



Attributed Graph Convolutional Network for Enhanced Social Recommendation Through Hybrid Feedback Integration

Xiaoyi Deng 

Business School, Huaqiao University, Quanzhou 362021, China

Corresponding Author Email: londonbell@hqu.edu.cn

<https://doi.org/10.18280/ts.400509>

ABSTRACT

Received: 15 August 2023
Revised: 8 October 2023
Accepted: 24 October 2023
Available online: 30 October 2023

Keywords:

social recommendation, attribute information, variational graph embedding, graph convolution network (GCN), representation learning

Social recommendation, a technique aimed at predicting user preferences by harnessing social ties, has frequently employed collaborative filtering (CF) due to its demonstrated efficiency and scalability. Nonetheless, a decline in performance of most extant CF techniques has been observed when confronted with extreme sparsity in explicit feedback. Past investigations predominantly merged both explicit and implicit feedback to mitigate the data scarcity issue, embedding based solely on explicit characteristics and formulating objective functions founded on user-item associations. Such a paradigm signifies a dependency on these interactions to compensate for deficient embeddings. Notably, a considerable discrepancy exists between implicit feedback and genuine user satisfaction in social recommendations, attributed to pervasive false positive interactions devoid of detailed user/item attributes. Furthermore, the establishment of connectivity between users/items has been partially dependent on users' inclinations, suggesting that the aggregation procedure might overlook certain neighbourhood preferences. In response to these challenges, a hybrid neural graph model endowed with attributive features has been introduced. This model amalgamates explicit/implicit feedback, attribute data, and a user-item interaction graph. To counteract data sparsity, a variational graph framework has been devised to extract latent representations from both feedback and attribute data. For the effective and explicit discernment of collaborative signals, the embedding incorporates a user-item interaction graph, which offers a potent modelling of elevated-order connectivities and the detection of latent user-item associations. The user and item embeddings are derived via an attentive propagation method, with the ultimate item embeddings being sourced through a linear weighted sum, eschewing non-linear activation functions. Comparative analyses on four real-world datasets have demonstrated the superior efficacy of the proposed methodology in relation to leading contemporary recommendation systems.

1. INTRODUCTION

Recommendation systems have been recognized as instrumental tools in addressing the challenges of information overload, guiding users towards valuable insights and predicting preferences across an array of items in alignment with individual tastes. In the evolving landscape of social commerce, a profusion of social connections, such as friendships, is generated within social networks. These connections are believed to significantly augment social engagements amongst individuals, whether they be acquaintances, colleagues, or even strangers [1]. In this digital realm, avenues for communication and the exchange of ideas are presented, fostering rich social interactions. Recommendations, derived from these social interactions, underscore the foundational importance of these social ties. An enhancement in the potency of social recommendations is, therefore, postulated to be closely tethered to these social connections. The principle of social correlation posits that within a given network, individuals are predisposed to manifest similar preferences or exhibit mutual influences, culminating in analogous choices within their interconnected web [1]. CF, which predicates user preferences on the notion that individuals with converging tastes are inclined towards

analogous items, emerges as a quintessential methodology in the realm of social recommendation [2]. For the actualization of this principle, user-item interactions are reconstituted through a process that necessitates the parameterization of both users and items. Conventionally, most CF models can be demarcated into two predominant facets: embedding and interaction modelling. In initial methodologies, entities were projected into a shared latent space, and hidden vectors were employed to represent either entity. Subsequently, the emphasis shifted towards reconstructing user-item interactions leveraging these embeddings, as evidenced by techniques such as matrix factorization (MF) [3]. Advances like the collaborative topic regression (CTR) [4] amalgamated with deep feature representation learning further refined these interactions. A paradigm shift was observed with the advent of Neural Collaborative Filtering (NCF) and its derivatives [5], which merged linear MF and nonlinear neural networks, forsaking the traditional MF interaction function.

While the aforementioned methodologies have demonstrated efficacy, inherent limitations emerge when confronted with data paucity, inhibiting the generation of optimal embeddings for CF. In response to the challenges posed by data scarcity, attribute data has been integrated into traditional CF methods by several researchers [6]. To more

accurately distil the essential elements from attribute data, techniques such as latent Dirichlet allocation, Bayesian personalized ranking (BPR), and autoencoders, among others, have been employed in previous studies [7]. However, a predominant reliance on inner products to emulate interactions between users and items has been identified, restricting these methodologies' capability to delineate non-linear relationships [8]. In a bid to rectify this, various approaches, inclusive of NCF, DeepFM [9], and Neural Factorization Machine [10], have harnessed deep neural networks, resulting in enhanced modelling of non-linear interactions. Notwithstanding their successes, the intricate ambiguities of latent representations pertaining to users and items have not been entirely captured by these profound neural architectures. Recently, the adoption of deep generative models, such as the Variational Autoencoder (VAE) [11], in CF tasks has been noted. Characterized by its probabilistic nature, VAE can encapsulate uncertainty, facilitating the exploration of non-linear probabilistic latent-variable models on extensive recommendation datasets, as evidenced by models like collaborative variational autoencoder (CVAE) [12] and VAE-based CF (VAECF) [13]. Nonetheless, certain caveats accompany these VAE-based CF methods. For instance, the dependency of CVAE on inner products for interaction representation has been observed to constrict its scope in capturing intricate non-linear interactions between users and items. Conversely, VAECF's exclusive reliance on rating matrices for modelling user behaviors and prediction generation is found to be suboptimal on exceedingly sparse rating matrices, further compromising its ability to furnish new users with accurate recommendations [14].

In the quest to ameliorate the shortcomings of insufficient embeddings, recent analyses have been undertaken [7, 15]. It was discerned that a preponderance of studies formulated embedding functions without the integration of implicit characteristics. As a result, the capability to efficaciously seize the fundamental collaborative signals, indicative of user-item interactions, was compromised. Furthermore, the employment of user-item interactions was discerned to be predominantly constrained to delineating training objectives, leading to the insufficiency in embeddings' capability to capture pertinent signals [8]. Such an observed disconnect between embedding and interaction modelling suggests that latent representations of users and items may not always be accurately discerned. Predominantly, information pertaining to users in social recommendation systems is gleaned from both social interactions and direct user-item interactions. However, this multifaceted information is not always holistically leveraged, often due to a lack of integrated analysis from diverse vantage points [1]. It should be emphasized that interactions between users and items not only denote direct exchanges but also encapsulate users' intrinsic item preferences. The nature of interactions within social networks is subject to the strength of user connections; stronger ties often manifest in more aligned preferences compared to weaker connections. A uniform treatment of these varied social interactions might inadvertently attenuate the efficacy of recommendation systems [16]. Hence, there emerges an intrinsic imperative to distinctly categorize these varied social ties, concurrently assimilating insights about both interactions and preferences. This comprehensive approach would involve amalgamating user information sourced from both social dynamics and direct user-item engagements, ensuring the embedding function is optimally primed to extrapolate the requisite collaborative

signals for enriched representation.

To address the outlined challenges, a hybrid attributive GCN structure tailored for recommendation in the realm of social e-commerce has been proposed. Within this structure, explicit collaborative signals in the user-item interaction graph are discerned, leveraging two supplementary VAEs to derive non-linear latent representations and establish high-order connectivity via attentive embedding propagation. A unified neural variational model has been employed to encapsulate the generative processes of both users and items, thereby facilitating the efficient extraction of non-linear latent representations for CF. By embedding the attribute information of users and items into their latent factors through a deep graph neural networks (GNN) for CF, data sparsity issues can be effectively addressed and the latent representations of users and items can be enhanced. Drawing from findings in GNN research [17, 18], an embedding propagation layer has been integrated to augment user/item embeddings. This augmentation is achieved by aggregating embeddings from interacting items or users. Sequentially, these embeddings are optimized to discern cooperative indications in higher-order connectivity, achieved through layer stacking for embedding propagation. Through data-driven training, the adopted neural network has demonstrated the capacity to incorporate both user preferences and item attributes into the latent factors of users and items. This manuscript is structured as follows: Section 2 offers a review of extant literature on CF models. In Section 3, the proposed models are elucidated, accompanied by a discussion on parameter learning methodologies. Section 4 presents the empirical outcomes and associated discussions, with Section 5 concluding the paper and suggesting avenues for future research.

2. RELATED WORKS

2.1 CF

Recommendation systems have often been crafted utilizing CF, a method widely acknowledged for its ubiquity in the domain. In such processes, data pertaining to users' actions, behaviors, or preferences are collected and scrutinized. The aim is to predict user preferences through juxtaposition with other users. Three primary paradigms emerge: memory-based CF, model-based CF, and hybrid CF models.

Memory-based models bifurcate into user-based and item-based categories. In these, CFs have been observed to hinge on ratings and the discernment of closely correlated neighbors. Such neighbors, identified as those manifesting analogous rating patterns to a given user, serve as sources of recommendations rooted in items previously endorsed by these similar entities. As per Shi et al. [2], a greater precision in user-based CF is often achieved when an expansive data pool is employed for similarity computations. However, challenges arise in scalability due to user volume and data sparsity. In contrast, the scalability of item-based CF is noted to be commendable, even with voluminous, sparse data, albeit potentially at the cost of precision.

Model-based CF, diverging from memory-based models, offers an alternate paradigm that does not predominantly lean on user or item similarity. Here, techniques such as matrix factorization have been employed, constructing predictive models rooted in user preferences, leading to notable boosts in

both accuracy and scalability. Koren et al. [3] have asserted the prevalence of matrix factorization techniques, underscoring their foundational premise: a minimal set of concealed factors often suffices to elucidate extant ratings and to prognosticate absent ones. Whilst model-based CF has been shown to augment prediction efficacy and offer insightful rationales behind recommendations, a significant loss of information is often incurred due to dimensionality reduction.

Hybrid models aim to augment recommendation precision by amalgamating CF with diverse methodologies including Bayesian hierarchical models, clustering, neural networks, and knowledge-based techniques [1]. As noted by He and Chua [10], such models have demonstrated enhanced precision in suggestions, adeptly addressing issues like data sparsity and the cold-start problem. Moreover, they present explicative recommendation rationales. Yet, complexities in implementation and the recurrent unattainability of external information requisites often overshadow their potential benefits.

Over recent decades, deep learning has experienced significant advancements in model-based CF. Restricted Boltzmann machines and deep belief networks have been applied to discern a probabilistic perspective of user-item interaction in rating contexts [11]. Other state-of-the-art models, including variations of VAEs, generative adversarial networks, and the like, have been explored due to their inherent capacity to embrace both uncertainty and non-linearity [13]. In such endeavors, deep neural networks have been employed, envisaging recommendation as either a classification or regression challenge, with the intent to autonomously discern obscured affiliations between users and items.

2.2 GNN for recommendation systems

GNNs, especially in recent developments, have been demonstrated to hold remarkable efficacy in processing graph data across various fields [19]. Their innate capacity to seamlessly handle typical graph data makes them particularly apt for the realm of social recommendation. GNNs are often perceived in certain recommendation models as instrumental tools for feature extraction, thereby harnessing additional attributes from social data [20, 21]. Central to this approach is the propensity to employ embedding propagation, thereby facilitating the amalgamation of neighborhood embeddings. Such a mechanism ultimately grants each node the capability to assimilate information from higher-order neighbors. Both user-user social graphs and user-item interaction graphs have been leveraged in social recommendations, revealing inherent benefits in the realm of embedding learning.

Recent investigations, as highlighted by Gao et al. [19], have denoted an inclination towards amalgamating neighbouring embeddings to refine the embeddings of target nodes within the spatial domain. Contemporary research has pivoted towards the deployment of GCN for the meticulous management of user-item interaction graphs. Such endeavors are underscored by the pursuit of unveiling CF signals within high-order neighbours, steered by the GCN's interpretability and efficiency. An exemplar of this approach is the graph convolutional matrix completion (GCMC) [22], wherein both the CF signal and the inherent graph structure are harnessed to curate embeddings. Through an effective synthesis of information stemming from the structural and attribute components of the user-item interaction graph, enhanced

accuracy is attained. PinSage [20] adopts a distinct methodology, employing random walks to sculpt bespoke item embeddings. This method meticulously captures adjacent graph configurations, subsequently consolidating the preferences of akin users. In a fusion of GCN with CF, neural graph CF (NGCF) [23] adeptly discerns intricate interconnections existing between users and items. By accruing embeddings from neighboring nodes, latent user preferences are capably grasped, with performance metrics outstripping conventional methodologies. To enhance the proficiency of GCN, attention mechanisms have been employed, as illustrated by the works of GAT [24] and KGAT [25]. In the former, varying levels of significance are assigned to interactions between users and items, enabling the capture of intricate associations within the knowledge graph. Conversely, in the latter, weights are allocated to proximate nodes, with a distinct focus on prioritizing nodes deemed more informative during the information propagation phase.

Recent advancements in the understanding of GCN have been observed, and based on the comprehensive survey by Gao et al. [19], the evolution of GCN-based recommendation methodologies can be categorized into three primary domains. Firstly, emphasis has been placed on the deployment of simplified single-layer GCN or streamlined message passing to augment efficiency without compromising performance. The second trend involves the integration of GCN models with supplementary modules, notably attention mechanisms or gating functions, thereby elevating the efficacy of representation learning. Finally, advancements have been witnessed in the fusion of GCN with auxiliary datasets like knowledge graphs and various side information, including attributes, trust connections, and social regularization, to refine user profiling. Among these evolving trends, the emphasis on designing streamlined GCN has garnered significant attention in recent years. For instance, the nexus between GCN and Page Rank has been explored in personalized propagation of neural predictions (PPNP) [26], which yielded an advanced propagation technique. This methodology was further employed to construct a streamlined GCN model, demonstrating compatibility with diverse neural network architectures. In parallel, it has been asserted by Wu et al. [17] in their introduction of the simplifying GCN (SGCN) that the conventional GCN poses unnecessary complexities. They advocate for a simplified design, achieved by omitting non-linear functions and amalgamating multiple weight matrices for successive layers. LightGCN [18], on the other hand, was developed to address the excessive smoothing issue intrinsic to GCN for collaborative filtering scenarios. By centering attention solely on the adjacency matrix of the user-item interaction graph and employing numerous graph convolution layers with a direct linear aggregation strategy, Light GCN illuminates the pitfalls associated with non-linearity and redundant weight matrices in the collaborative filtering paradigm. Corroborating this perspective, other academic pursuits [27-29] have also postulated that the exclusion of non-linear components can bolster recommendation outcomes. A distinct linear residual network architecture has been delineated specifically for collaborative filtering undertakings involving user-item interactions. This innovative structure efficaciously rectifies the over-smoothing quagmire encountered during the graph convolution aggregation phase, particularly in light of interaction matrix sparsity.

While numerous studies have endeavored to enrich GCN-

based recommendation, the pervasive influence of user and item attributes is often overlooked. The amalgamation of solely the interaction graph within GCN can inadvertently lead to over-smoothing and a dilution of semantic significance, issues potentially rectifiable through the inclusion of attribute and graph structures [19]. Attributes can proffer valuable side information, aiding in refining user profiles and decoding item attributes. Moreover, in scenarios typified by cold-start users/items, attribute data can be instrumental in offsetting the dearth of interaction history. Hence, the exploitation of user and item attributes in concert with GCN is indispensable.

3. METHODOLOGY

The forthcoming section elucidates the architecture of the

proposed Hybrid Attributive GCN (HAGCN), as visually represented in Figure 1. This model comprises three pivotal components: the initial feature extraction and embedding, an attentive GCN fortified with a dual attention mechanism, and the subsequent rating prediction. Initially, two ancillary VAEs are employed to discern users' and items' characteristics through a coherent deep generative framework. Nodes' embeddings are initialized employing both ID and feature embeddings. Thereafter, the node embeddings undergo enhancement by a refined GCN, incorporating dual attentive propagation mechanisms. Such a mechanism facilitates the assimilation of collaborative signals through attentive embedding propagation atop the user-item relation graph. The culminating stage involves generating predictions via the inner product of users' and items' final representations.

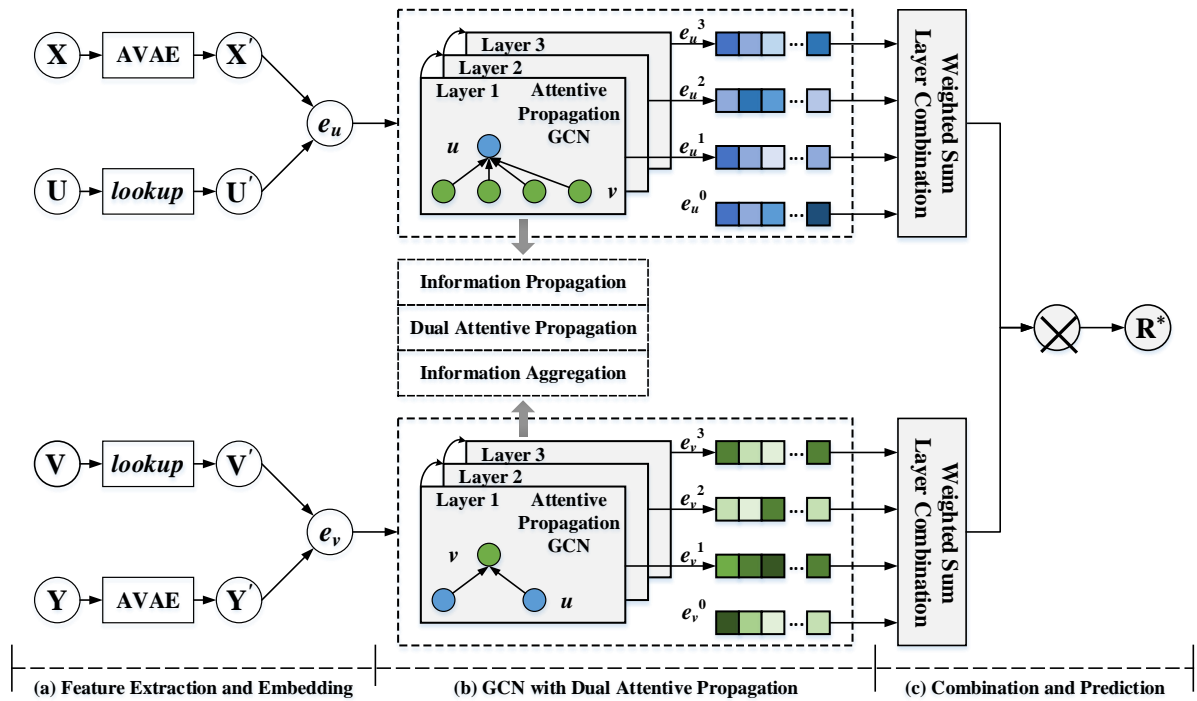


Figure 1. The architecture of the proposed HAGCN model

Table 1. Symbols and notations used in this paper

Symbols/Notations	Description
U, V	The sets of users and items
$R, R^* \in \mathbb{R}^{M \times N}$	The rating matrix and prediction matrix
$R_u \in \mathbb{R}^{K \times M}, R_v \in \mathbb{R}^{K \times N}$	The rating vectors of users and items
$X_u, X_u' \in \mathbb{R}^{P \times M}$	The attribute information and feature vectors of users
$X_v, X_v' \in \mathbb{R}^{Q \times N}$	The attribute information and feature vectors of items
$e_u, e_v \in \mathbb{R}^D$	The full embedding vectors of user and item
$e_u^{(l)}, e_v^{(l)} \in \mathbb{R}^D$	The l -th-hop embedding representations of user and item
M, N	The number of users and items
K, D	The dimensionality of latent space for users/items and node embedding
P, Q	The dimensionality of latent space for users' and items' attribute information

3.1 Notations

For clarity, a compendium of employed symbols and notations is presented, with comprehensive details delineated in Table 1.

The sets $U = \{u_i | i=1, \dots, M\} \in \mathbb{R}^{K \times M}$ and $V = \{v_j | j=1, \dots, N\} \in \mathbb{R}^{K \times N}$ are posited to symbolize the latent factors of users and items, respectively, with K representing the latent factors'

dimensions. In scenarios involving implicit feedback, both the rating matrix and its predictive counterpart are denominated as $R \in \mathbb{R}^{M \times N}$ and $R^* \in \mathbb{R}^{M \times N}$, respectively. Within this framework, $R_{ij} = 1$ is indicative of engagement between the i -th user and the j -th item, whereas $R_{ij} = 0$ denotes a lack thereof. Concurrently, rating vectors for users and items are represented as $R_u \in \mathbb{R}^{K \times M}$ and $R_v \in \mathbb{R}^{K \times N}$, respectively.

Vectors $X_u = \{X_{u_i} | i=1, \dots, M\} \in \mathbb{R}^{P \times M}$ and $X_v = \{X_{v_j} | j=1, \dots,$

$N\}$ $\in \mathbb{R}^{Q \times M}$ encapsulate attribute data pertinent to users and items, respectively. Dimensions of these attributes are demarcated by P for users and Q for items. Within this context, terms latent profile representation and latent content representation correspond to X_u and X_v , respectively. Embedding representations for users and items after l propagation steps are delineated as $e_u \in \mathbb{R}^D$ and $e_v \in \mathbb{R}^D$, respectively. The overarching endeavour is to deduce the latent factors u_i and v_j for users and items, predicated upon the extant variables R , U , V , X_u , and X_v , with the terminal objective being the prediction of hitherto uncharted ratings, R^* .

3.2 Feature extraction and node embedding

Traditionally, most model-based CF methodologies are dependent exclusively on user-item interactions for rating predictions. However, when certain techniques are employed to amalgamate user or item attribute data into rating prediction through linear regression, an observed limitation in accuracy emerges. In addressing this constraint, attribute information pertaining to both users and items is integrated into the feature learning process in the proposed model. This integration is surmised to enhance the precision in discerning latent factors germane to users and items. Subsequently, the ID and feature embeddings pertaining to users and items are amalgamated into two distinct embedding lookup tables. Utilizing the features of previously unobserved nodes, embeddings for such nodes are generated, a feat unattainable with conventional ID embedding. Furthermore, this method contributes to a reduction in the count of learnable parameters.

3.2.1 Generative model

In the pursuit of deriving robust user and item features, a dual modeling structure fortified with two supplementary VAEs is devised. This structure is primed to incorporate user profile and item content information, along with associated tag data. By this approach, the effective derivation of concealed user and item representations becomes feasible. The generative process espoused by the proposed model is analogical to the deep latent Gaussian model paradigm. Tag information for user profiles and item contents are symbolized by $S \in \mathbb{R}^{N \times S}$ and $T \in \mathbb{R}^{N \times T}$ respectively, being represented as binary matrices. The condition $S_{is} = T_{jt} = 1$ connotes the association of the s -th tag with user u_i and the t -th tag with item v_j . Conversely, $S_{is} = T_{jt} = 0$ denotes a void in association. For each user u_i , a K -dimensional latent representation $z_u \sim N(0, \mathbb{I}^K)$ is sampled from a standard Gaussian prior. Post this sampling, a conditional sample variable $X \sim p_\theta(X|z_u)$ is generated via a decoder, where θ represents the generative parameter. Depending on data type, $p_\theta(X|z_u)$ might originate from a multivariate Bernoulli distribution (binary data) or a Gaussian distribution (real-valued data).

The latent representation $z_u \sim N(0, \mathbb{I}^K)$ is drawn from a Gaussian prior distribution with an identity covariance matrix. The latent representation of user u_i amalgamates the latent user offset with the latent user profile vector: $u_i = \varepsilon_i + z_{ui}$. The methodology employed for generating item content mirrors that of the user profile. Thus, the latent representation of item v_j is the fusion of the latent item offset and the latent item content vector: $v_j = \varepsilon_j + z_{vj}$.

3.2.2 Inference model

The inference framework encompasses an encoder network that bears resemblance to the generative model. Within this

construct, the generative network prescribes the posterior distribution $p_\theta(z_u|X)$, which, due to its intractable nature, necessitates approximation during the inference phase for each user. By deploying the SGVB estimator, the posterior of the latent user profile variable z_u can be effectively approximated through a tractable variational distribution $q_\phi(z_u|X)$. The variational parameters are subsequently derived by optimizing the evidence lower bound (ELBO) objective:

$$q_\phi(z_u | X_i) = N(\mu_\phi(X)_i, \text{diag}(\sigma_\phi^2(X_i))) \quad (1)$$

where, the mean and standard deviation of the approximate posterior are denoted as $\mu_\phi \in \mathbb{R}^K$ and $\sigma_\phi^2 \in \mathbb{R}^K$, respectively.

Outputs generated by the inference mode were identified as nonlinear functions of X_i and the variational parameter ϕ . The representation $z_u \sim N(\mu_i, \text{diag}(\sigma_i^2))$ was elucidated in a manner aligned with the findings presented by Liang et al. [13]. Drawing on the work of Deng et al. [14], it is posited that the SGVB estimator holds potential for approximating the ELBO of X_i .

$$\begin{aligned} \mathcal{L}(\theta, \phi; X_i, S_s) &= \mathbb{E}_{q_\phi(z_u|X_i, S_s)} [\log p(u_i|z_u) + \log p_\theta(X_i, S_s|z_u)] \\ &\quad - \beta \text{KL}(q_\phi(z_u|X_i, S_s) \| p(z_u)) \\ &\approx \log p(u_i|z_{u,l}) + \frac{1}{L} \sum_{l=1}^L \log p_\theta(X_i, S_s|z_{u,l}) \\ &\quad - \beta \cdot \text{KL}(q_\phi(z_u|X_i, S_s) \| p(z_u)) \end{aligned} \quad (2)$$

$$\begin{aligned} \text{KL} &= \frac{1}{2} \sum_{i=1}^M (\mu_i^2 + \sigma_i^2 - \log \sigma_i^2 - 1), \\ z_{u_i,l} &= \mu_i + \sigma_i \odot \varepsilon_{i,l} \end{aligned} \quad (3)$$

The term KL is understood to denote the Kullback-Leibler divergence. In this context, the parameter $\beta \in [0,1]$ has been identified as a pivotal component, employed to modulate the robustness of regularization, a strategy adopted to address the challenge of posterior collapse, as documented by Lee et al. [11]. It was also observed that $\varepsilon_{i,l}$ adheres to the normal distribution $N(0, \mathbb{I}^K)$, and \odot is interpreted to symbolize the element-wise multiplication. Inferences related to items were drawn in a manner mirroring the inferences pertaining to users. Furthermore, the derivation of the ELBO for the item-oriented network was found to emulate established protocols.

$$\begin{aligned} \mathcal{L}(\theta, \phi; Y_j, T_t) &= \mathbb{E}_{q_\phi(z_v|Y_j, T_t)} [\log p(v_j|z_v) + \log p_\theta(Y_j, T_t|z_v)] \\ &\quad - \beta \text{KL}(q_\phi(z_v|Y_j, T_t) \| p(z_v)) \\ &\approx \log p(v_j|z_{v,j,t'}) + \frac{1}{L} \sum_{l=1}^L \log p_\theta(Y_j, T_t|z_{v,j,l'}) \\ &\quad - \beta \cdot \text{KL}(q_\phi(z_v|Y_j, T_t) \| p(z_v)) \end{aligned} \quad (4)$$

$$\begin{aligned} \text{KL} &= \frac{1}{2} \sum_{j=1}^N (\mu_j^2 + \sigma_j^2 - \log \sigma_j^2 - 1), \\ z_{v_j,l'} &= \mu_j + \sigma_j \odot \varepsilon_{j,l'} \end{aligned} \quad (5)$$

3.2.3 Node embedding

In the absence of supplementary features such as user profiles or item attributes, ID embeddings were traditionally employed. To augment the precision of rating predictions, the integration of attribute data pertaining to users and items into node embeddings was proposed, as such an approach was observed to enhance the inference of latent factors for users and items. Two embedding look-up tables, denoted as $E_u = \{e_{u1}, \dots, e_{uM}\}$ and $E_v = \{e_{v1}, \dots, e_{vN}\}$, were utilized, providing an initial state for both user and item embeddings.

3.3 Simplified GCN with dual attentive propagation

Guided by the foundational principles presented in Light GCN, an innovative mechanism for the transmission of information was incorporated, with the primary objective being the capture of collaborative signals intricately woven within the graph structure. Such an endeavour was undertaken to fortify the quality and precision of user and item embeddings.

3.3.1 First-order propagation

As postulated by Wang et al. [23], an observable interaction between a user and an item often provides valuable insights into the user's predilection for said item. Furthermore, an engagement of a user with a particular item can be interpreted, not merely as an isolated event, but as an intrinsic attribute of that item. This attribute, in turn, can be instrumental in evaluating the shared collaborative inclinations between that item and others. During the embedding propagation phase, attention-driven mechanisms were deployed to determine the relevance of data exchanged across a consistent user and varying items. In the proposed model, a meticulous propagation of embedding was executed, predicated predominantly on the documented interactions between users and items. This structured propagation strategy encompassed three core procedures: the initial dissemination of information, the duality of attentive transmission, and the final synthesis of the amassed information.

Information Propagation. The information transmitted from v to u is defined as follows for a connected user-item pair (u, v) .

$$\text{IP}(u, v) = f(e_u, e_v, p_{u,v}) = \frac{e_v}{\sqrt{|\mathbb{N}_u| |\mathbb{N}_v|}} \quad (6)$$

where, function $\text{IP}(\cdot)$ is delineated as the conduit for the propagation of information embeddings, and the role of function f is to facilitate the encoding of this information. The primary inputs to this function are represented by the embeddings, namely e_u and e_v . To modulate the decay factor of each propagation occurring on the edge (u, v) , the parameter $p_{u,v}$ is introduced. This parameter essentially quantifies the historical influence items exert on user preferences. The set of items and users interacting with user u and item v are symbolized by \mathbb{N}_u and \mathbb{N}_v , respectively. These denote the first-hop neighbors for both u and v . Intriguingly, $p_{u,v}$ is conceptualized as the square root of the graph Laplacian norm $(|\mathbb{N}_u| |\mathbb{N}_v|)^{-1/2}$, offering further insights into the extent of the influence of historical items on user inclinations. In the pursuit of amplifying both performance and representational

capability, the distinction of NGCF lies in its opting for an identity matrix over a feature transformation matrix during the information aggregation phase. Such an approach has been observed to enhance recommendation performance.

Dual Attentive Propagation. Deficiencies in user-item interaction signals are noted to impact the precision of the CF task, as underscored by Shi et al. [29]. Given such challenges, a dual attentive network, geared towards assimilating both user and item attentions, has been integrated to enhance the acquisition of their embeddings. The embeddings within this network are iteratively updated based on calculated attentions corresponding to diverse items. Subsequent to this, predictions are formulated by amalgamating the embeddings corresponding to self-connection information with those linked to interactive data.

Upon procuring the node embeddings, attention mechanisms are deployed to discern a user's inclination towards an assortment of items, with subsequent updates applied to the embeddings. The ensuing step involves a fusion of the self-connection message embedding with the interactive message embedding to proffer the concluding prediction. In this study, the dual attention mechanism is articulated through the subsequent equations:

$$\begin{aligned} H_j &= \text{LeakyRelu}(W_a e_{v_j} + b_a), \\ \text{AT}_j &= \exp(e_u^\top H_j) / \sum_{k \in |\mathbb{V}|} \exp(e_u^\top H_k) \end{aligned} \quad (7)$$

where, H_j denotes the concealed representation procured from the dense, low-dimensional embedding associated with item $v_j \in |\mathbb{V}|$. The LeakyReLU function has been enlisted to bolster the nonlinear potential of the attention model, complemented by the designated hyper-parameters $W_a \in \mathbb{R}^{K \times K}$ and $b_a \in \mathbb{R}^{K \times 1}$. Given a sequential labelling of items from 1 to N , the dense embedding vector corresponding to item v_j is represented as e_{v_j} .

Diverging from traditional attention models which uniformly employ context vectors for each input, the adopted context vector in this methodology is identified as the embedding e_u . The attention score AT_j , indicative of the significance of v_j for u , is discerned by harnessing the Softmax function. This assists in computing the normalized similarity between H_j and e_u . The representation of the user, denoted as e_u , is then derived by integrating the item embeddings, each weighed by their corresponding attention scores. Such a computational approach is designed to encapsulate both inherent traits of u_i and the interactions noted between the user and the item.

$$e_u = e_{u_i} + \sum_{j \in |\mathbb{V}|} \text{AT}_j e_{v_j} \quad (8)$$

Information Aggregation. During the aggregation phase, information from the neighborhoods of u is assimilated to refine its representation. The accuracy of predictions is purported to be augmented when information propagation and aggregation are undertaken using a straightforward weighted sum aggregator. Notably, there is no reliance on feature transformation or nonlinear activation functions in this context [17]. The corresponding aggregation functions are depicted as:

$$\begin{aligned}
e_u^{(l)} &= \sum_{v \in \mathbb{N}_u} \text{IP}(u, v) = \sum_{v \in \mathbb{N}_u} \frac{1}{\sqrt{|\mathbb{N}_u| |\mathbb{N}_v|}} e_v^{(0)}, \\
e_v^{(l)} &= \sum_{u \in \mathbb{N}_v} \text{IP}(v, u) = \sum_{u \in \mathbb{N}_v} \frac{1}{\sqrt{|\mathbb{N}_v| |\mathbb{N}_u|}} e_u^{(0)}
\end{aligned} \tag{9}$$

$$\begin{aligned}
e_u^{(l)} &= \sum_{v \in \mathbb{N}_u} \frac{1}{\sqrt{|\mathbb{N}_u| |\mathbb{N}_v|}} e_v^{(l-1)}, \\
e_v^{(l)} &= \sum_{u \in \mathbb{N}_v} \frac{1}{\sqrt{|\mathbb{N}_v| |\mathbb{N}_u|}} e_u^{(l-1)}, \quad l \geq 1
\end{aligned} \tag{11}$$

In the presented Eq. (9), the representations of user u and item v , symbolized by $e_u^{(l)}$ and $e_v^{(l)}$, are procured by channelling information from interconnected users and items during the foundational embedding propagation stages. The representations at the 0-th layer for the user and item are depicted as $e_u^{(0)}$ and $e_v^{(0)}$ respectively. As highlighted by He et al. [18], superior outcomes can be achieved through the application of the square root of the symmetric graph Laplacian normalization in contrast to other norms. This method also serves to circumvent any escalation in the embedding scale attributable to graph convolution operations.

Within the model described, the aggregation functions solely take into account direct connections of users and items, deliberately overlooking their self-connection. It is posited that layer combination operations can fulfil the roles typically ascribed to self-connections. Such an approach presents a deviation from a majority of extant graph convolution operations [17, 23] that encompass extended neighbors and acknowledge self-connections. Upon the culmination of the K -th embedding propagation layers, the final representations for users and items are synthesized by amalgamating the embeddings discerned at each successive layer, as illustrated:

$$\begin{aligned}
e_u &= \xi_0 e_u^{(0)} + \xi_1 e_u^{(1)} + \dots + \xi_K e_u^{(K)} = \sum_{k=0}^K \xi_k e_u^{(k)}, \\
e_v &= \xi_0 e_v^{(0)} + \xi_1 e_v^{(1)} + \dots + \xi_K e_v^{(K)} = \sum_{k=0}^K \xi_k e_v^{(k)}
\end{aligned} \tag{10}$$

where, ξ_k stands as a modifiable hyper-parameter reflecting the importance of the embedding in the k -th layer for the synthesis of the concluding embedding. Disparate embedding layers are understood to encapsulate varied representations. For instance, the foundational layer is tailored for fostering unimpeded interactions between users and items. Subsequent layers accentuate the congruity between users and items exhibiting shared interactions. Elevated layers, in contrast, are adept at capturing higher-order proximities. Such a conglomerate approach ensures a more encompassing final representation. The infusion of the attention mechanism into the embedding propagation is believed to amplify the dynamics of user-item interactions.

3.3.2 High-order propagation

Upon completion of the primary information propagation, representations of both users and items are ascertained. To further delve into higher-order interactions between users and items, supplementary layers of embedding propagation are introduced. It is posited that these interactions are instrumental in encapsulating collaborative signals and discerning the affiliations between users and items. By invoking multiple layers of embedding propagation, both users and items stand to assimilate information conveyed by their neighbors within a defined range of hops. Within the model, the recursion formulas are employed to determine the representations of the user and item at the l -th iteration.

By iteratively implementing multiple layers of embedding propagation, the user and item representations, delineated as $e_u^{(l-1)}$ and $e_v^{(l-1)}$ from the $(l-1)$ -th propagation phase, are believed to inculcate collaborative signals from their respective $(l-1)$ -hop neighbors. Such an integration is purported to augment representation learning and thereby bolster performance.

3.3.3 Propagation in matrix form

To facilitate the execution of HAGCN, the matrix formulation is introduced. Let $E^{(0)} \in \mathbb{R}^{(M+N) \times d_l}$ stand as the embedding matrix at the initial layer, and let $d_l < \min(M, N)$ represent the size of the embedding. The corresponding matrix form is delineated as:

$$\begin{aligned}
E^{(l)} &= \left(D^{-\frac{1}{2}} \mathbb{A} D^{-\frac{1}{2}} \right) E^{(l-1)}, \quad D = \begin{pmatrix} 0 & \mathbb{R} \\ \mathbb{R}^T & 0 \end{pmatrix}, \\
E &= \sum_{k=0}^l \xi_k E^{(k)} = \xi_0 E^{(0)} + \xi_1 E^{(1)} + \dots + \xi_l E^{(l)}
\end{aligned} \tag{12}$$

where, \mathbb{A} serves to represent the adjacency matrix pertaining to the user-item graph. Furthermore, the diagonal degree matrix D encapsulates the count of non-zero elements present within each row of the matrix \mathbb{A} . The parameter ξ_k is posited to demarcate the significance of the embedding at the k -th layer within the holistic embedding.

3.4 Prediction and optimization

Upon the culmination of the attentive embedding propagation, a series of user representations is acquired from multiple layers. These representations underscore the nuances of information propagation through distinct connections. These sets of user representations, represented as $\{e_u^{(1)}, e_u^{(2)}, \dots, e_u^{(l)}\}$, signify varied importances concerning user preferences. The eventual user embeddings are formulated by collating $e_u^{(1)}, e_u^{(2)}, \dots, e_u^{(l)}$. Analogously, the item representations are derived by amalgamating the item representations $e_v^{(1)}, e_v^{(2)}, \dots, e_v^{(l)}$ amassed across diverse layers. The ultimate representation for both users and items is established as:

$$\begin{aligned}
e_u^* &= e_u^{(0)} \otimes e_u^{(1)} \otimes \dots \otimes e_u^{(l)}, \\
e_v^* &= e_v^{(0)} \otimes e_v^{(1)} \otimes \dots \otimes e_v^{(l)}
\end{aligned} \tag{13}$$

where, \otimes symbolizes a concatenation operation, absent of learning any ancillary parameters. Such an operation has been demonstrated as efficacious in GCN [22]. Through the deployment of concatenation operations, initial embeddings are enhanced via attentive embedding propagation, with the propagation breadth modulated by the parameter l . The focal point remains on embedding learning, with a fundamental inner product interaction function utilized for deriving the ranking score for recommendations. This implies that the preference of a user for a specific item is evaluated through

$R^* = e_u^{*\top} e_u^*$. Future investigations are poised to delve into more sophisticated mechanisms, encompassing VAE-based interaction functions.

Following the prediction of ratings, optimization of the loss functions is conducted to achieve optimal performance. Typically, these loss functions encompass both the feature reconstructing error and the rating prediction error. Training of parameters in HAGCN is confined to the embeddings of the initial layer, making its complexity comparable to that of the standard MF. Therefore, the pairwise BPR loss is adopted as the principal loss function for HAGCN, in line with findings by He et al. [18]. This loss function accounts for the relative order of both observed and unobserved user-item interactions. Notably, the pairwise BPR loss tends to favour predictions of observed interactions with elevated values in comparison to their unobserved counterparts. The formulation of HAGCN's loss function is delineated as:

$$\mathcal{L}_{\text{BPR}} = \sum_{u=1}^M \sum_{i \in \Psi^+} \sum_{j \in \Psi^-} -\ln \sigma(R_{ui}^* - R_{uj}^*) + \lambda_{L2} \|E^{(0)}\|^2 \quad (14)$$

In Eq. (15), Ψ^+ signifies the set of observed interactions, while Ψ^- denotes the set of unobserved instances. The sigmoid

function is represented as $\sigma(\cdot)$. All trainable parameters within the 0-th layer are encapsulated by $E^{(0)}$, with λ_{L2} being employed to modulate the intensity of L2 regularization, thereby mitigating overfitting.

Drawing parallels with SGCN and Light GCN, a minimal increase in parameters is observed in the model to facilitate the modelling of high-order connectivity. Given that HAGCN's propagation layers encompass a limited number of parameters, and that the size of the parameters matrix is defined by $d_l \times d_{l-1}$ (where d_l represents the embedding size and $d_l < \min(M, N)$), the model's size is approximated as $2C \times d_l \times d_{l-1}$. Herein, C typically remains a minor integer, seldom surpassing four, as these embedding matrices are derived from the underlying user-item graph structure and corresponding weight matrices.

4. EXPERIMENTS AND DISCUSSIONS

4.1 Experimental settings

4.1.1 Datasets

The proposed method was assessed using datasets sourced from GroupLens, Yelp, Epinions, and Amazon. An overview of the characteristics of these datasets is presented in Table 2.

Table 2. Statistics of real-world datasets

Dataset	Users	Items	Ratings	Sparsity	User Features	Item Features
ML1M	6,040	3,706	1,000,209	95.53%	Demographics	Genres
Yelp	31,668	38,048	1,561,406	99.87%	Social relations	Categories
Epinions	49,289	139,738	664,823	99.99%	Trust relations	Topics
Douban	278,297	21,359	1,048,576	99.98%	Topics	Genres

The MovieLens-1M dataset (ML1M), supplied by GroupLens, has traditionally been utilised for evaluating recommendation algorithms. This dataset encompasses 6,040 users, 3,706 films, and over 1,000,209 ratings, with each user having provided ratings for at least 20 movies. Ratings span a range of 1 to 5. For the purpose of assessing learning performance based on implicit feedback, ML1M was transformed into implicit data. Collaborative information was inferred from user attributes such as age, occupation, and gender, whereas auxiliary item information was derived from movie genres.

Data from Yelp in 2018 incorporated customer feedback, tied to a 1 to 5 rating scale. To structure a user-item matrix, a binary conversion was employed using a threshold of 3. Retained reviews were limited to English language content and were exclusively restaurant-related. Sparsity reduction measures entailed excluding users with fewer than 5 reviews and establishments rated by fewer than 30 users. Duplicate ratings were unified by their earliest time-stamp. The refined dataset comprised 25,815 users, 25,677 businesses, and totalled 730,791 ratings.

The dataset from Epinions.com, a notable consumer opinion platform, permits users to convey feedback, appraise diverse items (ranging from literature and music to gadgets), and share experiences and expertise with their counterparts. Within this platform, the establishment of trust boundaries by users facilitates the formation of distinct communities. Within these communities, reviews and ratings, as provided by the users, have been repeatedly evidenced to hold significant value. The dataset is characterized by its notable sparsity, encompassing

49,289 users, 139,738 items, 664,823 ratings, and 487,183 trust connections. The attributes of this dataset comprise trust relationships between users and the topics of the items reviewed.

Data sourced from Douban, a renowned online platform, in 2019, incorporates over 140,000 films, 70,000 performers, 600,000 users, 4.16 million movie ratings, and 4.42 million film reviews. Among all datasets, Douban exhibits the highest level of sparsity. It is observed that a significant portion of users and movies have fewer than 6 interactions. In the pursuit of ensuring comprehensive user feedback on movies, entities and users with fewer than 10 interactions were excluded from the dataset. The refined dataset encapsulates 278,297 users, 21,359 items, and 1,048,576 ratings.

For the empirical evaluation, each dataset was bifurcated: the training set, comprising a random selection of 80% of user ratings, and the test set, reserved with the remaining 20%.

4.1.2 Baselines and evaluation metrics

For the purpose of assessing the model under discussion, a selection of five CF models was chosen as benchmarks.

- ANCF [6] represents an attribute-driven model. Through the deployment of an attention mechanism, this model discerns the significance levels of attribute data and subsequently assimilates attributes during feature acquisition, aiming for a holistic feature representation.
- NGCF [23] stands as a refined GCN model amalgamating NCF with graph convolutional networks. This combination is devised to capture intricate relationships inherent in user-item interactions, thereby enhancing

recommendation precision.

- KGAT [25], an avant-garde model, amalgamates graph convolutional networks, knowledge graph embeddings, and attention mechanisms, all of which synergise to refine recommendation accuracy.
- LRGCCF [27] presents itself as an evolved GCN method. Characteristically, it eschews non-linearity and infuses a residual network structure, an innovation targeting the mitigation of over-smoothing issues during graph convolution aggregation.
- Light GCN [18], evolved from the NGCF model, leverages both linear and attentive embedding propagation, devised to judiciously diffuse embeddings from adjacent nodes.

For the quantification of the model's efficacy, a set of four established evaluation metrics tailored for social recommendation were employed: mean absolute error (MAE), root mean square error (RMSE), Recall, and normalized discounted cumulative gain (NDCG). While MAE and RMSE are harnessed to gauge prediction accuracy, Recall@20 and NDCG@20 serve as tools to assess the performance of top-tier recommendations.

4.1.3 Parameter settings

Within the training dataset, for every positive instance, four negative instances were sampled. An embedding size of 64 was established, and the Xavier method was applied for the initialization of embedding parameters across all methodologies. Parameters for the proposed model were initialized by drawing samples from a Gaussian distribution characterized by a mean of 0 and a standard deviation of 0.01. HAGCN's optimization was undertaken employing a mini-batch Adam method, akin to the approach delineated by He et al. [18]. A learning rate of 0.001 was fixed, with a default mini-batch size determined at 1024. During feature extraction endeavors, K was fixed at 128. The AVAE comprises two latent layers equipped with LeakyReLU activation, both functioning as generative networks. Adjustments were made for the dropout ratio and the parameter β within the confines

of $\{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$ and $\{0.2, 0.4, 0.6, 0.8, 1.0\}$ respectively. With β fixed at 0.2, the most commendable performance by AVAEs was noted. In the span of the simplified GCN with attentive propagation, the L2 regularization coefficient was investigated within the spectrum of $\{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. Concurrently, the layer combination coefficient was designated as $(1+\kappa)^{-1}$. Here, κ , symbolizing the number of layers, underwent testing in the interval [1, 4]. Optimal results across all datasets were achieved when κ was equated to 3.

4.2 Experimental results and discussions

4.2.1 Overall performance

Table 3 presents the performance evaluations of HAGCN and five other methods on all datasets, comparing rating prediction and top-ranking based on MAE/RMSE. Additionally, Table 4 displays the results for Recall@20 and NDCG@20.

The comparison of rating prediction performance on MAE and RMSE with all methods is presented in Table 3, emphasizing the highest scores achieved by each method. The exceptional effectiveness of the proposed HAGCN is evident as it consistently outperforms other methods on MAE and RMSE for all datasets, validating its remarkable efficacy achieved through a straightforward yet logical design. Comparing to Light GCN and other GCN models, our model consistently achieves promising performance by leveraging attributes boosting based on variational graph framework. Light GCN outperforms LRGCCF, KGAT, and NGCF among the GCN-based baselines. The performance of KGAT is comparable to that of NGCF, but it barely outperforms NGCF. The performance comparison of all methods on Recall@20 and NDCG@20 is shown in Table 4. The superiority of HAGCN over the other five baselines is apparent, with recall rates ranging from 2.23% to 7.33% and NDCG rates ranging from 2.22% to 7.53% across all datasets. Additionally, HAGCN considerably outshines ANCF by at least 7.84%, which indicates the GCN structure can facilitate capturing collaborative signals for CF task.

Table 3. Performance comparison among all methods on MAE and RMSE

Dataset	Metrics	ANCF	NGCF	KGAT	LRGCCF	LightGCN	HAGCN
ML1M	MAE	0.8725	0.8141	0.8043	0.7812	0.7431	0.6874
	RMSE	0.9807	0.9126	0.9011	0.875	0.8316	0.7681
Yelp	MAE	0.9171	0.8562	0.8505	0.8101	0.7709	0.7232
	RMSE	1.167	1.0889	1.0813	1.0294	0.9644	0.9151
Epinions	MAE	1.0012	0.9554	0.9528	0.8912	0.8512	0.8193
	RMSE	1.1398	1.0881	1.0855	1.0149	0.9691	0.9325
Douban	MAE	0.5749	0.5634	0.5672	0.5561	0.5457	0.5298
	RMSE	0.7663	0.7508	0.7557	0.7413	0.7266	0.7051

Table 4. Performance comparison among all methods on Recall@20 and NDCG@20

Dataset	Metrics	ANCF	NGCF	KGAT	LRGCCF	LightGCN	HAGCN
ML1M	Recall@20	0.2316	0.2514	0.2475	0.2552	0.2601	0.2696
	NDCG@20	0.2079	0.2407	0.2376	0.2517	0.2560	0.2623
Yelp	Recall@20	0.0551	0.0563	0.0572	0.0585	0.0614	0.0659
	NDCG@20	0.0812	0.0829	0.0843	0.0861	0.0903	0.0971
Epinions	Recall@20	0.6863	0.7110	0.7004	0.7333	0.7535	0.7703
	NDCG@20	0.7261	0.7526	0.7539	0.7896	0.8111	0.8291
Douban	Recall@20	0.0703	0.0698	0.0757	0.0820	0.0703	0.0875
	NDCG@20	0.0817	0.0852	0.0845	0.0921	0.0998	0.1061

Table 5. The performance comparison in cold-start scenarios on Recall@20

Dataset	Scenario	ANCF	KGAT	LRGCCF	HAGCN
ML1M	Cold-U	0.1623	0.1656	0.1888	0.2068
	Cold-V	0.1725	0.1745	0.1919	0.2101
Yelp	Cold-U	0.0341	0.0344	0.0385	0.0435
	Cold-V	0.0376	0.0389	0.0414	0.0468
Epinions	Cold-U	0.3992	0.4068	0.4471	0.4809
	Cold-V	0.4374	0.4438	0.4816	0.5223
Douban	Cold-U	0.0420	0.0434	0.0502	0.0551
	Cold-V	0.0433	0.0446	0.0512	0.0562

Table 6. The performance comparison in cold-start scenarios on NDCG@20

Dataset	Scenario	ANCF	KGAT	LRGCCF	HAGCN
ML1M	Cold-U	0.1965	0.2017	0.2225	0.2405
	Cold-V	0.2052	0.2110	0.2346	0.2543
Yelp	Cold-U	0.0287	0.0302	0.0322	0.0333
	Cold-V	0.0321	0.0336	0.0364	0.0378
Epinions	Cold-U	0.4929	0.4951	0.5442	0.5790
	Cold-V	0.5389	0.5430	0.5945	0.6329
Douban	Cold-U	0.0500	0.0513	0.0605	0.0688
	Cold-V	0.0517	0.0525	0.0627	0.0722

Referring to Tables 3 and 4, it becomes evident that most GCN-based methodologies consistently surpass ANCF. This highlights the intrinsic ability of GCNs to derive enhanced latent user and item representations. Further examination reveals that simplified GCN models, notably HAGCN and Light GCN, excel in comparison to conventional GCNs across all metrics and datasets. Such findings suggest that by simplifying graph convolution and layer combinations within GCNs, model complexity is reduced, which in turn mitigates the risk of overfitting during graph-based learning. This enhances both the performance and scalability of social recommendation systems. HAGCN, being rooted in a streamlined GCN structure, consistently surpasses the other five cutting-edge baselines across all the metrics on every dataset, underscoring the potency of its hybrid attributive framework.

4.2.2 Performance in cold-start scenarios

To rigorously evaluate the performance under diverse cold-start conditions, test sets were curated with varying proportions of cold data. By randomly selecting 30% of the samples from the test datasets to represent 30% cold users, each sample was accorded a distinct user id exclusive to that instance. Evaluation was conducted under scenarios wherein 30% of users (Cold-U) and 30% of items (Cold-V) were novel, spanning all datasets. Metrics of evaluation employed were Recall@20 and NDCG@20. In juxtaposition, Light GCN and NGCF, which are primarily contingent on feedback data and devoid of attribute information, were observed to be less adept at navigating cold-start contexts (Tables 5 and 6).

For users experiencing cold-start scenarios, where interaction data are notably sparse, it has been found that harnessing auxiliary social connections within HAGCN can potentially discern latent preferences. Similarly, when addressing cold-start items with limited user interactions, the HAGCN mechanism is shown to capitalize on supplementary attribute data, refining similarity computations for CF tasks. Through this approach, embedding learning based on

Bayesian inference is harnessed, suggesting that cold-start embeddings benefit from the statistical correlations between users and items.

4.2.3 Analysis of layer combination

Experiments were conducted to probe the potential advantages of integrating multiple embedding propagation layers within HAGCN. The objective was to gauge the impact of varying layer counts, spanning a range of 1 to 4 layers, across all datasets. For the clarity of representation, only Recall and NDCG were selected as evaluative metrics, and the corresponding findings are depicted in Figure 2.

With an increase in the number of layers, a gradual enhancement in the performance of HAGCN is observed. Optimal outcomes on metrics such as Recall and NDCG across all datasets are noted when the layer number is three. These significant improvements can be attributed to the proficient encapsulation of the CF effect, whereby the second-order connectivities inherently grasp the collaborative user similarity and the third-order connectivities capture the collaborative signals.

Upon extending to four layers, the performance on the ML1M and Yelp datasets consistently showed an upward trend. However, a marginal overfitting is noticed on the Epinions dataset and a pronounced decline in performance on the Douban datasets. Such performance fluctuations might be traced back to the potential introduction of noise as the network architecture deepens during graph representation learning. The nuanced gains on ML1M and Yelp datasets suggest that the impact of a four-layer propagation essentially mirrors that of a three-layer propagation, deemed adequate for encapsulating CF signals. This three-layer combination was identified to be as effective in counteracting over-smoothing as seen in the PPNP model by Gasteige et al. [26]. The layer combination coefficient for HAGCN was set in a manner akin to Light GCN, being designated as $(1+\kappa)^{-1}$. This raises the possibility that further enhancements in performance might be unlocked by fine-tuning the parameter $(1+\kappa)^{-1}$.

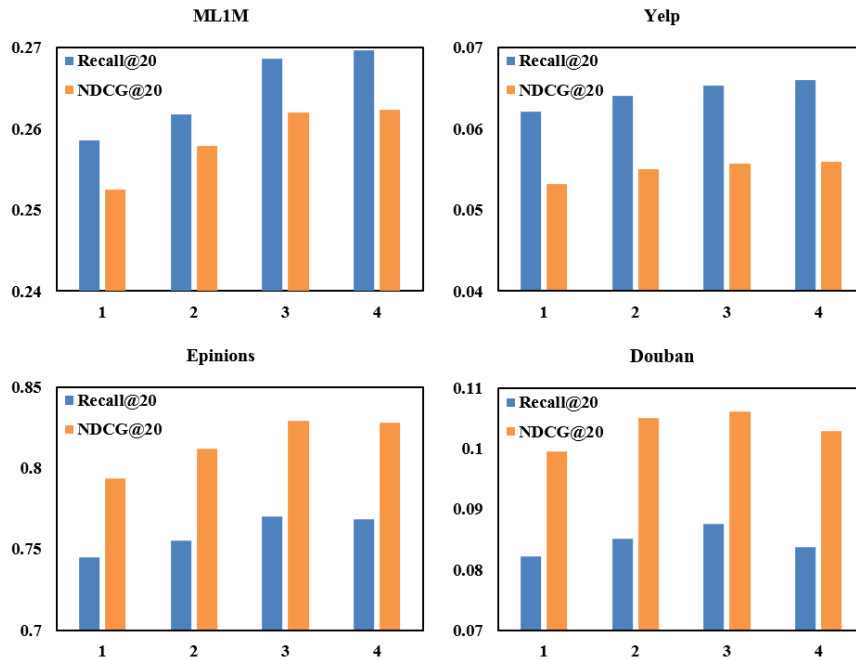


Figure 2. Performance comparison of HAGCN with different layers on four datasets

5. CONCLUSION

In the study presented, a hybrid attributive neural graph structure was introduced for enhanced social recommendation. Dual AVAEs structure was employed for feature extraction, adeptly capturing attribute details from both users and items. Through this method, linear and non-linear latent representations of users and items for feature embedding were derived. Initial embeddings were inferred from shared distributions anchored on Bayesian inference, allowing cold-start embeddings to derive benefit from statistical strength among users and items.

A simplified GCN, embedded with attentive feature propagation across layers, was integrated to ascertain a profound high-order connectivity, unveiling explicit collaborative signals within the interaction graph. By amalgamating user and item attribute data into latent factors via a deep neural graph network, data sparsity challenges were addressed, and enhanced user and item representation was achieved. Attention was given to refining the embeddings of users and items, by including a layer of attentive embedding propagation that amalgamates the embeddings of interactive users and items. Multiple embedding propagation layers, layered with weighted sum, were further augmented to harness collaborative indications in advanced connectivity. This provision was invaluable in ameliorating over-smoothing.

As a result, HAGCN was postulated to effectively tackle the cold-start problem and augment the learning of user and item latent factors. This was achieved by synergising the intrinsic Bayesian probabilistic perspective evident in VAE and GCN. Experimental results validated the superiority of HAGCN over contemporaneous methods such as NGCF, KGAT, LRGCCF, and Light GCN. An improvement of at least 2.9% on MAE/RMSE for rating prediction performance was observed, while top-ranking performance witnessed an upliftment of at least 2.2% on Recall/NDCG. The efficacy of HAGCN in addressing challenges in cold-start scenarios, particularly those bereft of preliminary data, was further corroborated through extensive experimentation.

Given the mounting inclination towards graph-based models in social recommendations, harnessing supplementary data from social commerce or media platforms has emerged as a novel trajectory. Yet, extant GCN models grapple with complexities in their GCN design. Prospective efforts will pivot on probing into the simplified graph convolutional layers within these GCNs to amplify recommendation outputs. The focus might shift towards harnessing limited rather than infinite message passing and reducing regularization for expedited training. Emphasis will also be placed on refining the variational graph architecture to deduce users/items embeddings from attribute distributions via Bayesian inference, coupled with the application of VAE-based interaction functions for reconstructing elusive preference embeddings.

ACKNOWLEDGMENT

This work was supported by the Office for Philosophy and Social Science of Fujian Province, China (grant number FJ2020B047); and the Natural Science Foundation of Fujian Province, China (grant number 2020J01112035).

REFERENCES

- [1] Eirinaki, M., Gao, J., Varlamis, I., Tserpes, K. (2018). Recommender systems for large-scale social networks: A review of challenges and solutions. *Future Generation Computer Systems*, 78: 413-418. <https://doi.org/10.1016/j.future.2017.09.015>
- [2] Shi, Y., Larson, M., Hanjalic, A. (2014). Collaborative filtering beyond the user-item matrix: A survey of the state of the art and future challenges. *ACM Computing Surveys (CSUR)*, 47(1): 1-45. <https://doi.org/10.1145/2556270>
- [3] Koren, Y., Bell, R., Volinsky, C. (2009). Matrix factorization techniques for recommender systems.

- Computer, 42(8): 30-37. <https://doi.org/10.1109/MC.2009.263>
- [4] Wang, C., Blei, D.M. (2011). Collaborative topic modeling for recommending scientific articles. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Diego California USA, pp. 448-456. <https://doi.org/10.1145/2020408.2020480>
- [5] He, X., Liao, L., Zhang, H., Nie, L., Hu, X., Chua, T.S. (2017). Neural collaborative filtering. In Proceedings of the 26th International Conference on World Wide Web, China, pp. 173-182. <https://doi.org/10.1145/3038912.3052569>
- [6] Chen, H., Qian, F., Chen, J., Zhao, S., Zhang, Y. (2021). Attribute-based neural collaborative filtering. Expert Systems with Applications, 185: 115539. <https://doi.org/10.1016/j.eswa.2021.115539>
- [7] Zhang, S., Yao, L., Sun, A., Tay, Y. (2019). Deep learning based recommender system: A survey and new perspectives. ACM Computing Surveys (CSUR), 52(1): 1-38. <https://doi.org/10.1145/3285029>
- [8] Tay, Y., Tuan, L., Hui, S.C. (2018). Latent relational metric learning via memory-based attention for collaborative ranking. In Proceedings of the 2018 World Wide Web Conference, Lyon France pp. 729-739. <https://doi.org/10.1145/3178876.3186154>
- [9] Guo, H., Tang, R., Ye, Y., Li, Z., He, X. (2017). DeepFM: a factorization-machine based neural network for CTR prediction. In Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp. 1725-1731. <https://doi.org/10.48550/arXiv.1703.04247>
- [10] He, X., Chua, T.S. (2017). Neural factorization machines for sparse predictive analytics. In Proceedings of the 40th International ACM SIGIR conference on Research and Development in Information Retrieval, Shinjuku Tokyo Japan, pp. 355-364. <https://doi.org/10.1145/3077136.3080777>
- [11] Lee, W., Song, K., Moon, I. C. (2017). Augmented variational autoencoders for collaborative filtering with auxiliary information. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, Australia, pp. 1139-1148. <https://doi.org/10.1145/3132847.3132972>
- [12] Li, X., She, J. (2017). Collaborative variational autoencoder for recommender systems. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax NS Canada, pp. 305-314. <https://doi.org/10.1145/3097983.3098077>
- [13] Liang, D., Krishnan, R.G., Hoffman, M.D., Jebara, T. (2018). Variational autoencoders for collaborative filtering. In Proceedings of the 2018 World Wide Web Conference, Lyon France, pp. 689-698. <https://doi.org/10.1145/3178876.3186150>
- [14] Deng, X., Zhuang, F., Zhu, Z. (2019). Neural variational collaborative filtering with side information for top-K recommendation. International Journal of Machine Learning and Cybernetics, 10: 3273-3284. <https://doi.org/10.1007/s13042-019-01016-2>
- [15] Rendle, S., Krichene, W., Zhang, L., Anderson, J. (2020). Neural collaborative filtering vs. matrix factorization revisited. In Proceedings of the 14th ACM Conference on Recommender Systems, Virtual Event Brazil, pp. 240-248. <https://doi.org/10.1145/3383313.3412488>
- [16] Deng, X., Wu, Y.J., Zhuang, F. (2020). Trust-embedded collaborative deep generative model for social recommendation. The Journal of Supercomputing, 76: 8801-8829. <https://doi.org/10.1007/s11227-020-03178-1>
- [17] Wu, F., Souza, A., Zhang, T., Fifty, C., Yu, T., Weinberger, K. (2019). Simplifying graph convolutional networks. In International Conference on Machine Learning, pp. 6861-6871. PMLR.
- [18] He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., Wang, M. (2020). LightGCN: Simplifying and powering graph convolution network for recommendation. In Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval, pp. 639-648. <https://doi.org/10.1145/3397271.3401063>
- [19] Gao, C., Zheng, Y., Li, N., Li, Y., Qin, Y., Piao, J., Li, Y. (2023). A survey of graph neural networks for recommender systems: Challenges, methods, and directions. ACM Transactions on Recommender Systems, 1(1): 1-51. <https://doi.org/10.1145/3568022>
- [20] Ying, R., He, R., Chen, K., Eksombatchai, P., Hamilton, W.L., Leskovec, J. (2018). Graph convolutional neural networks for web-scale recommender systems. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, London United Kingdom, pp. 974-983. <https://doi.org/10.1145/3219819.3219890>
- [21] Fan, W., Ma, Y., Li, Q., He, Y., Zhao, E., Tang, J., Yin, D. (2019). Graph neural networks for social recommendation. In The world wide web conference, San Francisco CA USA, pp. 417-426. <https://doi.org/10.1145/3308558.3313488>
- [22] Berg, R., Kipf, T., Welling, M. (2018). Graph convolutional matrix completion. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1-7. <https://doi.org/10.48550/arXiv.1706.02263>
- [23] Wang, X., He, X., Wang, M., Feng, F., Chua, T.S. (2019). Neural graph collaborative filtering. In Proceedings of the 42nd international ACM SIGIR conference on Research and development in Information Retrieval, Paris France, pp. 165-174. <https://doi.org/10.1145/3331184.3331267>
- [24] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y. (2018). Graph attention networks. In Proceedings of 6th International Conference on Learning Representations, pp. 1-12. <https://doi.org/10.48550/arXiv.1710.10903>
- [25] Wang, X., He, X., Cao, Y., Liu, M., Chua, T.S. (2019). KGAT: Knowledge graph attention network for recommendation. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage AK USA, pp. 950-958. <https://doi.org/10.1145/3292500.3330989>
- [26] Gasteiger, J., Bojchevski, A., Günnemann, S. (2019). Predict then propagate: Graph neural networks meet personalized pageRank. In Proceedings of the 7th International Conference on Learning Representations, New Orleans, LA, USA, pp. 1-15. <https://doi.org/10.48550/arXiv.1810.05997>
- [27] Chen, L., Wu, L., Hong, R., Zhang, K., Wang, M. (2020). Revisiting graph based collaborative filtering: A linear residual graph convolutional network approach. In Proceedings of the AAAI conference on artificial

- intelligence, 34(1): 27-34.
<https://doi.org/10.1609/aaai.v34i01.5330>
- [28] Guo, Z., Wang, H. (2020). A deep graph neural network-based mechanism for social recommendations. *IEEE Transactions on Industrial Informatics*, 17(4): 2776-2783. <https://doi.org/10.1109/TII.2020.2986316>
- [29] Shi, S., Zhang, M., Liu, Y., Ma, S. (2018). Attention-based adaptive model to unify warm and cold starts recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management, Torino Italy*, pp. 127-136. <https://doi.org/10.1145/3269206.3271710>