaws & Towards Better Practices. In *Proceedings of the ACM Web Conference 2022*; ACM: New York, NY, USA, 2022; pp. 1538–1549. [CrossRef]
13. Wei, C.; Wang, Y.; Bai, B.; Ni, K.; Brady, D.; Fang, L. Boosting Graph Contrastive Learning via Graph Contrastive Saliency. In *Proceedings of the 40th International Conference on Machine Learning*; PMLR; JMLR: Cambridge, MA, USA, 2023; pp. 36839–36855.
14. Liu, Y.; Kertkeidkachorn, N.; Miyazaki, J.; Ichise, R. Review-enhanced contrastive learning on knowledge graphs for recommendation. *Expert Syst. Appl.* **2025**, *277*, 127250. [CrossRef]
15. Zhao, C.; Yang, E.; Liang, Y.; Zhao, J.; Guo, G.; Wang, X. Symmetric Graph Contrastive Learning against Noisy Views for Recommendation. *ACM Trans. Inf. Syst.* **2025**, *43*, 80:1–80:28. [CrossRef]
16. Zhang, Y.; Zhu, J.; Zhang, Y.; Zhu, Y.; Zhou, J.; Xie, Y. Social-aware graph contrastive learning for recommender systems. *Appl. Soft Comput.* **2024**, *158*, 111558. [CrossRef]
17. Jiang, W.; Gao, X.; Xu, G.; Chen, T.; Yin, H. Challenging Low Homophily in Social Recommendation. In *Proceedings of the ACM on Web Conference 2024*; WWW '24; Association for Computing Machinery: New York, NY, USA, 2024; pp. 3476–3484. [CrossRef]
18. Xue, J.; Yang, Z.; Lin, H.; Zhang, Z.; Wang, L.; Gu, Y.; Xu, Y.; Li, X. HGCL: Hierarchical Graph Contrastive Learning for User-Item Recommendation. *arXiv* **2025**, arXiv:2505.19020. [CrossRef]
19. Wei, H.; Wang, J.; Ji, Y.; Guang, M.; Yan, C. Hierarchical neighbor-enhanced graph contrastive learning for recommendation. *Knowl.-Based Syst.* **2025**, *315*, 113263. [CrossRef]

20. Xu, J.; Chen, Z.; Yang, S.; Li, J.; Wang, H.; Wang, W.; Hu, X.; Ngai, E. NLGCL: Naturally Existing Neighbor Layers Graph Contrastive Learning for Recommendation. In *Proceedings of the 19th ACM Conference on Recommender Systems*; ACM: New York, NY, USA, 2025. [CrossRef]
21. Zhuo, J.; Lu, Y.; Ning, H.; Fu, K.; Niu, B.; He, D.; Wang, C.; Guo, Y.; Wang, Z.; Cao, X.; et al. Unified Graph Augmentations for Generalized Contrastive Learning on Graphs. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 37473–37503. [CrossRef]
22. Ni, X.; Xiong, F.; Zheng, Y.; Wang, L. Graph Contrastive Learning with Kernel Dependence Maximization for Social Recommendation. In *Proceedings of the ACM Web Conference 2024. Association for Computing Machinery*; WWW '24; ACM: New York, NY, USA, 2024; pp. 481–492. [CrossRef]
23. Zhou, X.; Sun, A.; Liu, Y.; Zhang, J.; Miao, C. SelfCF: A Simple Framework for Self-supervised Collaborative Filtering. *ACM Trans. Recomm. Syst.* **2023**, *1*, 9:1–9:25. [CrossRef]
24. Yu, P.; Bao, B.K.; Tan, Z.; Lu, G. Improving Graph Collaborative Filtering with Directional Behavior Enhanced Contrastive Learning. *ACM Trans. Knowl. Discov. Data* **2024**, *18*, 183:1–183:20. [CrossRef]
25. Zhang, D.; Geng, Y.; Gong, W.; Qi, Z.; Chen, Z.; Tang, X.; Shan, Y.; Dong, Y.; Tang, J. RecDCL: Dual Contrastive Learning for Recommendation. In *Proceedings of the ACM Web Conference 2024*; WWW '24; Association for Computing Machinery: New York, NY, USA, 2024; pp. 3655–3666. [CrossRef]
26. Zhao, R.; Chen, L.; Sun, S.; Peng, J.; Ju, S. Data Collaborative Contrastive Recommendation model with self-adaptive noise. *Expert Syst. Appl.* **2024**, *256*, 124899. [CrossRef]
27. Liu, X.; Wen, G.; Wang, A.; Liu, C.; Wei, W.; Meo, P.D. GGDHSCL: A Graph Generative Diffusion With Hard Negative Sampling Contrastive Learning Recommendation Method. *IEEE Trans. Comput. Soc. Syst.* **2025**, *12*, 2784–2799. [CrossRef]
28. Chen, Z.; Xu, J.; Wei, Y.; Peng, Z. Squeeze and Excitation: A Weighted Graph Contrastive Learning for Collaborative Filtering. In *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*; SIGIR '25; Association for Computing Machinery: New York, NY, USA, 2025; pp. 2769–2773. [CrossRef]
29. Zhang, C.; Han, Q.; Tan, Q.; Wang, S.; Zhao, X.; Chen, R. DimCL: Dimension-Aware Augmentation in Contrastive Learning for Recommendation. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.1*; KDD '25; Association for Computing Machinery: New York, NY, USA, 2025; pp. 1913–1923. [CrossRef]
30. Wang, S.; Zhou, Y.; Fan, X.; Li, J.; Lei, Z.; Gong, M. Personalized Federated Contrastive Learning for Recommendation. *IEEE Trans. Comput. Soc. Syst.* **2025**, *12*, 2986–2998. [CrossRef]
31. Dai, J.; Li, Q.; Nong, T.; Bi, Q.; Chu, H. Joint contrastive learning of structural and semantic for graph collaborative filtering. *Neurocomputing* **2024**, *586*, 127547. [CrossRef]
32. Li, X.; Tian, Y.; Dong, B.; Ji, S. MD-GCCF: Multi-view deep graph contrastive learning for collaborative filtering. *Neurocomputing* **2024**, *590*, 127756. [CrossRef]
33. Liu, X.; Hao, Y.; Zhao, L.; Liu, G.; Sheng, V.S.; Zhao, P. LMACL: Improving Graph Collaborative Filtering with Learnable Model Augmentation Contrastive Learning. *ACM Trans. Knowl. Discov. Data* **2024**, *18*, 1–24. [CrossRef]
34. Li, C.; Cheng, D.; Zhang, G.; Zhang, S. Contrastive learning for fair graph representations via counterfactual graph augmentation. *Knowl.-Based Syst.* **2024**, *305*, 112635. [CrossRef]
35. Zhang, Y.; Yu, J.; Liu, Z.; Wang, G.; Nguyen, M.; Sheng, Q.Z.; Wang, N. Improving graph collaborative filtering with network motifs. *Neural Comput. Appl.* **2025**, *37*, 9413–9432. [CrossRef]
36. Wang, G.; Yu, J.; Nguyen, M.; Zhang, Y.; Yongchareon, S.; Han, Y. Motif-based graph attentional neural network for web service recommendation. *Knowl.-Based Syst.* **2023**, *269*, 110512. [CrossRef]
37. Sun, Y.; Zhu, D.; Du, H.; Tian, Z. Motifs-based recommender system via hypergraph convolution and contrastive learning. *Neurocomputing* **2022**, *512*, 323–338. [CrossRef]
38. Cui, Z.; Cai, Y.; Wu, S.; Ma, X.; Wang, L. Motif-aware Sequential Recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*; SIGIR '21; Association for Computing Machinery: New York, NY, USA, 2021; pp. 1738–1742. [CrossRef]
39. Zhang, S.; Hu, Z.; Subramonian, A.; Sun, Y. Motif-Driven Contrastive Learning of Graph Representations. *IEEE Trans. Knowl. Data Eng.* **2024**, *36*, 4063–4075. [CrossRef]
40. Wu, X.; Wang, C.D.; Lin, J.Q.; Xi, W.D.; Yu, P.S. Motif-Based Contrastive Learning for Community Detection. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 11706–11719. [CrossRef] [PubMed]
41. Li, C.; Guo, Y.; Tang, Z.; Li, W.; Chen, Y.; Cao, J.; Zhang, Y.; Jiang, J. MHGCL: Combining Motif and Homogeneity in graph contrastive learning. *World Wide Web* **2025**, *28*, 47. [CrossRef]
42. Liu, Y.; Jin, M.; Pan, S.; Zhou, C.; Zheng, Y.; Xia, F.; Yu, P.S. Graph Self-Supervised Learning: A Survey. *IEEE Trans. Knowl. Data Eng.* **2023**, *35*, 5879–5900. [CrossRef]
43. Hamilton, W.L.; Ying, R.; Leskovec, J. Representation Learning on Graphs: Methods and Applications. *IEEE Data Eng. Bull.* **2017**, *40*, 52–74.
44. Diestel, R. *Graph Theory*, 5th ed.; *Graduate Texts in Mathematics*; Springer: New York, NY, USA, 2017; Volume 173.

45. Milo, R.; Shen-Orr, S.; Itzkovitz, S.; Kashtan, N.; Chklovskii, D.; Alon, U. Network Motifs: Simple Building Blocks of Complex Networks. *Science* **2002**, *298*, 824–827. [CrossRef] [PubMed]
46. Benson, A.R.; Gleich, D.F.; Leskovec, J. Higher-order organization of complex networks. *Science* **2016**, *353*, 163–166. [CrossRef] [PubMed]
47. Yu, J.; Xia, X.; Chen, T.; Cui, L.; Hung, N.Q.V.; Yin, H. XSimGCL: Towards Extremely Simple Graph Contrastive Learning for Recommendation. *IEEE Trans. Knowl. Data Eng.* **2023**, *36*, 913–926. [CrossRef]
48. Abdollahpouri, H.; Mansoury, M.; Burke, R.; Mobasher, B.; Malthouse, E. User-centered Evaluation of Popularity Bias in Recommender Systems. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*; ACM: New York, NY, USA, 2021; pp. 119–129. [CrossRef]
49. Zhang, X.; Deng, X.; Yuan, H.; Wei, C.; Fan, Y.; Zhang, J. Graph-Based Diffusion Model for Service Recommendation. *IEEE Trans. Serv. Comput.* **2026**, *19*, 396–409. [CrossRef]
50. Shi, T.; Si, Z.; Xu, J.; Zhang, X.; Zang, X.; Zheng, K.; Leng, D.; Niu, Y.; Song, Y. UniSAR: Modeling User Transition Behaviors between Search and Recommendation. In Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval; SIGIR '24; Association for Computing Machinery: New York, NY, USA, 2024; pp. 1029–1039. [CrossRef]
51. Xu, W.; Gao, F.; Su, J.; Wu, Q.; Wang, M. Dual-view Enhanced Knowledge Contrastive Learning for Recommendation. In *Proceedings of the Database Systems for Advanced Applications*; Springer: Singapore, 2024; pp. 528–538. [CrossRef]