Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 3 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i03.49427

PAPER

Knowledge Graph-Based Recommender Systems to Mitigate Data Sparsity: A Systematic Literature Review

Bouchra Boualaoui[1](✉), Ahmed Zellou[1], Mouna Berquedich[2]

[1]SPM Research Team, ENSIAS, Mohammed V University in Rabat, Rabat, Morocco

[2]Green Tech Institute, Mohammed VI Polytechnique University, Benguerir, Morocco

bouchra_boualaoui@um5.ac.ma

ABSTRACT

Recommender systems (RSs) have become important tools in the modern lifestyle; they have been integrated into all domains, spanning from entertainment (music, films, etc.) to more sensitive fields such as security and health care. Their success does not mean that they are ideal or flawless; quite the opposite, RSs suffer from plenty of drawbacks and challenges that need to be resolved. Data sparsity is a common problem in recommender systems; it has been of top interest among researchers. Numerous approaches from different perspectives have been proposed to mitigate it, including knowledge graphs (KGs), which quickly gained popularity due to the rich semantics residing in their components. In this paper, we will conduct a systematic literature review to explore and analyze in depth the existing contributions. Our work focuses on investigating the effectiveness of KGs to mitigate the data sparsity in RSs by discovering the techniques used, understanding how KGs are exploited, and what type of knowledge is extracted from them, besides studying their evaluation measures and discussing future directions that can strengthen the application of KGs to mitigate data sparsity in recommender systems.

KEYWORDS

recommender systems (RSs), sparsity, systematic literature review, knowledge graphs (KGs), embedding

1 INTRODUCTION

Understanding the human being has always been a question of existence, a goal, and a purpose in the history of human civilizations. It has evolved to be manifested in many forms, especially in the era of high-tech, where recommender systems (RSs) took place to deeply understand humans and their behaviors and try to offer the best personalized experience based on this knowledge.

Over the course of developing theory on RSs, several generations have been developed. The first generation of RSs was based on content and collaborative filtering, followed by techniques that leveraged contextual information such as time, location, etc. However, another generation is growing and becoming increasingly interesting,

Boualaoui, B., Zellou, A., Berquedich, M. (2025). Knowledge Graph-Based Recommender Systems to Mitigate Data Sparsity: A Systematic Literature Review. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(3), pp. 115–140. https://doi.org/10.3991/ijim.v19i03.49427

Article submitted 2024-05-21. Revision uploaded 2024-07-29. Final acceptance 2024-10-01.

which focuses on semantic models and the use of all the knowledge components involved in the process of generating recommendations. Each class was delegated to deal with specific challenges, yet some common issues were addressed in all of them. Of which we can cite cold start, serendipity, filter bubble, and data sparsity.

The data sparsity problem arises because of how limited the interaction with the system is in comparison to the huge number of options, leading to insufficient supervised signals and difficulties in modeling users' behavior. Therefore, the generated recommendations do not perfectly reflect users' preferences yet. Matrix factorization is one of the widely used techniques to mitigate data sparsity in the literature by learning the latent factors of items and users and producing a significantly lower-dimensional matrix while removing uncharacteristic elements. Despite their effectiveness, MF techniques still suffer from overfitting issues when the rating matrix is highly sparse [1], [2]. Other contributions proposed user clustering methods with enhanced similarity functions to not only rely on co-interacted items but also include items only rated by one of the users [3], [4]. While these suggested techniques can provide accurate recommendations with minimal ratings, their high dependency on rating information makes them unsuitable for absolute cold start scenarios with no start ratings available. To handle these issues, side information was introduced to fill the gaps and overcome the inaccuracies resulting from the sparse user-item interaction matrix [5], [6]. However, most of the works focus on auxiliary data in one context, limiting the usage of the large-scale heterogeneous data created every day by the different interactions within the digital world. Hence, researchers have proposed knowledge graphs (KGs) as a new strategy to tackle data sparsity.

The increased attention to KGs as a promised approach to alleviating the data sparsity is justified by KGs' features. KGs are considered as large, structured sources to store data in the form of triplets from diverse data inputs, offering the ability to mine users' interests more accurately because of their rich heterogeneous semantics residing in the multi-relational triplets and their connectivity [7]–[9]. Nevertheless, despite the widespread adoption of KGs in RSs to improve the quality of recommendations in general and in sparse settings in particular, few systematic surveys were dedicated to investigating and analyzing the current research landscape, but none of them mainly focused on the integration of KGs as a key solution to mitigate the sparseness of data in RSs despite the outstanding works published in this regard, which makes the field fruitful and promising, but at the same time full of gaps and lacking comprehensive analysis on the effectiveness of exploiting KGs to handle the sparsity of data in recommender systems.

Our contribution consists of offering a systematic literature review (SLR) conducted to investigate the efficiency of KG-based RSs under sparse scenarios. We have focused our analysis on the last decade to understand where recent scientific research is directed in this field. The study analyzes 51 academic studies published between 2013 and 2023, selected among 459 articles across major scholarly data sources (ACM, Scopus, IEEE, and ScienceDirect). In this regard, we provide a comprehensive review guided by key research questions to uncover the techniques and methodologies employed in this area, investigate how KGs were integrated and leveraged to mitigate the data sparsity problem, explore how the performance was evaluated, and guide the research community by highlighting the emerging trends and providing insights into limitations and open opportunities that can contribute to the field. The paper proceeds as follows: Section 2 presents the background of recommender systems. Section 3 outlines the methodology employed and the research questions. Followed by Section 4, where the key findings are represented and the selected studies are analyzed. In Section 5, we discuss the findings and identify potential avenues for future research. Finally, Section 6 offers a conclusion and future work.

2 BACKGROUND

The foundation of RSs could be marked by the evolution of information retrieval's field, where researchers started paying more attention to personalized information retrieval and illuminated the way towards the first recommendation approach, which is collaborative filtering, where recommendation is performed based on the users' similarities and their historical interactions [10], but the limitation of item-user interaction opened the doors for another approach, known as content-based recommendation, that basically relies on items' attributes and users' profiles [11]. Each of these approaches has seen great improvements; one of the improvements was to combine them to benefit from both at the same time. These methods were later called hybrid methods. In recent years, hybrid methods have been characterized as a fruitful research field due to the possibility of combining different techniques such as neighborhood-based techniques [12], association rule mining [13], and contextual information [14]–[16]. Once again motivated by outstanding results in the field of information retrieval, RSs leverage KGs to enhance their performance and overcome several issues.

Knowledge graphs are a practical approach to representing large-scale information and heterogeneous data, where their main atomic components are nodes and edges. The nodes are used to represent the different entities, while the edges serve as links between these entities and denote the relation between them. The knowledge is then stored in a sort of triplet (head entity, relation, tail entity), also known as a fact. For example, (rabat, capital of, Morocco) indicates the fact that Rabat is the capital of Morocco; Figure 1 is a simple representation of this triple; in a formal way, a KG G can be represented as the following:

$$G = (h, r, t) \,/\, h, t \in E \text{ and } r \in R \tag{1}$$

Where E is the entity set and R is the relation set.

The efficiency of KGs was demonstrated in different domains, whether for text analysis [17], question answering [18], or recommender systems, where some works were delegated to enhance the general performance, while others were guided to more specific issues such as cross-domain recommendations and explainability. Major achievements were also made to mitigate the issue of data sparsity.

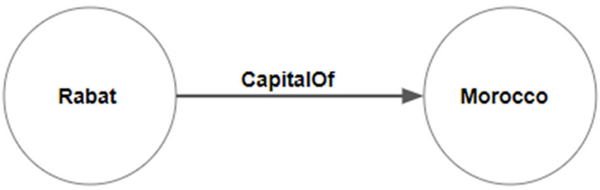

Fig. 1. A simple representation of the fact triplet

Data sparsity in RSs is a logical consequence of the information explosion, where the total of available items is much bigger than the items that any user has interacted with, which leads to big gaps and difficulties in modeling the users' preferences, hence inaccurate results on what could fit users the most. To illustrate the impact of this phenomenon, let's take the user-item interaction data, represented by a user-item matrix M, where each entry $M(i, j)$ of the matrix represents the interaction (with good or bad feedback) of the user i with the item j. In a sparse scenario, the matrix M is sparsely populated, with many missing values denoted as $M(i, j) = \phi$ to indicate the absence of interaction for most item-user pairs. Where the sparsity could be calculated based on [19] as the following:

$$sparsity = \frac{\#nonEmpty\,entries}{\#entries} \tag{2}$$

As a result of the lack of interaction with many items, accurately inferring user preferences becomes challenging [20]. Multiple studies addressed this issue using deep learning techniques [21], [22], or even including side information such as tags [23], and more recently, KGs, as it will be explored later in this work.

3 METHODOLOGY

The present study follows the protocol shared in [24] to conduct a transparent, rigorous review of previous research efforts. It defines the review's scope, develops a search strategy, selects primary studies, assesses their quality, analyzes data, and presents findings in a structured manner as illustrated in Figure 2.

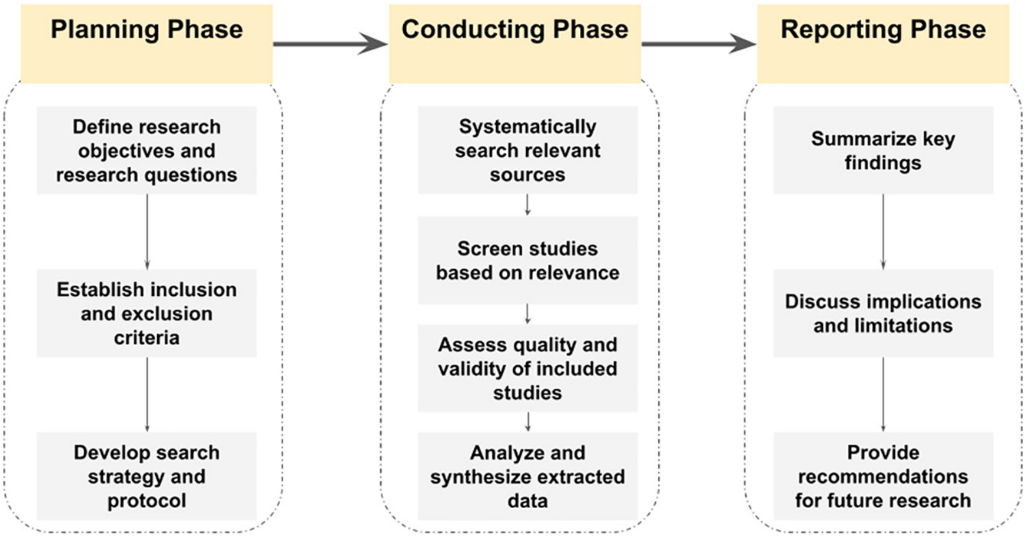

Fig. 2. The SLR in phases

3.1 The need for the systematic review and research questions

This work aims to offer a comprehensive survey for KG-based RSs, addressing the data sparsity challenge. To the best of our knowledge, among the many studies that perform SLRs in the domain of recommendation systems [25]–[28]; only a few addresses KG-based RSs [29], but none of them conduct such a specific investigation in a systematic way to treat the issue of data sparsity in these hybrid systems. To this end, several key research questions were defined to deeply investigate how KGs are used to mitigate the data sparsity in recommendation systems, as presented in Table 1.

Table 1. Research questions

Research Questions	Objectives
Q1. How do KGs help to overcome the issue of data sparsity in recommendation systems?	To explore the usage of KGs and their components, what type of knowledge is learned?
Q2. Which approaches are used in the knowledge graph-based recommendation area for data sparsity?	To explore the different approaches used in KG-based RSs to overcome data sparsity
Q3. What KGs are used to evaluate knowledge graph-based RSs with sparsity issues?	To understand how far the models cover real-world KGs
Q4. What evaluation procedures are used, and what metrics are employed?	To identify how each work assesses the effectiveness of the proposed solution

3.2 Search strategy

To capture all the work done in the field, an enumeration of all the candidate keywords was done to build a generalized query and retrieve all the needed information with no exclusion. The query should then be a combination of three axes:

1. **Recommendation:** recommendation system, recommender system, recommendation technique, recommendation engine, user modeling, personalization.
2. **Knowledge graph:** knowledge graph, semantic network, relational learning, knowledge-aware.
3. **Sparsity:** sparsity problem, sparse data, data sparsity, limited content, spareness, data scarcity.

The final used search string is as follows:

(“knowledge graph” OR “semantic network” OR “relational learning”) AND (“sparsity” OR “Spare” OR “spareness”) AND (“recommender” OR “recommendation”)

The chosen query is characterized by its generality and its ability to fetch exhaustive results. The search was performed in four electronic data sources known for their extensive coverage and quality of sources, which are Scopus, ACM Digital Library, IEEE, and Science Direct.

3.3 Selection of primary studies

By running the search query through the diverse scientific digital sources, 459 papers were retrieved. To objectively evaluate whether to proceed with each study or not for further processing, we have defined a set of inclusion and exclusion criteria stated in Table 2. These criteria help us distill the most relevant studies, along with carefully analyzing the abstract of each work.

Table 2. Selection criteria

Inclusion Criteria	Exclusion Criteria
Relevance of the paper based on the questions (Q1 to Q4)	Papers do not address the KG-based RSs.
Papers from journals or conferences	Papers that address KG-based RSs but not data sparsity
Papers published between 2013 and 2023	Papers lacking detailed information (only abstracts, presentations, etc.)
Language: English	Internal reports

3.4 Quality assessment

To enhance the robustness of the SLR, we have also assessed the quality of the selected studies based on the questions enumerated in Table 3. We choose to denote each question with three possible values: 0, 0.5, and 1, which represent answers "no," "partially" and "yes" respectively. To reflect the fact that not all questions have the same degree of importance, we have settled on assigning coefficients to weight them.

Table 3. Quality questions

Quality Assessment	Weight
Are the motivation and contributions clearly described	1
Did the work review the previous attempts devoted to dealing with data sparsity?	0.5
Did the paper provide a solution for sparsity issues using knowledge graph-based recommender systems?	1.5
Is the proposed solution clearly explained and can be implemented	2
Is there a discussion about the results or provide an empirical evaluation?	1.5
Are the limitations explicitly discussed or recommend any further research?	1

The final quality scores were calculated using the following equation:

$$Score = \sum_{j=1}^{6} v_j * w_j / 7.5 \tag{3}$$

Where, w_j is the weight of question j and v_j is the vote for question j.

Following our assessment and including studies with a score equal to or greater than the average, we ended up with a set of 51 papers that were ultimately included.

3.5 Data extraction strategy and synthesis

Once the relevant data have been extracted to answer our research questions cited in Table 1, the process of data synthesis begins, which requires the consolidation and summarizing of the findings. To establish a coherent and consistent synthesis, narrative techniques are used to describe the results and explore the different key themes across the selected papers. As we conceptualize clustering the studies to offer clear classification from different aspects. Furthermore, to convey the results, the findings will be represented in a descriptive statistical manner.

4 DATA EXTRACTIONS RESULTS

4.1 Overview of the reviewed literature

This section represents a general review of the selected studies, where the papers are analyzed based on their publication year and application domain. Figure 3 shows the evolution of papers through the selection process, while Figure 4a demonstrates the increased interest in the topic over the years. KGs were introduced for the first time by Google in 2012 [30] and started gaining popularity from then, and more interesting studies were published to deal with data sparsity in RS starting in 2017.

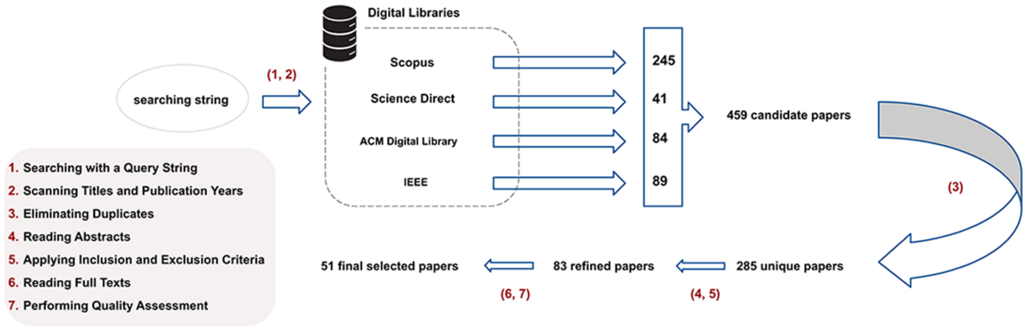

Fig. 3. Papers' selection in steps

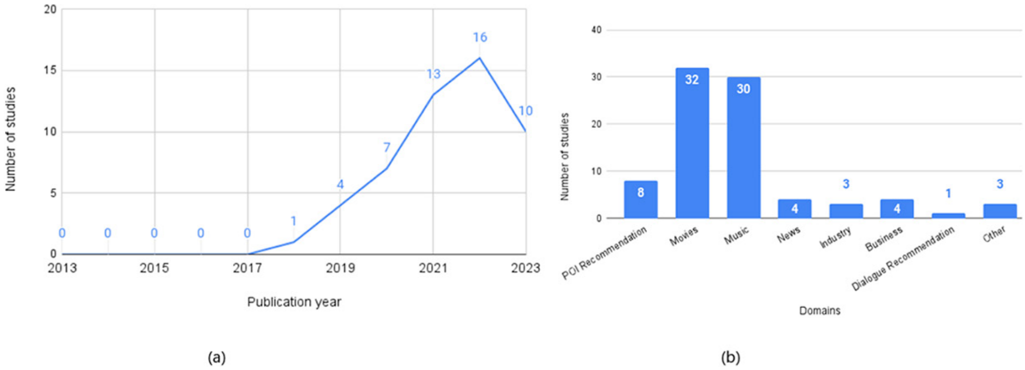

Fig. 4. Distribution of studies by publication year and their application domains

The proposed approaches have been tested in different recommendation contexts, as illustrated in Figure 4b. We have also investigated the quality of the selected studies over the years, as shown in Figure 5a, while the distribution of the works across the four data sources is represented in Figure 5b.

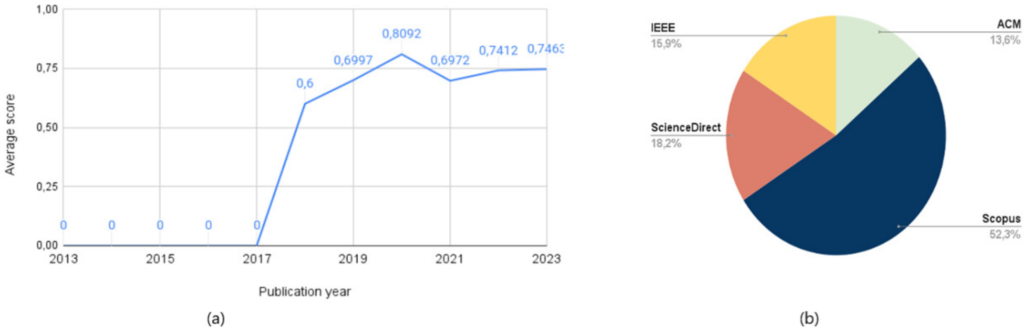

Fig. 5. Average quality scores over the years and their distribution by data source

4.2 How do KGs help to overcome the issue of data sparsity in recommender systems?

Data sparsity is inevitable, but techniques such as using KGs as side information can improve recommendation performance in sparse scenarios. As demonstrated in

Figure 6, with an example of movie recommendation based on KG, the system will recommend Movie 4 along with Movie 3, even though it is only seen by User 3 and there is insufficient interaction with it.

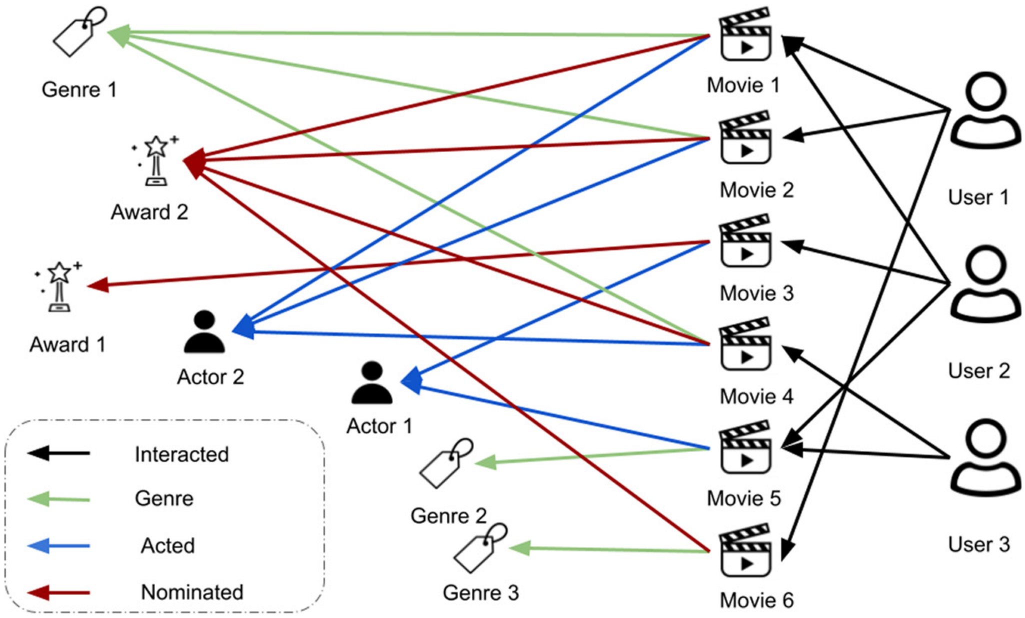

Fig. 6. Illustration of the importance of KGs within sparse scenarios

Knowledge graphs were designed to capture multiple types of knowledge in the form of entities and relations. Table 4 classifies studies according to the type of the main knowledge leveraged, whereas Figure 7 highlights their distribution.

Table 4. Types of knowledge leveraged in KGs to mitigate the data sparsity issue

Type	Description	Studies
Semantic Knowledge	Focuses on learning entity concept-level information by mining the meaning and interactions among entities in a KG.	[31]–[73]
Temporal Knowledge	Learning how user-item interactions evolve over time and capturing the dynamics of user preferences.	[35], [50], [52], [60], [61], [74], [75]
Spatial knowledge	Modeling user preferences in terms of location-based information.	[32], [35], [50], [61], [72], [74], [75]
Social Knowledge	Used to capture the influence of social factors on user preferences.	[35], [50], [74]
Structural knowledge	Focuses on learning graph-level information by mining the graph structure itself, capturing patterns, topologies, and connectivity entities in the graph.	[31], [32], [34], [36]–[49], [51]–[65], [67]–[72], [74]–[76]

Considering items as one of the main components of KGs, [37] handles sparsity using an aggregation layer of user-interacted items for feature smoothing and employs graph filtering to avoid overfitting. While [77] uses manifold regularization to smooth transitions between adjacent graph entities and maintain the smooth distribution of the mapping function even in sparse scenarios.

To enhance the user-specific preferences, [52] uses short-term to capture users' recent behavior and KG to model contextual knowledge from, thereby providing a more comprehensive representation that can handle data sparsity. While [58] enhances user preferences by utilizing feature-based and structure-based information in the KG, enabling reliable recommendations in sparse settings. On the other hand, [63] employs meta-path-based similarity measures to identify high-order associations between users, items, and attributes, even in sparse interactions, and enrich users' preferences.

In an attempt to incorporate more user characteristics and not only focus on item enhancement, [34] proposes using static user features to model new users and applying a relation-type attention network to enhance item modeling. User's modeling is categorized into static profile modeling and implicit preference modeling. Static profile modeling uses a constructed user graph to capture similar user information for new users, while implicit preference modeling uses historical interactions.

The inter-task relationship between the RS and the KG, which connects items and entities, served as the starting point for multiple works, where [31]–[33], [51], [56], [64], [65], [67], [68], and [76] propose multi-task learning approaches to support the recommendations task with KG representation task, where the heterogeneity of KGs helps to mitigate the limited item-user interaction. [43] integrates deep learning to empower the sparse item-user interaction by capturing the relation between it and the item-item relations using a cross-information sharing layer. While [53], [55] adapt the link prediction and KG completion tasks to assist the recommendation task.

The nature and characteristics of KGs in modeling data in terms of triples allow them to also be merged and cooperate based on common and bridging entities, as manifested in [70] and [71], where data sparsity was handled by leveraging the extracted knowledge from multiple and cross domains.

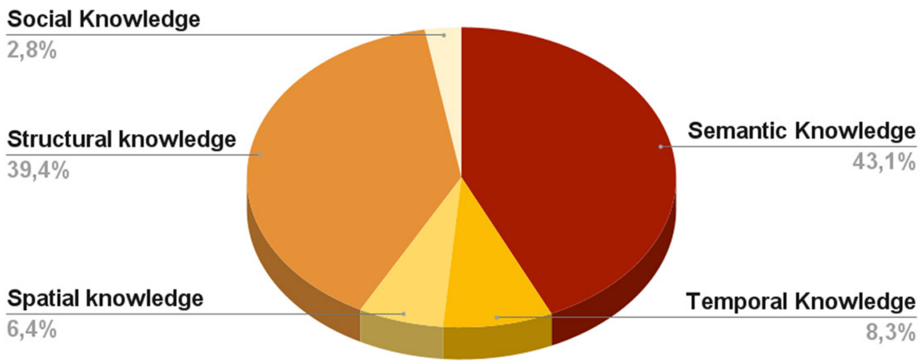

Fig. 7. Knowledge type distribution by article

With the same base idea of exploiting the heterogeneity of KGs, [61] proposes a spatio-temporal relation path embedding to learn the spatial correlations between Point-of-Interests (POIs), with the temporal dependencies between them. On the other side [35] and [75] joined location and time as a relation of the triple, focusing on user check-in sequences without introducing external attributes or points of interest to avoid sparse data. Moreover, [35] exploits correlations between cold-start users, POIs, spatiotemporal contexts, and their warm-start counterparts to provide additional training data to treat the cold start and data sparsity issues. While [50] and [74] focus on overcoming the data sparsity by mining the hidden knowledge in interactions between users and/or POIs represented not only in spatial, and temporal but also in social context, the proposed framework by [74] aims to construct personalized sequential KGs from users' check-in sequences, and to learn fast from sparse training data, a meta-learner was introduced to generate user-specific parameters based on check-in sequences.

To address the data sparsity in group recommendation, [57] use a graph convolution network (GCN) to capture hidden knowledge in group recommendations. They map each item to its equivalent entity in the KG, incorporating an attention mechanism to determine the influence of each user in the group based on user-user interaction and candidate item influence. Using the same approach of aggregation, [62] treats the data sparsity in individual recommendations by learning the complex relationships between users, items, and their attributes. While [66] used GCNs to capture inter-item and inter-user correlations, incorporating an attention mechanism to assess similarities between candidate items and users.

With the aim of overcoming the data sparsity issue, [36] introduces knowledge-grounded user and item representations from KG to improve user interests in sparse interaction scenarios. The first representation encodes a user's interest in a specific item, while the second encodes the item's attributes for additional knowledge. The model focuses on generic or grounded representations based on sparsity.

To seek enhanced representations of users and items, [59] uses graph propagation to capture the dependencies and relationships between users and items. Additionally, the model generates item-aware user and user-aware item representations by adaptively selecting relevant information from both ends of the user-item graph through an attention mechanism. While [72] exploits rich ontology and multidimensional representation to construct a tourism KG, allowing understand tourists' behaviors based on them and capturing them using the preference-propagation paradigm.

Focusing on items' attributes [73] tackles the data sparsity by integrating global and local interest features that are often overlooked by other approaches and focusing on the target item's neighborhood without considering the impact of other items, while [44] enhances sparse multiple user-facet interactions; instead of a simple concatenation of user preferences across multiple facets, the model utilizes hidden knowledge correlations by generating a KG to model facets and their relationships. [60] uses a recurrent neural network (RNN) to analyze the sequential and transition regularities of similar locations, enabling the model to understand complex and nuanced relationships. The model also features an attention mechanism to learn from relevant information and reduce noise.

To empower the connectivity of KG, [46] merges bipartite user-item interaction graphs and KG into a collaborative KG, enhancing deep connectivity by linking similar users and items if similarity passes a threshold. Three aggregators were used to reduce feature sparsity: a node aggregator, a relation aggregator, and a type aggregator to capture high-level features and long-range dependencies between node types in the graph.

Other approaches benefit from the knowledge expressed in the connectivity patterns of KGs and their characteristics to deal with data sparsity. But injecting non-linear topologies into linear ones will automatically lead to information loss and not exploit the full power of KGs. In this direction, [38] tries to enhance the embedding in sparse scenarios by characterizing the user-item interactions into subgraphs. The subgraphs (u, i) were constructed by conducting random walks, then exploring these subgraphs to learn first the entities' embedding by encoding both the semantics and topology of the subgraph, followed by an aggregation of the embedding of the most relevant entity to get the holistic subgraph. Focusing on the semantics of properties, [39] constructs property-specific subgraphs to model each relation independently. To overcome the fact that numerous relations usually link one item to a few entities, which leads to poor knowledge excretion, [39] proposes to integrate users' feedback as a pivot relation in every single constructed subgraph, wherein it helps to empower connectivity as well as explore in a direct way user-item relevance for all properties. The embedding of the graphs are aggregated to compute the global relatedness used for recommendation.

Aligned with the idea of exploring the high-order connectivity in KG, [40], [48] support the sparse item-user interaction by learning a meaningful user and item representation, where item embedding is generated from multi-hop relation paths in the KG. The models also focus only on the most informative. While [60] improves user modeling by considering indirect interactions and potential interest in un-interacted items, it computes similarity-based items rated by one user and common items rated by both, reducing the impact of insufficient interactions and enabling preferences for unobserved and unexposed items. The model enhances item representation by incorporating multi-hop neighbors, recognizing their role in capturing valuable information, and expanding relationships based on semantic similarity between different-order relations.

To preserve the semantics hidden in a long-term path that is ignored in traditional RNNs, [41] uses a residual recurrent network (RRN) on relational paths constructed from entities and relations instead of relying only on the current input and the previous hidden state to predict the next element, which does not model the structure of KGs' paths and their triplets and does not distinguish between entities and relations. Additionally, the model can mine the different user interaction behaviors. and learns the complex nonlinear interaction between users and items, which helps mitigate the data sparsity issue. [42] uses a self-supervised learning (SSL) approach to enhance a model by utilizing users' active feedback through conversation, overcoming limited dialogue samples, and extracting structural knowledge based on knowledge graphs.

Another paradigm that was leveraged to enhance the representation of the different components is the contrastive learning employed in [78]–[80] by mapping similar entities close to each other, capturing more shared knowledge, and minimizing the similarity between negative ones. The model is forced to effectively distinguish between entities based on their similarity and not overly generic representations.

While all the above-mentioned works concentrate on using KG as auxiliary information, [45] considers the KG as the core component by conducting schema-based reasoning as a compilation task for KG, which helps increase the amount of the training data. Along with treating the long-tail recommendations as a few-shot relational learning problem using a multi-head attention-based meta-relational learning model to capture the existing higher-level patterns and correlations between these relations and make recommendations based on the few available interactions.

Overcoming the data sparsity with KG completion was the motivation behind the model of [47] by using a transudative learning approach to model links between unknown items, taking advantage of the fact that KGs are dynamic and not static, which further enriches the embedding of the entities by leveraging information from the extrapolated KG. While [49] enhances representation capability by weighting vector representation dimensions, with more informative dimensions given greater weight, and using Mahalanobis distance to account for covariance between dimensions. Some works took the RSs one step further by treating the data sparsity not only in terms of user-item interactions but also in terms of KG itself, where [81] treats the long-tail distribution of entities and [54] deals with the over-smoothing.

4.3 Which approaches are used in the knowledge graph-based recommendation area for data sparsity?

Owing to their heterogeneity and flexibility, KGs were used with different approaches in RSs to offer accurate and effective recommendations. These approaches can be divided into four key categories.

Translation-based methods. Translation-based methods aim to encode the semantics hidden behind entities and relations; these methods generally capture the local knowledge by treating the triplets or facts (h, r, t). Where Where [35], [50] joined location and time as a relation to represent triplets as the following: (u, $< t, l >$, v) where u is the user, $< t, l >$ is the relation, and t: time, l: location, and v: points of interest (POI), in a way that the spatiotemporal transition among POIs and users is encapsulated by the following translation $\overrightarrow{u_{tl}} + \overrightarrow{tl} = \overrightarrow{v_{tl}}$. Where $\overrightarrow{u_{tl}}$ and $\overrightarrow{v_{tl}}$ are the projected vectors of users and POIs in the spatiotemporal context space, and $\overrightarrow{tl}$ is the embedding of spatiotemporal context. On the other hand, [49] also employed the relation as a translation from the head to the tail, but they improved their score method by using the Mahalanobis distance [82] instead of the Euclidean distance. Table 5 highlights the techniques used.

Table 5. The techniques used in translation-based studies

Work	Distance	Score Function
[50], [35]	Euclidean distance	$s_{tl}(u, v) = \left\Vert \overrightarrow{u_{tl}} + \overrightarrow{tl} - \overrightarrow{v_{tl}} \right\Vert_2^2$
[49]	Mahalanobis distance	$f_r(h, t) = (h + r - t)^T W_r (h + r - t)$ where the weight matrix $W_r = L_r^T D_r L_r$, L_r is the transformation matrix, and the diagonal matrix D_r the corresponding weight values of the corresponding dimensions

Path-based methods. Path-based approaches focus on leveraging the structural information in KGs by encoding paths or sequences of relations between entities in KGs to infer unseen or relevant connections and make personalized recommendations. [63] uses the constructed meta paths to assess the similarity between users and offer a personalized recommendation. This work relied on two types of similarity measurement:

- Single-path similarity: to reveal the similarity of users under a single path, along with considering the importance of each feature. This similarity is obtained with:

$$simM = \sum_{V \in V_u} \alpha_i * V_i \left(\sum_{i=1}^{N} \alpha_i = 1 \right) \tag{4}$$

Where, i is the number of dimensional features, V represents the similarity of two users in a particular dimension, and α is the weight value in each dimension feature.

- Multipath fusion similarity: to calculate the similarity between users based on multiple paths.

$$simFSCM = \sum_{M \in M_n} \beta_i * simM_i \left(\sum_{i=1}^{n} \beta_i = 1 \right) \tag{5}$$

Where, $simM_i$ is the similarity of users in each path and β represents its weight.

Propagation-based methods. The popularity of deep learning-based approaches is due to their ability to learn high-order representations and patterns; several recent deep learning techniques have been applied in KG-based RSs. Where MLP (multi-layer perceptron) learning was employed by [31], [33], [51], [53], [56], [58], [65], [67], and [68] to learn more complex nonlinear dependencies and KG representations. Cross & compress units were used as bridges in [31], [53], [56], [58], [65], [67], and [68]. In these multi-task models, the cross and compress modules were designed to transfer knowledge between the different tasks and learn their relevance, as shown in Figure 8.

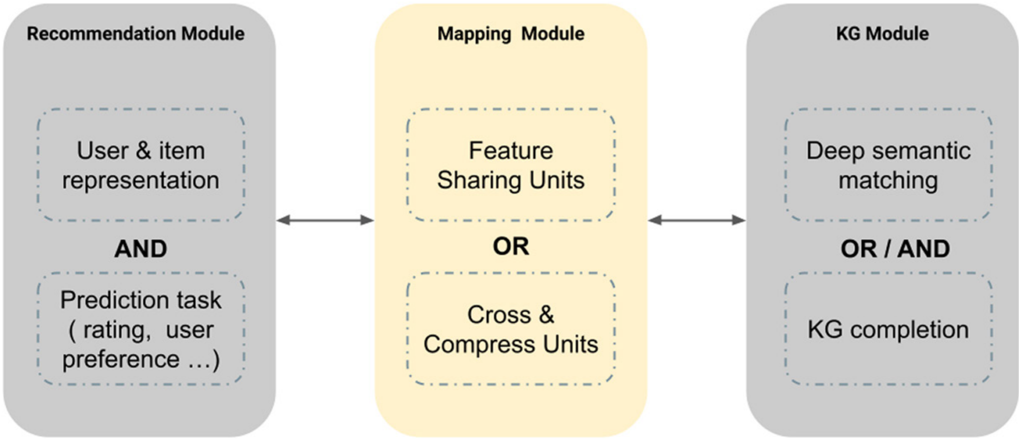

Fig. 8. High-level overview of multi-task learning KG-based RS architecture

Other works utilize graph convolutional network-based methods to encode their graphs [34], [37], [40], [42], [46], [54], [57], [59], [66], [71], [73], [76], [78], and [80]. Where [71], [78], [80] strength the representations with contrastive learning and minimize noisy interference within the KG, while [42] utilizes a very specific type of GCN, the relational graph convolution network, to learn structural knowledge. [62] implements the model based on neural graph collaborative filtering. [41] develops an RNN network with an attention mechanism to learn the sequential dependencies between entities and relations along item-user connecting paths. [43] proposed a framework with four layers in their work: a cross-information sharing layer to share knowledge between entities and item embedding; a deep neural network (DNN) layer to improve the model's generalization ability and learn high-level features; a knowledge-enhanced network (KEN) layer to capture user preferences; and a deep ripple network (DRN) layer to map the interactions of the DNN layer and KEN layer. [47] combines KG extrapolation networks with transudative learning to enrich feature representations. [48] uses different propagation strategies in their model to assist the recommendation process. [52] worked with GRU (gated recurrent units), as it's an extension of RNN that addresses the vanishing gradient problem and allows to capture short-term and long-term preferences. [72] applied a preference-propagation paradigm based on KG inspired by Ripplenet [83]. [74] enforces their RNN model with a meta-learning technique to model the user-specific sequential behavior even in sparse scenarios. On the other hand, [34], [37], [40], [42], [46]–[48], [51]–[53], [57]–[59], [62], [66], [68], [72], and [76] incorporate attention mechanisms to guide the proposed models to focus on learning the most relevant knowledge. Table 6 classifies these works based on their attention scope.

Table 6. A general classification of the employed attention mechanisms

Scope	Research Work: Attention Strategy
Entity level	[51], [70], [72], [74], [77]: graph attention, [38], [42], [68]: self-attention, [47]: attentive entity propagation, [48]: asymmetrical semantic attention, [52]: context entity attention, [53]: multi-head attention, [59]: multiple attentive propagation layers, [60]: salient neighbors attention, [57]: attentive users' influence among a group, [66]: neighbor aware attentive mechanism, [58]: neighbors' information aware attention aggregator, [40]: user-aware item aggregator, [67]: users' historical behavior aware attention aggregator
Triple-aware	[41]: KG-aware attention mechanism
Context-based path	[37]: self-attention mechanism
Relation level	[34]: user-item attention, [75]: short-term relation attention, [45]: multi-head attention, [46], [81]: relation-aware attention aggregator, [40], [54]: relation type-aware attention, [55]: relation-aware attention aggregator, [62], [76] user-relation aware attention mechanism

Hybrid methods. Hybrid methods are a promising solution that combine multiple approaches. Some studies have integrated translation-based and deep learning-based methods; [32] employs the KG embedding (KGE) module of their multi-task model TransR for entities and relations embedding, then feed them into a K-layer MLP to predict tails of triples. While [75] learns users' preferences by adopting the TransR and GRU methods. [44] designed an autoencoder model with two main components: the encoder to learn the different representations using TransE and the decoder to predict ratings using a multiple-layer neural network. For semantic representation improvement, [81] employs a training approach that alternates between the relation-aware knowledge aggregator and the TransE model.

Other works incorporate both path-based and deep learning-based methods; [38] and [39] extract the meaningful and informative subgraphs that characterize the connectivity of user-item, while [38] generates paths for user-item pairs (u, i) using random walks with a maximum depth of six. Only paths leading to item i are retained, and these paths are assembled to create a subgraph. Subgraphs are encoded using a hierarchical attentive subgraph module, which uses a layer-wise propagation mechanism and aggregates embedding for the subgraph embedding.

[39] creates property-specific subgraphs using the random walk technique too. Thereafter, the model computes property-specific embedding by applying node2vec to property-specific subgraphs. [60] proposes an attentional recurrent neural network with applying random walk to design meta-paths. [55] combined the three approaches, where the model used TransD to obtain entities and relations embedding, a breadth-first search algorithm for subgraph construction, node labeling to facilitate capturing the structural knowledge, and finally a GNN to learn structural characteristics and perform the relation completion task. [70] used TransD with the GNN to capture multi-type relations with fewer training parameters and allow the model to be applied to large-scale KGs in cross-domain RSs, and [79] empower these techniques with multi-contrastive learning. While [69] combined the GNN with RotatE to gain more structural knowledge, [77] proposed a semi-supervised method by initializing the attentive KG neural network with translation-based pre-trained embedding using TransE [84]. Further research draws upon the approaches of matrix factorization and deep learning, where [36] used matrix factorization to generate generic representations of items and users and supplement them with knowledge-grounded representations using encoders.

4.4 What KGs are used to evaluate knowledge graph-based RSs with sparsity issues?

Introducing KGs as a key solution to alleviate data sparsity in RSs is the main idea behind all selected studies. In the following analysis, we will focus on scrutinizing the works in terms of the used KGs to identify the most frequently employed, their construction method, and their sparsity level. The used KGs can be classified as follows:

- **Generic KGs:** used in the first place to model entities and their relations in the form of triplets.
- **Spatial KGs:** extend traditional KGs by incorporating spatial relationships between entities, including geographic categories, locations, and distances.
- **Social KGs:** used to capture social connections, relationships between users, and social influence.
- **Spatiotemporal KGs:** combine both spatial and temporal information to mine the evolution of knowledge over time in different locations.
- **Domain-specific KGs:** delegated to a specific domain with the aim of capturing domain-specific knowledge.
- **Ontology-based KGs:** These KGs are built based on ontologies with a well-defined structure, including properties, relations, and concepts.

The approaches have been evaluated on real extensions and instances of the adopted class. Figure 9 illustrates the data sparsity level of each used dataset.

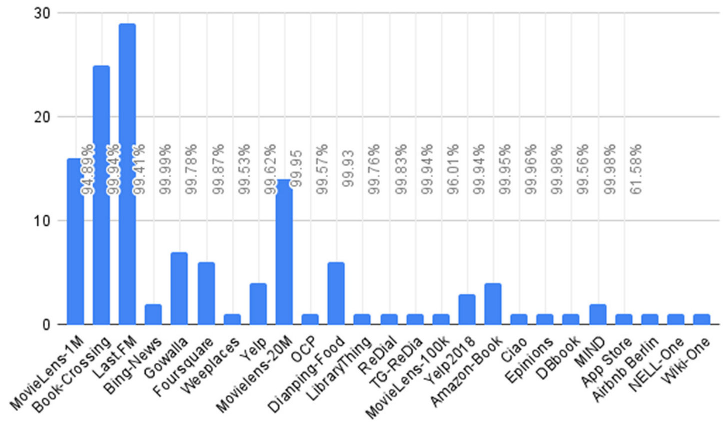

Fig. 9. The sparsity of the utilized datasets

4.5 What evaluation procedures are used, and what metrics are employed?

After exploring the selected approaches dedicated to mitigating the sparsity of data and the limited content from different aspects, it's crucial to investigate their efficiency and analyze how well they performed in such cases. The majority of works assess their models in two major experiment scenarios; the first experiment focuses on evaluating the performance of click-through rate (CTR) prediction, while the second pertains to top-K recommendation, where the studies ran multiple tests, in which k varied generally between K = 2, 5, 10, and 20. Several evaluation metrics were used that could be classified as the following:

- **Classification metrics:** It assess the performance of the model, since the recommendation task can be seen as a binary classification; recommended or not recommended. Where the accuracy (ACC) was employed in 53% of the works to measure the models' confidence by calculating the proportion of correct predictions to the total predictions, AUC or Area under the ROC Curve represents the most widely used indicator with 45% as a usage percentage; it is used to analyze the overall performance and not just at a specific classification threshold. Recall is used in 49% of the cases to check the systems' ability to capture all relevant recommendations. Precision in 21.5% of the studies to check if the recommendations are truly relevant for users. F1 was used in 29.5% of the papers to represent a balanced evaluation that considers both completeness (recall) and correctness (precision). It evaluates how well the system can gather all relevant items at the same time and provide reliable suggestions.

- **Ranking metrics:** It evaluates how well the model can provide relevant recommendation lists and assess their ranking quality. Hit was employed in 17.5% of the works to assess the models' ability to provide relevant suggestions within the list, the Normalized Discounted Cumulative Gain (NDCG) was checked in 17.5% of the papers to quantify the system's ranking quality based on the relevance of the items and their position in the list. Mean Reciprocal Rank (MRR) was calculated in 7.5% of the approaches to offer an overall assessment of the system's performance in terms of identifying the first relevant item from the top K suggestions.
- **User Engagement metrics:** PVCTR (CTR on page views) and UVCTR (CTR on unique user views) indicators are used in [43] to evaluate the performance of the system in CTR prediction by measuring the degree of interest and user interaction with the provided recommendation.
- **Serendipity and novelty metrics:** It is used in [39] to evaluate the system's ability to generate diverse and unexplored recommendations. Where serendipity is used to check the capacity of the model to recommend unexpected but relevant items. The measure of novelty is intended to assess if the model can recommend unknown items for the user.
- **Rating metrics:** It assess the capacity to predict users' preferences by quantifying the prediction errors. Two rating metrics were used in this sense: MAE (mean absolute error) and RMSE (root mean square error).

Table 7 provides a classification method based on the evaluation type.

Table 7. Classification of the used metrics based on evaluation type

Evaluation Type	Description	Metrics
Offline evaluations	Analyzing the performance by using pre-collected historical data	AUC, ACC, Precision, Recall, F1, RMSE, MAE, MRR, NDCG, Hit, SER, NOV
Online evaluations	Analyzing the performance based on the users' direct interactions and evaluating feedback in real-time	PVCTR, UVCTR

In addition to using sparse datasets and to further examine the models' sensitivity to sparse scenarios, some studies have manipulated the training ratio r to replicate different levels of data sparsity and evaluate the model's performance under varying degrees of data sparsity. In the experiments, the ratio was set from 100% to 10% of the training sets, while the validation and test sets were kept fixed. This allows for fair settings and a controlled comparison across different training set ratios as they are conducted on the same set of data; moreover, fixing the validation and test set while reducing the training set represents real-world scenarios where the models are trained on limited data but then applied to new and unseen data, so reducing them means that the simulation is not adequate to represent the diversity of user-system interactions anymore. Figure 10 summarizes how much the data sparsity affects the performance of each work that has conducted this experiment, regardless of the used dataset, since the impact of the data sparsity is identified by calculating the percentage of the decrease in the AUC score.

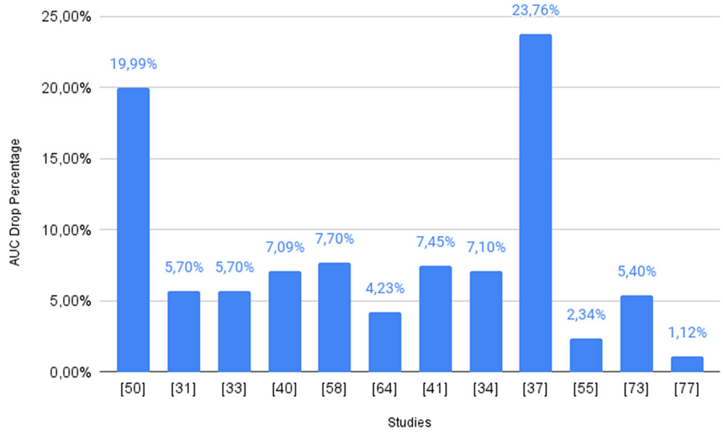

Fig. 10. Sensitivity to data sparsity analyzed by AUC drop percentage

5 RESULTS AND DISCUSSION

This section represents an interpretation and analysis of findings in depth, the results were gathered in one global entry view, as represented in Figure 11.

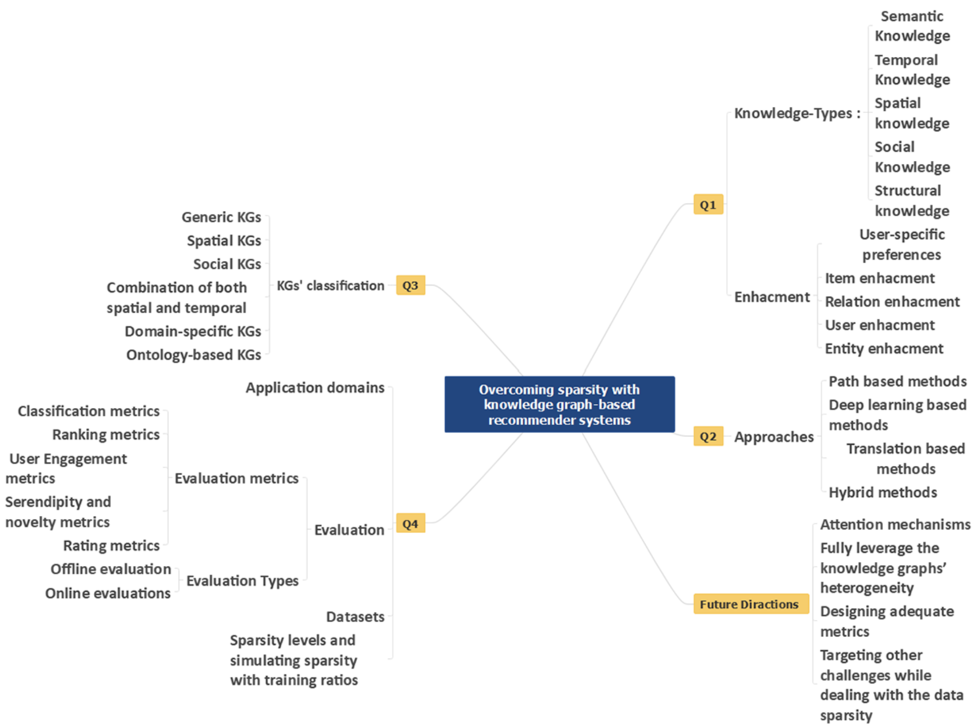

Fig. 11. High-order analysis overview

5.1 Interpreting the findings

Through this review, we seek to better understand the actual impact of leveraging KGs in RSs to mitigate the data sparsity. Where an overwhelming majority of the works employed KGs as side information, which is justified by the heavy reliance on user-item interactions in the domain of recommendation in general. The studies have used KGs as a secondary source of knowledge to support the recommendation task in sparse conditions. Only [42] has considered the KG as the main component of its approach. Dealing with KGs as side information does not diminish their effectiveness. In contrast, it could be seen as an advantage as it offers the possibility of employing multiple techniques and gives the freedom to apply the most suitable solution depending on the context. Where the best results were achieved in studies that used deep learning techniques due to their ability to learn higher-order representations and reveal hidden knowledge. Another valid reason for incorporating KGs as supporting resources is that items are parts of both item-user interactions and KGs; on this basis, researchers have proposed to jointly learn from them. As it was explored in Q1 and Q2, new trends are being used to deal with the sparsity, such as few-shot relational learning, meta-learning, and transfer learning.

From Figure 10, the research works that have shown more stability even in sparse settings are the ones using multitask techniques, which is justified by the ability of exchanging knowledge between the different tasks with the mapping module during the joint learning, where mainly items link the tasks and allow a joint learning to enhance the performance degraded because of the sparse rating matrices. Moreover, there is a growing interest in using attention mechanisms with different scopes in recent studies as an efficient solution to tackle the data sparsity, where the works that have integrated the attention mechanism outperform as they focused on the most relevant knowledge for the user and quantified the importance of the data using score functions and normalizing the scores with the SoftMax function in order to interpret the values as weights and reflect their relevance to the user [85]. Contrastive learning is also one of the promising approaches where the augmentation of data reduces the high dependency on sparse data by generating multiple views of the same data point and trying to learn its representation based on the generated views that are considered as similar or positive samples since they were generated from the same instance. However, all the works considered the rest of the data as a negative pair or dissimilar samples that should be placed further away in the embedding space, but it's not always semantically true. As a result, more efforts should be devoted to proposing approaches to define the data used as positive and negative pairs during the learning process.

It is essential to bring attention to the fact that the heterogeneity of KGs also plays a crucial role in mitigating the data sparsity, where this diversity helps to fill the gaps resulting from the limited data, especially when using more than one class of knowledge. Noting that semantic and structural knowledge are widely used among the selected studies, while temporal and spatial knowledge are only used in the context of point-of-interest recommendations, even if it is not less important to understand how the user behaves over time and location in other recommendation scenarios. Another way to benefit from the heterogeneity of KGs that was strongly missed by the filtered set of studies is the type of data where the papers focused on short textual items' features and attributes without dealing with long descriptions, users' feedback, or even images and videos, which also represent large hidden territories of knowledge; therefore, exploiting multi-modal KGs [86] is highly recommended to improve the performance of the recommender system with limited user-item interactions.

To fully leverage the KGs' heterogeneity in the future, it is important to highlight the truth that none of the studies have integrated multiple KGs in their approaches; they opted to utilize monographs even if their heterogeneity highly allows them to be treated and exploited in parallel. One of the major limitations of the totality of papers is the evaluation part, where they assess their work with metrics used in traditional RSs and even classification and ranking tasks in general. The only works that use more specialized evaluation metrics are [43], which used PVCTR and UVCTR to assess the system in CTR prediction. Besides, they represent the only online evaluation used across all the studies. While [39] used serendipity and novelty metrics along with the other performance indicators. Therefore, these limitations can be considered as a promising research field to develop more metrics for KG-based RSs in general and for dealing with data sparsity in particular, and why not elevate the improvement even further by designing an exhaustive framework to assess their quality.

Finally, all the works have proposed sophisticated contributions to mitigate the data sparsity in KG-based RSs in one way or another, by enhancing users, items, entities, and relation representation or by leveraging the different knowledge types hidden in the graph. Despite this, just a minority of them try to address other challenges such as cold start, noise, and long tail issues. As a result, it is notable to state the importance of targeting other challenges while dealing with the data sparsity, such as explainability, user's preference dynamicity, cross-domain recommendation, and data quality assessment of the integrated KG [87], [88].

5.2 Limitation

It is crucial to note that the present work relies on a systematic process of study selection and analysis, which may have introduced certain biases, and it could not be taken as an exhaustive study since it covers only four data sources, which is logical due to the huge number of available digital sources. Moreover, valuable insights from non-English studies might have been overlooked, as might contributions that are not within the scope of the period from 2013 to 2023. Therefore, we highly suggest taking these limitations into account while interpreting this work.

6 CONCLUSION AND FUTURE WORK

This work represents a systematic literature review (SLR) of 51 contributions selected from four different data sources (ACM, Scopus, IEEE, and ScienceDirect). It offers a comprehensive review of the literature between 2013 and 2023. Where valuable insights are provided into how KGs are used to alleviate the data sparsity in recommender systems, the techniques used, and recent trends, besides uncovering possible future directions. Through this work, we aim to offer a concise and coherent summary of the existing literature in a systematic manner, which could help researchers in their future work addressing data sparsity in RSs. Where deep learning and hybrid techniques have proven their efficiency in mitigating the sparsity of data, especially the works that were implemented based on joint learning and contrastive paradigms. Forthcoming efforts could be directed towards leveraging the outcomes of this SLR, where researchers can dedicate more attention to the evaluation techniques and protocols. Most of the papers just used the general metric to evaluate their work, except for the works that tried to simulate different stages of the sparsity and check how their approaches would react. A possible enhancement is developing standardized evaluation measures for KG-based RSs to objectively assess the performance. This could be by comparing the performance of the proposed technique in different stages of sparsity or by assessing the overall error deviation for different sparsity levels in comparison with the results of the original dataset. In addition, more protocols need to be implemented to not only compare by sparsity levels but also compare by specific items or users for the same sparsity level to check if the approaches are also able to make diverse suggestions for the same users in all different sparsity levels, as serendipity and novelty are also crucial criteria in recommender systems. Future studies could also focus on incorporating multimodal KGs. From our analysis made through this review, we have constantly identified the need to fully leverage the heterogony of the KGs, as most of the works explored only the KGs to store items and their short descriptive attributes, ignoring the other sort of knowledge that can be extracted from users' comments, audios, images, and videos. Additionally, although the notable results of the multi-tasking works, they do not fully leverage the power of this learning paradigm as they represent bi-task approaches, and only two tasks were implemented in the different papers, except for one work that implemented three tasks; thereby, it's very important to investigate and include more tasks to support learning representations of users and items in sparse settings. Also, one of the interesting paths is to extend the idea of joining correlated relations and form new triplets based on them, since the only relations joined in the works were the time and location and mainly explored in POI-RSs. Studying the possibility of combining more relations in other contexts is a very promising direction, as it would help to mine more hidden knowledge and improve the quality of the different representations. Finally, contrastive learning is one of the newly developed trends that has shown great results with limited data. Consequently, more attention should be paid to these techniques by introducing new strategies for data augmentation that are more oriented to recommender systems; this could be by including creating views without the most popular items or removing the less important relations detected with attention mechanisms and getting inspiration from the outstanding results in NLP and computer vision fields.

7 REFERENCES

[1] G. Behera and N. Nain, "Handling data sparsity via item metadata embedding into deep collaborative recommender system," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 10, Part B, pp. 9953–9963, 2022. https://doi.org/10.1016/j.jksuci.2021.12.021

[2] R. Duan, C. Jiang, and H. K. Jain, "Combining review-based collaborative filtering and matrix factorization: A solution to rating's sparsity problem," *Decision Support Systems*, vol. 156, p. 113748, 2022. https://doi.org/10.1016/j.dss.2022.113748

[3] Y. Wang, J. Deng, J. Gao, and P. Zhang, "A hybrid user similarity model for collaborative filtering," *Information Sciences*, vol. 418–419, pp. 102–118, 2017. https://doi.org/10.1016/j.ins.2017.08.008

[4] B. K. Patra, R. Launonen, V. Ollikainen, and S. Nandi, "A new similarity measure using Bhattacharyya coefficient for collaborative filtering in sparse data," *Knowledge-Based Systems*, vol. 82, pp. 163–177, 2015. https://doi.org/10.1016/j.knosys.2015.03.001

[5] L. Guo, L. Sun, R. Jiang, Y. Luo, and X. Zheng, "Recommendation based on attributes and social relationships," *Expert Systems with Applications*, vol. 234, p. 121027, 2023. https://doi.org/10.1016/j.eswa.2023.121027

[6] M. J. Hazar, M. Zrigui, and M. Maraoui, "Learner comments-based recommendation system," *Procedia Computer Science*, vol. 207, pp. 2000–2012, 2022. https://doi.org/10.1016/j.procs.2022.09.259

[7] F. Wang, Z. Zheng, Y. Zhang, Y. Li, K. Yang, and C. Zhu, "To see further: Knowledge graph-aware deep graph convolutional network for recommender systems," *Information Sciences*, vol. 647, p. 119465, 2023. https://doi.org/10.1016/j.ins.2023.119465

[8] S. Zhang, N. Hui, P. Zhai, J. Xu, L. Cao, and Q. Wang, "A fine-grained and multi-context-aware learning path recommendation model over knowledge graphs for online learning communities," *Information Processing & Management*, vol. 60, no. 5, p. 103464, 2023. https://doi.org/10.1016/j.ipm.2023.103464

[9] G. Balloccu, L. Boratto, G. Fenu, and M. Marras, "Reinforcement recommendation reasoning through knowledge graphs for explanation path quality," *Knowledge-Based Systems*, vol. 260, p. 110098, 2023. https://doi.org/10.1016/j.knosys.2022.110098

[10] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work*, 1994, pp. 175–186. https://doi.org/10.1145/192844.192905

[11] M. Pazzani and D. Billsus, "Learning and revising user profiles: The identification of interesting web sites," *Machine Learning*, vol. 27, no. 3, pp. 313–331, 1997. https://doi.org/10.1023/A:1007369909943

[12] Y. Kilani, A. F. Otoom, A. Alsarhan, and M. Almaayah, "A genetic algorithms-based hybrid recommender system of matrix factorization and neighborhood-based techniques," *Journal of Computational Science*, vol. 28, pp. 78–93, 2018. https://doi.org/10.1016/j.jocs.2018.08.007

[13] A. A. Kardan and M. Ebrahimi, "A novel approach to hybrid recommendation systems based on association rules mining for content recommendation in asynchronous discussion groups," *Information Sciences*, vol. 219, pp. 93–110, 2013. https://doi.org/10.1016/j.ins.2012.07.011

[14] M. Afzal *et al*., "Personalization of wellness recommendations using contextual interpretation," *Expert Systems with Applications*, vol. 96, pp. 506–521, 2018. https://doi.org/10.1016/j.eswa.2017.11.006

[15] N. Idrissi, A. Zellou, O. Hourrane, Z. Bakkoury, and E. H. Benlahmar, "Addressing cold start challenges in recommender systems: Towards a new hybrid approach," in *2019 International Conference on Smart Applications*, 2019, pp. 1–6. https://doi.org/10.1109/SmartNets48225.2019.9069801

[16] S. Ouaftouh, A. Zellou, and A. Idri, "Social recommendation: A user profile clustering-based approach," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 20, p. e5330, 2019. https://doi.org/10.1002/cpe.5330

[17] B. Probierz and J. Kozak, "Knowledge graphs to an analysis and visualization of texts from scientific articles," *Procedia Computer Science*, vol. 225, pp. 4324–4333, 2023. https://doi.org/10.1016/j.procs.2023.10.429

[18] C. Liu, X. Ji, Y. Dong, M. He, M. Yang, and Y. Wang, "Chinese mineral question and answering system based on knowledge graph," *Expert Systems with Applications*, vol. 231, p. 120841, 2023. https://doi.org/10.1016/j.eswa.2023.120841

[19] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *Proceedings of the 10th International Conference on World Wide Web*, 2001, pp. 285–295. https://doi.org/10.1145/371920.372071

[20] M. Grčar, D. Mladenič, B. Fortuna, and M. Grobelnik, "Data sparsity issues in the collaborative filtering framework," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 4198, 2006, pp. 58–76. https://doi.org/10.1007/11891321_4

[21] W. Zhang, X. Zhang, H. Wang, and D. Chen, "A deep variational matrix factorization method for recommendation on large scale sparse dataset," *Neurocomputing*, vol. 334, pp. 206–218, 2019. https://doi.org/10.1016/j.neucom.2019.01.028

[22] N. Idrissi, O. Hourrane, A. Zellou, and E. H. Benlahmar, "A restricted Boltzmann machine-based recommender system for alleviating sparsity issues," in *2019 1st International Conference on Smart Systems and Data Science (ICSSD)*, 2019, pp. 1–5. https://doi.org/10.1109/ICSSD47982.2019.9003149

[23] J. Li, Y. Tang, and J. Chen, "Leveraging tagging and rating for recommendation: RMF meets weighted diffusion on tripartite graphs," *Physica A: Statistical Mechanics and its Applications*, vol. 483, pp. 398–411, 2017. https://doi.org/10.1016/j.physa.2017.04.121

[24] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Information and Software Technology*, vol. 55, no. 12, pp. 2049–2075, 2013. https://doi.org/10.1016/j.infsof.2013.07.010

[25] D. Roy and M. Dutta, "A systematic review and research perspective on recommender systems," *Journal of Big Data*, vol. 9, p. 59, 2022. https://doi.org/10.1186/s40537-022-00592-5

[26] N. Idrissi and Z. Ahmed, "A systematic literature review of sparsity issues in recommender systems," *Social Network Analysis and Mining*, vol. 10, 2020. https://doi.org/10.1007/s13278-020-0626-2

[27] C. Salazar, J. Aguilar, J. Monsalve-Pulido, and E. Montoya, "Affective recommender systems in the educational field. A systematic literature review," *Computer Science Review*, vol. 40, p. 100377, 2021. https://doi.org/10.1016/j.cosrev.2021.100377

[28] P. Alamdari, N. Navimipour, M. Hosseinzadeh, A. Safaei, and A. Darwesh, "A systematic study on the recommender systems in the e-commerce," *IEEE Access*, vol. 8, pp. 115694–115716, 2020. https://doi.org/10.1109/ACCESS.2020.3002803

[29] Y. Deng, "Recommender systems based on graph embedding techniques: A review," *IEEE Access*, vol. 10, pp. 51587–51633, 2022. https://doi.org/10.1109/ACCESS.2022.3174197

[30] D. Fensel *et al.*, "Introduction: What is a knowledge graph?" in *Knowledge Graphs: Methodology*, 2020, pp. 1–10. https://doi.org/10.1007/978-3-030-37439-6_1

[31] H. Wang, F. Zhang, M. Zhao, W. Li, X. Xie, and M. Guo, "Multi-task feature learning for knowledge graph enhanced recommendation," in *The World Wide Web Conference*, 2019, pp. 2000–2010. https://doi.org/10.1145/3308558.3313411

[32] B. Hu, Y. Ye, Y. Zhong, J. Pan, and M. Hu, "TransMKR: Translation-based knowledge graph enhanced multi-task point-of-interest recommendation," *Neurocomputing*, vol. 474, pp. 107–114, 2022. https://doi.org/10.1016/j.neucom.2021.11.049

[33] J. Sun and M. D. M. Billa Shagar, "MUKG: Unifying multi-task and knowledge graph method for recommender system," in *Proceedings of the 2020 2nd International Conference on Image Processing and Machine Vision*, 2020, pp. 14–21. https://doi.org/10.1145/3421558.3421561

[34] S. Yang and X. Cai, "Bilateral knowledge graph enhanced online course recommendation," *Information Systems*, vol. 107, p. 102000, 2022. https://doi.org/10.1016/j.is.2022.102000

[35] T. Qian, B. Liu, Q. V. H. Nguyen, and H. Yin, "Spatiotemporal representation learning for translation-based POI recommendation," *ACM Trans. Inf. Syst.*, vol. 37, no. 2, pp. 1–24, 2019. https://doi.org/10.1145/3295499

[36] Y. Chen, S. Mensah, F. Ma, H. Wang, and Z. Jiang, "Collaborative filtering grounded on knowledge graphs," *Pattern Recognition Letters*, vol. 151, pp. 55–61, 2021. https://doi.org/10.1016/j.patrec.2021.07.022

[37] Q. Dai *et al.*, "Personalized knowledge-aware recommendation with collaborative and attentive graph convolutional networks," *Pattern Recognition*, vol. 128, p. 108628, 2022. https://doi.org/10.1016/j.patcog.2022.108628

[38] X. Sha, Z. Sun, and J. Zhang, "Hierarchical attentive knowledge graph embedding for personalized recommendation," *Electronic Commerce Research and Applications*, vol. 48, p. 101071, 2021. https://doi.org/10.1016/j.elerap.2021.101071

[39] E. Palumbo, D. Monti, G. Rizzo, R. Troncy, and E. Baralis, "entity2rec: Property-specific knowledge graph embeddings for item recommendation," *Expert Systems with Applications*, vol. 151, p. 113235, 2020. https://doi.org/10.1016/j.eswa.2020.113235

[40] L. Sang, M. Xu, S. Qian, and X. Wu, "Knowledge graph enhanced neural collaborative recommendation," *Expert Systems with Applications*, vol. 164, p. 113992, 2021. https://doi.org/10.1016/j.eswa.2020.113992

[41] L. Sang, M. Xu, S. Qian, and X. Wu, "Knowledge graph enhanced neural collaborative filtering with residual recurrent network," *Neurocomputing*, vol. 454, pp. 417–429, 2021. https://doi.org/10.1016/j.neucom.2021.03.053

[42] S. Li *et al.*, "Self-supervised learning for conversational recommendation," *Information Processing & Management*, vol. 59, no. 6, p. 103067, 2022. https://doi.org/10.1016/j.ipm.2022.103067

[43] X. Guo *et al.*, "DKEN: Deep knowledge-enhanced network for recommender systems," *Information Sciences*, vol. 540, pp. 263–277, 2020. https://doi.org/10.1016/j.ins.2020.06.041

[44] S. Chantamunee, K. W. Wong, and C. C. Fung, "Relation-aware collaborative autoencoder for personalized multiple facet selection," *Knowledge-Based Systems*, vol. 246, p. 108683, 2022. https://doi.org/10.1016/j.knosys.2022.108683

[45] Y. Liu, F. Gu, X. Gu, Y. Wu, J. Guo, and J. Zhang, "Resource recommendation based on industrial knowledge graph in low-resource conditions," *International Journal of Computational Intelligence Systems*, vol. 15, 2022. https://doi.org/10.1007/s44196-022-00097-2

[46] X. Liu, R. Song, Y. Wang, and H. Xu, "A multi-granular aggregation-enhanced knowledge graph representation for recommendation," *Informastion*, vol. 13, no. 5, p. 229, 2022. https://doi.org/10.3390/info13050229

[47] R. Ma *et al.*, "Knowledge graph extrapolation network with transductive learning for recommendation," *Applied Sciences*, vol. 12, no. 10, p. 4899, 2022. https://doi.org/10.3390/app12104899

[48] H. Zhang, Y. Wang, C. Chen, R. Liu, S. Zhou, and T. Gao, "Enhancing knowledge of propagation-perception-based attention recommender systems," *Electronics*, vol. 11, no. 4, p. 547, 2022. https://doi.org/10.3390/electronics11040547

[49] X. Zhao, C. Liu, S. Zhang, and X. You, "A novel science and technology resource recommendation service based on knowledge graph and collaborative filtering," in 2022 *International Conference on Service Science (ICSS)*, 2022, pp. 196–202. https://doi.org/10.1109/ICSS55994.2022.00037

[50] T.-Y. Qian, B. Liu, L. Hong, and Z.-N. You, "Time and location aware points of interest recommendation in location-based social networks," *Journal of Computer Science and Technology*, vol. 33, pp. 1219–1230, 2018. https://doi.org/10.1007/s11390-018-1883-7

[51] A. Li and B. Yang, "GSIRec: Learning with graph side information for recommendation," *World Wide Web*, vol. 24, pp. 1411–1437, 2021. https://doi.org/10.1007/s11280-021-00910-6

[52] J.-L. Li, Z.-J. Du, and J.-T. Zhou, "A hybrid recommendation algorithm with short-term preference and knowledge preference," in *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications*, 2020, pp. 593–600. https://doi.org/10.1109/ISPA-BDCloud-SocialCom-SustainCom51426.2020.00100

[53] X. Ma, L. Dong, Y. Wang, Y. Li, and H. Zhang, "MNI: An enhanced multi-task neighborhood interaction model for recommendation on knowledge graph," *PLoS ONE*, vol. 16, no. 10, p. e0258410, 2021. https://doi.org/10.1371/journal.pone.0258410

[54] R. Ma, F. Guo, Z. Li, and L. Zhao, "Knowledge graph random neural networks for recommender systems," *Expert Syst. Appl.*, vol. 201, p. 117120, 2022. https://doi.org/10.1016/j.eswa.2022.117120

[55] W. Zhao, Y. Li, T. Fan, and F. Wu, "Retracted article: A novel embedding learning framework for relation completion and recommendation based on graph neural network and multi-task learning," *Soft Computing*, vol. 28, p. 447, 2022. https://doi.org/10.1007/s00500-021-06617-0

[56] C. Yan, S. Liu, Y. Zhang, Z. Wang, and P. Wang, "A multi-task learning approach for recommendation based on knowledge graph," in *2021 International Joint Conference on Neural Networks (IJCNN)*, 2021, pp. 1–8. https://doi.org/10.1109/IJCNN52387.2021.9533556

[57] Z. Deng, C. Li, S. Liu, W. Ali, and J. Shao, "Knowledge-aware group representation learning for group recommendation," in *2021 IEEE 37th International Conference on Data Engineering (ICDE)*, 2021, pp. 1571–1582. https://doi.org/10.1109/ICDE51399.2021.00139

[58] H. Shu and J. Huang, "User-preference based knowledge graph feature and structure learning for recommendation," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*, 2021, pp. 1–6. https://doi.org/10.1109/ICME51207.2021.9428363

[59] L. Yang, E. Shijia, S. Xu, and Y. Xiang, "Interactive knowledge graph attention network for recommender systems," in *2020 International Conference on Data Mining Workshops (ICDMW)*, 2020, pp. 211–219. https://doi.org/10.1109/ICDMW51313.2020.00038

[60] Q. Guo, Z. Sun, J. Zhang, and Y.-L. Theng, "An attentional recurrent neural network for personalized next location recommendation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, vol. 34, no. 1, pp. 83–90. https://doi.org/10.1609/aaai.v34i01.5337

[61] X. Wang, F. D. Salim, Y. Ren, and P. Koniusz, "Relation embedding for personalised translation-based POI recommendation," in *Lecture Notes in Computer Science*, vol. 12084, 2020, pp. 53–64. https://doi.org/10.1007/978-3-030-47426-3_5

[62] M. Cai and J. Zhu, "Knowledge-aware graph collaborative filtering for recommender systems," in *2019 15th International Conference on Mobile Ad-Hoc and Sensor Networks (MSN)*, 2019, pp. 7–12. https://doi.org/10.1109/MSN48538.2019.00016

[63] X. Dong, T. Li, and Z. Ding, "An ontology enhanced user profiling algorithm based on application feedback," in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, 2019, pp. 316–325. https://doi.org/10.1109/COMPSAC.2019.00054

[64] Yuan, Y. Tang, L. Du, and X. Li, "Entity2item: Leveraging knowledge graph embedding for item recommendation," in *2021 International Joint Conference on Neural Networks (IJCNN)*, 2021, pp. 1–7. https://doi.org/10.1109/IJCNN52387.2021.9534384

[65] J. Ren, X. Wang, Q. He, B. Yi, and Y. Zhang, "A cultural resource recommendation model based on graph neural networks," in *2022 14th International Conference on Communication Software and Networks (ICCSN)*, 2022, pp. 119–125. https://doi.org/10.1109/ICCSN55126.2022.9817589

[66] Z. Hou, T. Li, H. Fu, Q. Liu, Z. Zhang, and M. Hu, "A model hybrid recommendation approach based on knowledge graph convolution networks," in *2021 4th International Conference on Artificial Intelligence and Big Data (ICAIBD)*, 2021, pp. 283–288. https://doi.org/10.1109/ICAIBD51990.2021.9459108

[67] M. Gao, J.-Y. Li, C.-H. Chen, Y. Li, J. Zhang, and Z.-H. Zhan, "Enhanced multi-task learning and knowledge graph-based recommender system," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 10, pp. 10281–10294, 2023. https://doi.org/10.1109/TKDE.2023.3251897

[68] H. Shu and J. Huang, "Multi-task feature and structure learning for user-preference based knowledge-aware recommendation," *Neurocomputing*, vol. 532, pp. 43–55, 2023. https://doi.org/10.1016/j.neucom.2023.02.023

[69] Z. He, B. Hui, S. Zhang, C. Xiao, T. Zhong, and F. Zhou, "Exploring indirect entity relations for knowledge graph enhanced recommender system," *Expert Systems with Applications*, vol. 213, p. 118984, 2023. https://doi.org/10.1016/j.eswa.2022.118984

[70] Y. Li, L. Hou, and J. Li, "Preference-aware graph attention networks for cross-domain recommendations with collaborative knowledge graph," *ACM Trans. Inf. Syst.*, vol. 41, no. 3, pp. 1–26, 2023. https://doi.org/10.1145/3576921

[71] Y. Li, L. Hou, D. Li, and J. Li, "HKGCL: Hierarchical graph contrastive learning for multi-domain recommendation over knowledge graph," *Expert Systems with Applications*, vol. 233, p. 120963, 2023. https://doi.org/10.1016/j.eswa.2023.120963

[72] J. Gao, P. Peng, F. Lu, C. Claramunt, and Y. Xu, "Towards travel recommendation interpretability: Disentangling tourist decision-making process via knowledge graph," *Information Processing & Management*, vol. 60, no. 4, p. 103369, 2023. https://doi.org/10.1016/j.ipm.2023.103369

[73] X. Song, J. Qin, and Q. Ren, "A recommendation algorithm combining local and global interest features," *Electronics (Switzerland)*, vol. 12, no. 8, p. 1857, 2023. https://doi.org/10.3390/electronics12081857

[74] Y. Cui, H. Sun, Y. Zhao, H. Yin, and K. Zheng, "Sequential-knowledge-aware next POI recommendation: A meta-learning approach," *ACM Trans. Inf. Syst.*, vol. 40, no. 2, pp. 1–22, 2021. https://doi.org/10.1145/3460198

[75] W. Chen *et al.*, "Building and exploiting spatial–temporal knowledge graph for next POI recommendation," *Knowledge-Based Systems*, vol. 258, p. 109951, 2022. https://doi.org/10.1016/j.knosys.2022.109951

[76] L. Long, Y. Yin, and F. Huang, "Graph-aware collaborative filtering for top-N recommendation," in *2021 International Joint Conference on Neural Networks (IJCNN)*, 2021, pp. 1–8. https://doi.org/10.1109/IJCNN52387.2021.9534309

[77] G. Ngo and N. N. Y. Vo, "Enhancing recommendation systems with hybrid manifold regularized knowledge graph," in *2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA)*, 2023, pp. 1–8. https://doi.org/10.1109/DSAA60987.2023.10302462

[78] X. Shen and Y. Zhang, "A knowledge graph recommendation approach incorporating contrastive and relationship learning," *IEEE Access*, vol. 11, pp. 99628–99637, 2023. https://doi.org/10.1109/ACCESS.2023.3310816

[79] F. Chen, Z. Kang, C. Zhang, and C. Wu, "Multi-contrastive learning recommendation combined with knowledge graph," in *2023 International Joint Conference on Neural Networks (IJCNN)*, 2023 pp. 1–8. https://doi.org/10.1109/IJCNN54540.2023.10191678

[80] Q. Tian, "Multi-channel contrastive learning for sequential recommendation," in *2022 8th International Conference on Systems and Informatics (ICSAI)*, 2022, pp. 1–6. https://doi.org/10.1109/ICSAI57119.2022.10005401

[81] Y. Yang, C. Huang, L. Xia, and C. Li, "Knowledge graph contrastive learning for recommendation," in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2022, pp. 1434–1443. https://doi.org/10.1145/3477495.3532009

[82] P. C. Mahalanobis, "On the generalized distance in statistics," *Sankhyā: The Indian Journal of Statistics*, vol. 80, pp. S1–S7, 2018.

[83] H. Wang *et al.*, "RippleNet: Propagating user preferences on the knowledge graph for recommender systems," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, 2018, pp. 417–426. https://doi.org/10.1145/3269206.3271739

[84] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, "Translating embeddings for modeling multi-relational data," in *Advances in Neural Information Processing Systems*, 2013, vol. 2, pp. 2787–2795. https://dl.acm.org/doi/10.5555/2999792.2999923

[85] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021. https://doi.org/10.1016/j.neucom.2021.03.091

[86] J. Peng, X. Hu, W. Huang, and J. Yang, "What is a multi-modal knowledge graph: A survey," *Big Data Research*, vol. 32, p. 100380, 2023. https://doi.org/10.1016/j.bdr.2023.100380

[87] O. Reda, I. Sassi, A. Zellou, and S. Anter, "Towards a data quality assessment in big data," in *Proceedings of the 13th International Conference on Intelligent Systems: Theories and Applications*, 2020, pp. 91–96. https://doi.org/10.1145/3419604.3419803

[88] X. Wang *et al.*, "Knowledge graph quality control: A survey," *Fundamental Research*, vol. 1, no. 5, pp. 607–626, 2021. https://doi.org/10.1016/j.fmre.2021.09.003

8 AUTHORS

Bouchra Boualaoui is with the SPM Research Team, ENSIAS, Mohammed V University in Rabat, Rabat, Morocco (E-mail: bouchra_boualaoui@um5.ac.ma).

Ahmed Zellou is with the SPM Research Team, ENSIAS, Mohammed V University in Rabat, Rabat, Morocco.

Mouna Berquedich is with the Green Tech Institute, Mohammed VI Polytechnique University, Benguerir, Morocco.