TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 21/24 pp1580–1592
DOI: 10.26599/TST.2023.9010118
Volume 29, Number 5, October 2024

Exploring Artistic Embeddings in Service Design: A Keyword-Driven Approach for Artwork Search and Recommendations

Jie Yuan, Fangru Lin, and Hae Yoon Kim*

Abstract: As living standards improve, the demand for artworks has been escalating, transcending beyond the realm of mere basic human necessities. However, amidst an extensive array of artwork choices, users often struggle to swiftly and accurately identify their preferred piece. In such scenarios, a recommendation system can be invaluable, assisting users in promptly pinpointing the desired artworks for better service design. Despite the escalating demand for artwork recommendation systems, current research fails to adequately meet these needs. Predominantly, existing artwork recommendation methodologies tend to disregard users' implicit interests, thereby overestimating their capability to articulate their preferences in full and often neglecting the nuances of their diverse interests. In response to these challenges, we have developed a weighted artwork correlation graph and put forth an embedding-based keyword-driven artwork search and recommendation methodology. Our approach transforms the keywords that delineate user interests into word embedding vectors. This allows for an effective distinction between the user's core and peripheral interests. Subsequently, we employ a dynamic programming algorithm to extract artworks from the correlation graph, thereby obtaining artworks that align with the user's explicit keywords and implicit interests. We have conducted an array of experiments using real-world datasets to validate our approach. The results attest to the superiority of our method in terms of its efficacy in searching and recommending artworks.

Key words: artwork; self-attention; recommendation; dynamic programming; implicit interest

1 Introduction

In the wake of continuous economic advancement in human society, living conditions have seen substantial improvements. As a result, the fundamental focus of survival has become less predominant, giving way to a broader pursuit of quality of life as a key aspiration for humankind[1]. This societal and economic evolution, in tandem with the relentless progression of industrial capabilities, has facilitated wider accessibility to a diverse array of artworks, enriching the daily lives of many individuals[2]. The realm of art, both traditional and emergent, has flourished under these conditions. Traditional forms of artistry such as music, painting, and architecture continue to inspire and provide aesthetic pleasure. Concurrently, emergent artistic expressions, including film, broadcasting, and photography, have witnessed burgeoning popularity. The ubiquitous presence and growing appreciation of art have transcended societal boundaries, heralding an era where people are paying unprecedented attention to artistic endeavors. This heightened awareness and interest have propelled the art industry into a rapid

• Jie Yuan, Fangru Lin, and Hae Yoon Kim are with College of Design, Dongseo University, Busan 47011, Republic of Korea. E-mail: yjrita0809@gmail.com; lfr19910929@gmail.com; haeyoonkim@empal.com.
• Jie Yuan is also with Shandong Provincial University Laboratory for Protected Horticulture, College of Arts, Weifang University of Science and Technology, Weifang 262700, China.
* To whom correspondence should be addressed.
Manuscript received: 2023-08-18; revised: 2023-09-30; accepted: 2023-10-13

cycle of growth, subsequently bolstering the turnover of the industry and creating expansive opportunities for growth in an array of art-related domains.

The escalating enthusiasm for artworks, while beneficial in numerous ways, also presents a series of unique challenges. The foremost of these is the rapid proliferation in the number and diversity of various artistic outputs including artifacts, music, literature, and film works[3]. The sheer volume, varied quality, and wide range of genres create a daunting landscape for users attempting to locate suitable choices in this vast artistic ecosystem for better service design. The complexity of selection is further compounded by the nuanced individual preferences and the evolving taste of users, making it a formidable task to pinpoint the right artwork that resonates with them. In light of this conundrum, numerous scholars have advanced various methodologies aiming at recommending items that align with users' needs. Broadly, these methodologies fall into three primary categories: (1) leveraging the similar relationships[4] of users or items, including techniques such as collaborative filtering[5] and topic modeling[6, 7]; (2) harnessing graph methods[8], such as graph neural networks; and (3) utilizing other information[9, 10], such as social information[11], tags, categories, etc. These methodologies, each with their respective merits, have proven effective in recommending items of interest to users to varying degrees. However, while these methods have certainly made strides in the realm of artwork recommendation, they are not without their shortcomings. Each approach, though effective in its own right, has specific limitations that prevent it from fully addressing the complexities of user preferences in the realm of art. These limitations often manifest as an inability to accurately interpret and respond to users' implicit interests, thereby leading to an overestimation of users' abilities to fully articulate their preferences[12]. Additionally, these methods often overlook the diverse and evolving focus of users' interests, resulting in recommendations that may not entirely satisfy users' multifaceted desires and expectations.

A critical observation in the realm of artwork recommendation is that user interests are typically unequal. Users often prioritize some interests over others. For instance, in a scenario where a user expresses interest in "science fiction", "action", and "drama" movies, a recommendation of a standalone "science fiction" movie may garner more engagement from the user than a singular "drama" movie. Through the interaction of these three keywords, we can infer a slightly diminished importance attributed to "drama". This example underscores the necessity to discern between the relative weightings of users' varied interests. Furthermore, existing recommendation methodologies place a considerable degree of trust in the users' ability to articulate their interests accurately. Techniques such as collaborative filtering, for instance, use the similarity between the keywords describing the user's interests and the keywords describing the artwork to generate recommendations. However, empirical research[13] highlighted that users often grapple with accurately and comprehensively articulating all of their needs. Consequently, the keywords provided by the user may not encompass the full spectrum of artworks potentially of interest to them. This gap can result in certain artworks, which would have otherwise piqued the user's interest, being overlooked in the recommendation process.

In response to the challenges identified in the artwork recommendation process, we propose an innovative embedding-based keyword-driven method for artwork search and recommendation. Our methodology's contribution is threefold, providing significant advancements in the following areas:

(1) We have constructed an artwork correlation graph to illustrate the intricate relationships between different artworks, thereby capturing the complex interplay of themes, styles, and genres inherent in the artistic landscape.

(2) We deploy a self-attention model, an innovative approach that enables the differentiation between users' core and non-core interests. This model is further enriched by a dynamic planning algorithm, which connects artworks that are strongly related to the user's core interests, thereby enhancing the relevance and appeal of the recommended artworks.

(3) To validate our methodology, we conducted a comprehensive series of experiments on a real-world dataset. The results from these experiments showcase our method's superiority over existing techniques in effectively searching and recommending artworks, further reinforcing the effectiveness of our approach.

Following this introduction, we proceed to explore pertinent research related to artwork recommendations and privacy data protection in Section 2. In Section 3, we present a detailed case study, graphically illustrating the motivation and relevance of our research.

Subsequently, in Section 4, we delve into the specifics of our proposed method for searching and recommending artworks. Section 5 outlines a series of experiments designed to verify the effectiveness and advantages of our novel approach. Finally, in Section 6, we conclude with a summary of our proposed methodology, its implications, and a glimpse into our prospective future work in the field.

2 Related Work

2.1 Recommendation system

A recommendation system is an information filtering system that utilizes algorithms and techniques to provide personalized suggestions and recommendations[14]. Its purpose is to assist users in discovering content that they may find interesting, thereby alleviating the problem of information overload. The core task of a recommendation system is to predict users' latent interests based on their historical behavior, preference, and item attributes, using data mining and machine learning techniques, and to provide personalized recommendations accordingly[15]. Understanding the working principles and algorithmic principles of recommendation systems is of crucial importance for enhancing and optimizing the performance of such systems.

The existing research on recommendation systems can be categorized into the following types: (1) collaborative filtering based recommendation systems[16, 17], which can be further divided into content-based collaborative filtering methods and user-based collaborative filtering methods[18]; (2) graph network based recommendation systems; (3) aggregated recommendation systems[19].

In order to solve the cold-start problem caused by sparse data, Wang et al.[20] proposed a new hybrid collaborative filtering recommendation method. They first proposed a trust-based collaborative filtering approach for rating prediction. Then they proposed a hybrid collaborative filtering recommendation method with user item trust records, which effectively improves the prediction effect of the model in the sparse data situation. On the other hand, Qi et al.[21] worked on graph modeling. They first constructed a Web Application Programming Interface (API) correlation graph, on the basis of which they proposed a Steiner tree algorithm for pre-chosen APIs. They modeled the compatibility relationship between Web APIs in the form of a graph, which in turn recommends a compatible Web API combination to software developers.

Recommendation systems face various challenges while also presenting vast prospects for future development. These challenges encompass issues such as data sparsity[22], anomaly detection[23], cold-start problem[24], long-tail recommendations[25], privacy protection[26, 27], etc. To overcome these challenges, future directions include research on the application of deep learning techniques, cross-domain and cross-media recommendations, and personalized explanation and interpret ability, aiming to enhance the performance and user experience of recommendation systems.

2.2 Artworks analysis and recommendation

Art product recommendations face specific challenges that differ from recommendation systems in other domains[7, 28]. Firstly, the subjective and emotional nature of artworks leads to more personalized and subjective user interests, making it difficult to accurately capture and predict[20]. Secondly, the diversity and complexity of artworks increase the difficulty of understanding and expressing artistic features in recommendation systems[29, 30]. Furthermore, the relative scarcity and high price nature of artworks present challenges, requiring considerations of users' budgets and purchasing power.

Darda and Chatterjee[31] argued that art is embedded in the historical, social, political, and cultural context in which it is situated and rarely assessed in isolation. Therefore the process of assessing artwork can be influenced by the semantic context created by providing text-based information about the artwork. Based on these ideas, they explored how contextual information affects the aesthetics of artworks through a series of experiments. Messina et al.[5] applied visual neural networks to content-based artwork recommendation, and proposed a new recommendation method that combines drawing metadata with neural and artificially designed visual features. They argued that despite the growing artwork market, the study of artwork recommendations has received relatively little attention. They contributed to the field of content-based artwork recommendation for physical paintings by investigating the impact of a number of features such as artwork metadata, neural visual features, and also artificially designed visual features such as naturalness, luminance, and contrast. Tian et al.[6] first used the attention mechanism for artwork recommendation, exploring the factors that influence users to like artworks. While these studies work to explore the factors that make artwork more influential or present approaches to recommending artwork that meets a user's needs, they still leave some questions unanswered. For example, it is not the case that artworks exist independently, and the complex relationships between different artworks have not received much attention from researchers. Also, users' needs are not always equally important, and they may be more concerned with some of them, e.g., paying more attention to content relevance (e.g., style and history) than to form relevance (e.g., painting and sculpture).

The future development of art product recommendations should focus on the following aspects. Firstly, multi-source data are integrated, such as textual descriptions, images[32], audio, and video of artworks, which provides a more comprehensive representation of artistic features. Secondly, sentiment analysis and emotion-based recommendation techniques are incorporated, the incorporation process considers users' emotional needs and feedback, enabling more accurate recommendations[33]. Additionally, social media data and user social network analysis are leveraged to explore art interactions and social influence among users, enhancing the effectiveness of personalized recommendations. Lastly, emphasis is placed on interpretability research to enable recommendation systems to provide explanations for recommended results, increasing user trust and satisfaction.

3 Motivation

In the context of a vast array of distinct types of artworks, users often find it a Herculean task to swiftly and accurately locate artworks aligning with their preferences. To illustrate this predicament more vividly, we reference an example depicted in Fig. 1.

As Fig. 1 demonstrates, users typically employ several keywords to delineate their needs. However, conventional search engines often return an overwhelmingly large number of results. This inundation of options necessitates users to invest considerable time and effort to meticulously sift through the search results, which can be both cumbersome and daunting. Collaborative filtering, a technique that recommends artworks to users based on the preferences of other users with similar tastes, has been a popular solution. Its efficacy has been underscored in numerous recommendation methodologies.

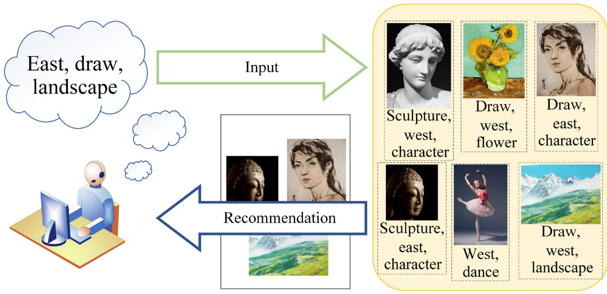

Fig. 1 Example for motivation.

Nevertheless, the recommendation process is not devoid of challenges. Two significant hurdles are particularly noteworthy:

(1) **The heterogeneity of users' interests:** Users typically prioritize certain interests over others. This hierarchical nature of interests is often overlooked by standard recommendation methodologies.

(2) **Implicit interests:** Existing artwork recommendation methodologies frequently underestimate the existence of users' implicit interests. This shortcoming results in an over-reliance on users' ability to fully articulate their interests, thereby potentially overlooking artworks that might align with users' unstated preferences.

Given these challenges, we propose a novel embedded keyword-driven methodology for artwork search and recommendation. Our approach is designed to address these shortcomings by acknowledging and accommodating the hierarchical nature of users' interests and incorporating a mechanism to identify and cater to implicit interests. This approach is expected to significantly improve the accuracy and relevance of artwork recommendations, thereby enhancing the user experience in navigating the vast landscape of artworks. The specifics of our approach will be discussed in detail in the following section.

4 Our Recommendation Method

In this section, we describe in detail the embedding-based keyword-driven artwork search and recommendation method. Before giving the method, we introduce the constructed artwork correlation graph.

4.1 Artwork correlation graph

The graph model is an effective method to reveal the relationship between entities which has been validated by many studies[21]. In this subsection, we define a graph to model the correlation between artworks, which are defined as follows:

Definition 1 Artwork correlation graph: The artwork correlation graph can be represented by $G(V, E)$, where G denotes that the graph itself consists of V (the set of nodes) and E (the set of edges).

Definition 2 Node: Each retrievable artwork can be represented by v, and they form the set of nodes V. It needs to be clear that each node v_i in V is described by a keyword set containing at least one keyword, denoted by conforming $\{k_j|k_j \in K\}_{v_i}$.

Definition 3 Node weight: For each node v_i in V, there is a dynamic weight $w(v_i)$. The weight $w(v_i)$ of the node is not fixed rather determined by the query keyword Q entered by the user and is calculated as follows:

$$w(v_i) = \sum_{j=1}^{n} w(k_j),\ n = |\{k_j|k_j \in K\}_{v_i}| \tag{1}$$

where $w(k_j)$ denotes the weight of a keyword describing v_i, and it is calculated in a way that we will provide in Section 4.2. n indicates the number of keywords describing v_i.

Definition 4 Edge: Each edge $e(v_i, v_j)$ in graph G indicates that the two nodes (artworks) v_i and v_j connected had been exhibited together, and they form the set E of edges.

Definition 5 Edge weight: Each edge $e(v_i, v_j)$ in E has a weight indicating the similarity between nodes v_i and v_j, which is calculated using the Jaccard similarity coefficient expressed by the following equation:

$$e_w(v_i, v_j) = \frac{\{k_i|k_i \in K\}_{v_i} \cap \{k_j|k_j \in K\}_{v_j}}{\{k_i|k_i \in K\}_{v_i} \cup \{k_j|k_j \in K\}_{v_j}} \tag{2}$$

where K denotes the set of all keywords in G.

As in Definition 2, each node (artwork) has a set of keywords describing it, so we can construct an offline index that allows to find the corresponding set of nodes by entering any of the keywords in K. Figure 2 is an example of an already constructed artwork correlation graph, which includes 12 nodes, each containing several descriptive keywords.

4.2 Calculating node weights

The self-attention model is a type of neural network architecture[34] used primarily in natural language processing (NLP) tasks, such as language translation

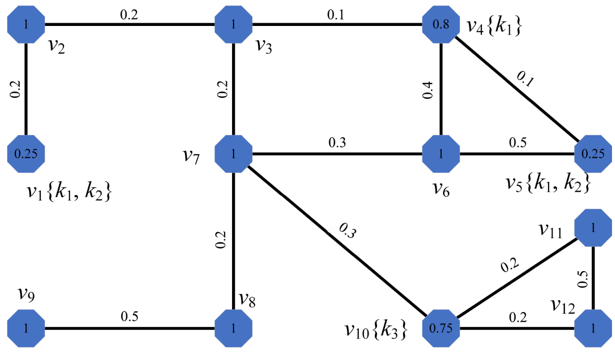

Fig. 2 Artwork correlation graph.

and sentiment analysis[35, 36]. It enables the network to selectively focus on different parts of the input sequence by computing the importance of each input element with respect to all the other elements. This is achieved by computing three matrices: the query matrix, the key matrix, and the value matrix. The dot product of the query and key matrices produces an attention matrix, which is then multiplied by the value matrix to obtain the output. The self-attention mechanism allows the model to learn long-range dependencies and capture contextual information, making it a powerful tool for NLP tasks. Here, we employ a self-attention model to compute the responsiveness of each node to users' requirements. The self-attention mechanism in its most basic form involves the following equations.

Given an input sequence of length N, each element of the sequence is represented by a d-dimensional vector X:

$$\text{Query matrix } Q = W_Q X \tag{3}$$

$$\text{Key matrix } K = W_K X \tag{4}$$

$$\text{Value matrix } V = W_V X \tag{5}$$

where X is the input sequence, W_Q is a learnable weight matrix of dimensions, W_K is a learnable weight matrix of dimensions, and W_V is a learnable weight matrix of dimensions. Compute the attention scores A as the dot product of Q and K transposed, scaled by the square root of the dimensional d, and then passed through a softmax function to obtain the normalized attention weights. The final output O is computed as the weighted sum of the values V, where the weights are given by the attention scores $A: O = AV$, as shown in Eq. (6):

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^{\text{T}}}{\sqrt{d}}\right)V \tag{6}$$

We take the co-occurrence frequencies of descriptive words used by users to describe artworks as the basis for training, which is easily understood: If a keyword co-occurs frequently with other keywords, it indicates that this keyword is likely a core requirement of the users. Therefore, the pre-defined node weights are calculated as follows:

$$w\,(v,\,q) = \sum w_{\text{attention}}(k_i) \tag{7}$$

where q represents the intersection of keywords contained in node v and keywords that describe user requirements Q_u. k is an element in q, and $w_{\text{attention}}(k_i)$ is calculated by Eqs. (8) and (9):

$$w_{\text{attention}}(k_i) = \frac{1}{1+\mathrm{e}^{-(x-c)}} \tag{8}$$

$$x = \frac{\sum\limits_{j=0}^{n}\text{Count}\,(q_i,\,q_j)}{\sum\limits_{i=0}^{n}\sum\limits_{j=0}^{n}\text{Count}\,(q_i,\,q_j)} \tag{9}$$

where c is a constant. It should be noted that we aim for greater diversity among nodes, so that smaller edge weights are preferred. For ease of computation, we transform $w\,(v,\,q)$ using Eq. (10), which means that smaller $w\,(v,\,q)$ is better in the retrieval process, and further and smaller node weights are better.

$$w\,(v,\,q) = 1-\sum w_{\text{attention}}(k_i) \tag{10}$$

4.3 Proposed method: KE-ArtR

In the previous two subsections, we discussed how to construct the artwork correlation graph, which forms the foundation of the subsequent work. In this subsection, we will describe in detail the process of recommending a diverse set of artworks based on user requirements.

As shown in Fig. 2, when a user provides a group of keywords Q that describe their interests, a set of nodes V is retrieved and contains at least one keyword from Q. If all nodes in V are recommended to the user, redundancy occurs, as in the case of v_1 and v_5. However, recommending a single node from V cannot fully satisfy the user's requirements. Therefore, we need to combine the nodes in V to meet the user's needs without redundancy, which is an NP-hard problem[37]. At the same time, in the combination process, other nodes need to serve as bridge nodes to connect the nodes in V. We use the Steiner tree algorithm for retrieval in this regard[38]. The Steiner tree is defined as follows:

Definition 6 Steiner tree: For a given connected graph $G(V, E)$ and a subset of nodes V', the Steiner tree is the tree T that connects all nodes in V'.

Definition 7 Tree weight: For a given Steiner tree T, the weight of T is the sum of the weights of all nodes and edges in T, denoted as $T(v)$.

Definition 8 Minimum weighted tree: For a given connected graph $G(V, E)$ and user demand Q, a set of Steiner trees $T_1, T_2, \ldots, T_n$ can be obtained, and the Steiner tree with the minimum weight is called the minimum weight Steiner tree, denoted as $T_{\min}(v)$.

As shown in Fig. 3, when a user provides a set of keywords Q describing his/her interests, we first compute the self-attention weights of these keywords to discover the user's core and non-core interests. After that, with the obtained self-attention weights, we compute the weights of each node in the artwork correlation graph in order to get the relatively important artworks (nodes with lower weights). Then we find all the nodes $\{v_1, v_4, v_5, v_{10}\}$ with weights less than 1 and store these initial nodes in the temporary queue. Next, we will repeat the tree growing and merging operations until the temporary queue is empty. The grow operation means that the tree connects one of the neighbors of the root node and uses the newly connected neighbor node as the root node of the new tree. The weights of the new tree are calculated as follows:

$$w\,(\text{nT}) = w\,(T)+e_w\,(v_i, v_j)+w\,(v_j) \tag{11}$$

where nT denotes the new tree with v_j as the root node, and T denotes the original tree with v_i as the root node. A merge operation is the process of combining two trees with the same root node into a new tree. The weights of the new tree are calculated as follows:

$$w\,(\text{nT}) = w\,(T_1)+w\,(T_2)-w\,(v) \tag{12}$$

where T_1 and T_2 denote the original tree with v as the root node.

As shown in Fig. 3b, we take each initial node as a tree and connect each tree to the neighboring nodes of the root node to get a new tree with the new node as the root node. This process is the tree growth operation. After obtaining the new trees, we need to calculate the weight of each tree and the intersection Q^1 between its set of keywords and Q. We then eliminate the trees that have the same root node, the same Q^1, and a larger weight, and put the remaining trees back into the

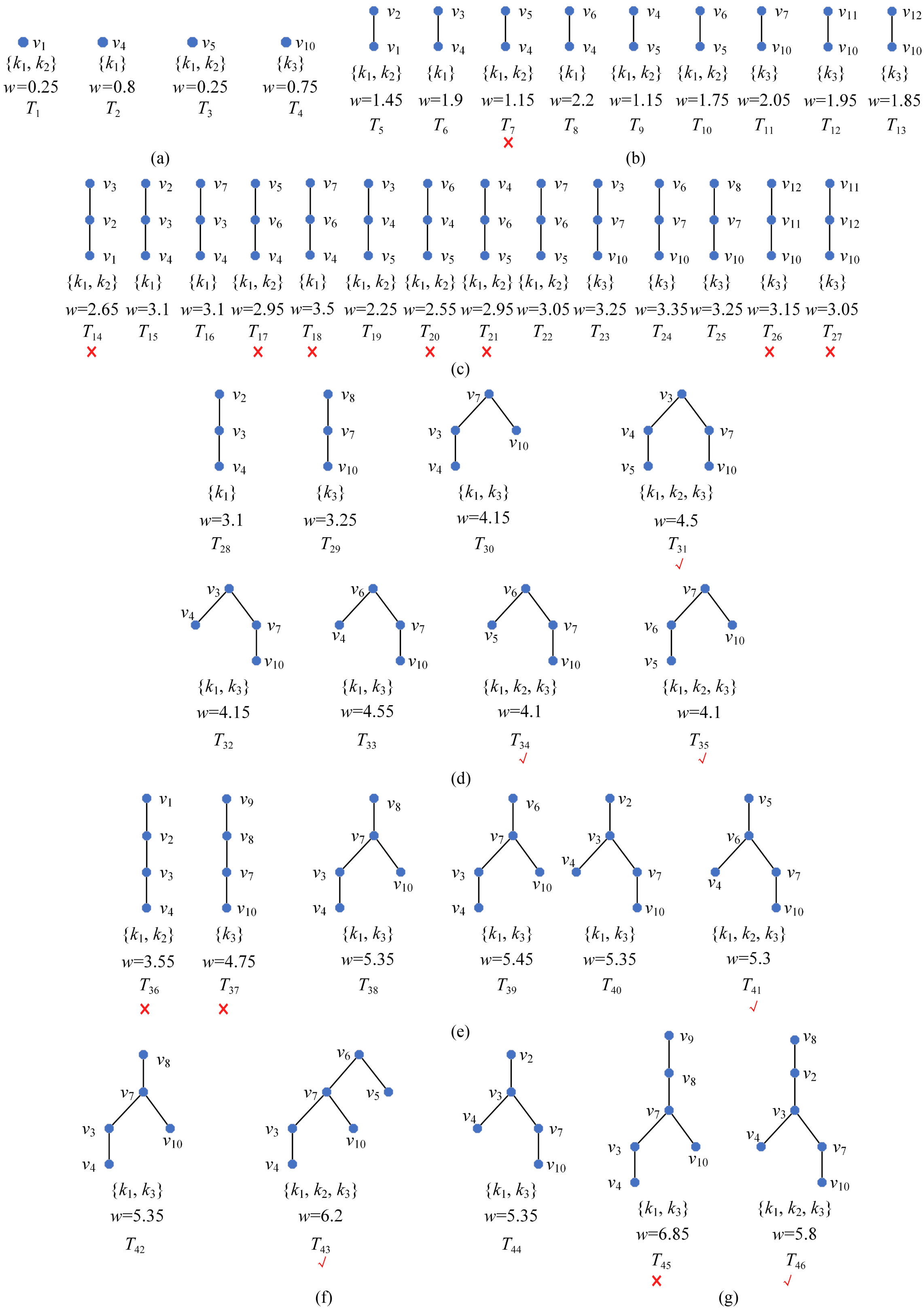

Fig. 3 Case study of Algorithm 1.

temporary queue. After completing the growth operation, trees in the temporary queue with the same root node v_1 but different Q^1 are merged into a new tree with v_1 as the root node. The weight of the new tree is recalculated, and this operation is called the merge operation. After the merge operation, the merged tree is put back to the temporary queue. After each growth and merge operation, if the Q^1 of the tree is a subset of Q, then the tree meets the user's interest and is stored in the final queue. The process is repeated until there are no more elements in the temporary queue or no more growth, and merge operations can be performed. The pseudo-code for the retrieval process is shown in Algorithm 1.

5 Experiment

5.1 Experiment setup

To validate the performance of our proposed method, we conducted a series of simulation experiments on the real-world dataset of Wu et al.[39] First, we selected two baseline methods as comparison methods: the random algorithm[40] that randomly selects nodes from the graph to form a connected tree and the greedy algorithm[40] that tends to select nodes with smaller weights during the selection process. To measure the performance of all the methods, we compared the metrics as follows:

Algorithm 1 Minimum weight tree

Input: $G(V,E)$: the artwork correlation graph; Q: query keywords; $w_{\text{attention}}(k_i)$: attention weight of k_i ($k_i \in Q$); and $N(v)$: v's neighbors.
Output: $T_{\min}(v, Q)$
initialization: Let $\text{Queue}_{\text{temp}} = \varnothing$ and $\text{Queue}_{\text{result}} = \varnothing$;
for $v_i \in V$ **do**
  if $K(v_i) \cap Q \neq \varnothing$ **then**
    Put $T(v, K(v_i))$ into $\text{Queue}_{\text{temp}}$;
    $w(v, K(v_i)) = 1 - \sum w_{\text{attention}}(q_i),\ q_i \in K(v_i) \cap \varnothing$;
  else
    $w(v, K(v_i)) = 1$
  end
end
while $\text{Queue}_{\text{temp}} \neq \varnothing$ **do**
  Get $T(v, K)$ from $\text{Queue}_{\text{temp}}$;
  if $K = Q$ **then**
    Put $T(v, K)$ into $\text{Queue}_{\text{result}}$
  else
    for $v_j \in N(v_i)$ **do**
      if $w(T(v_i, Q^1)) + w(e(v_i, v_j)) + w(T(v_i, Q^2)) < w(T(v_j, Q^1 \cap Q^2))$ **then**
        Growth $(T(v_i, Q^1))$ to $T(v_j, Q^1 \cap Q^2)$
      end
    end
    for $T(v, Q^1), T(v, Q^2) \in \text{Queue}_{\text{temp}}$ **do**
      if $Q^1, Q^2 \subsetneq (Q^1 \cap Q^2)$ **then**
        Merge$(T(v, Q^1), T(v, Q^2))$ to $T(v, Q^1 \cap Q^2)$
      end
    end
  end
end

(1) **Precision:** For measuring the ability of the model to recommend results that meet the input constraints, the larger the better;

(2) **Diversity:** For measuring the ability of the model to recommend more diverse results, the larger the better;

(3) **Coverage:** For measuring the range of keywords covered by the model in terms of the diversity of recommended results, which used to aid in measuring diversity of recommendation results;

(4) **Number of artwork groups:** A measure of the number of combinations obtained, the smaller the better;

(5) **Number of artworks:** For measuring diversity of recommendation results, the smaller the better;

(6) **Time cost:** A measure of the model's efficiency, the smaller the better.

We repeated each algorithm 50 times to take the average value to ensure that the test results are stable and valid. The experiments were conducted on a computer with 32 G RAM, RTX3060Ti 8 G graphics card, running Windows 10 OS, and Python version 3.7.

5.2 Experiment analysis

Profile 1: Precision

In this profile, we compare the performance of KE-ArtR with two baseline methods in precision, as shown in Fig. 4. The results show that the greedy and random algorithms have high precision with fewer input keywords, but the precision decreases to varying degrees as the number of input keywords increases. On the other hand, the precision of Keyword-Driven Approach for Artwork Search and Recommendations (KE-ArtR) is slightly lower with fewer input keywords, but as the number of input keywords increases, the precision improves significantly, and is significantly higher than the other methods. This is because with fewer input keywords, more candidate groups can be retrieved, which reduces the precision of the retrieval. When typing more input keywords, fewer candidate

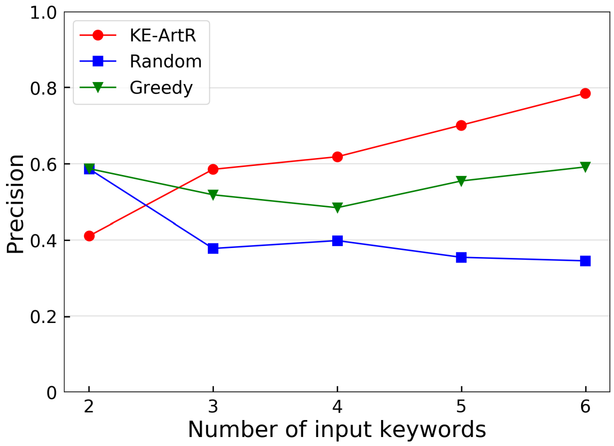

Fig. 4 Precision comparison of KE-ArtR, random, and greedy.

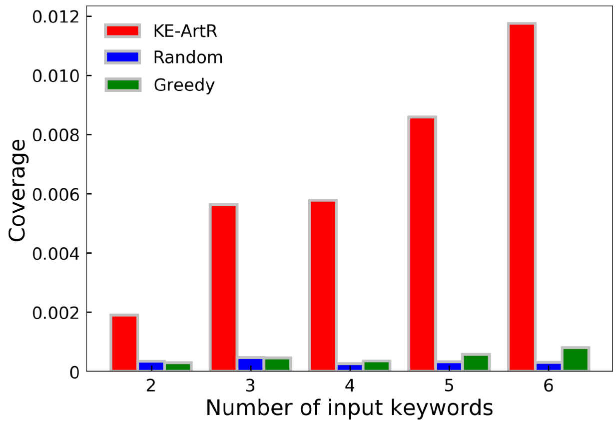

Fig. 6 Coverage comparison of KE-ArtR, random, and greedy.

groups satisfy the condition, and the precision of KE-ArtR increases. When fixing the number of input keywords, KE-ArtR is more precise than the other methods, which means that it recommends artworks that are more in line with the user's interests. Overall, KE-ArtR had better precision performance compared to the other two baseline methods.

Profile 2: Diversity

In this profile, we evaluate the diversity performance of all methods through two evaluation metrics: diversity and coverage, as shown in Figs. 5 and 6. As can be seen in Fig. 5, the diversity of all three methods shows a decreasing trend as the number of input keywords increases. This is because the candidate groups that satisfy the constraints decrease as the number of input keywords increases. When fixing the number of input keywords, KE-ArtR outperforms all other methods in terms of diversity. As shown in Fig. 6, the coverage of KE-ArtR is much higher than other methods. Specifically, the coverage of KE-ArtR increases significantly as more input keywords are typed, while the coverage of the other two methods does not change much. This is because as more input keywords are typed, the number of artworks retrieved increases. When the number of input keywords is fixed, the coverage of KE-ArtR is much higher than other methods. This is because the number of artworks in each group retrieved by the greedy and random algorithms is relatively small. Thus, the overall diversity of KE-ArtR is superior to other methods.

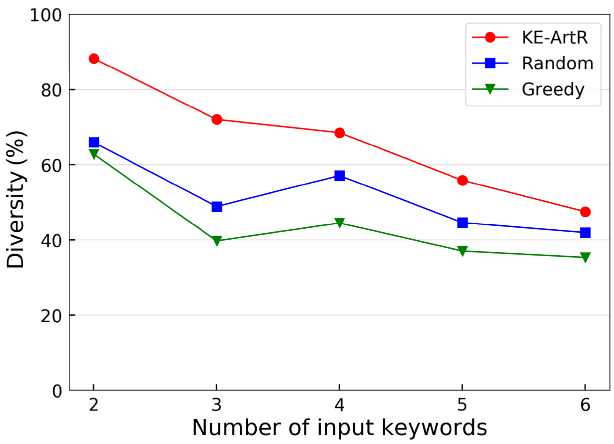

Fig. 5 Diversity comparison of KE-ArtR, random, and greedy.

Profile 3: Number of results

In this profile, we compare the number of artworks and artwork groups obtained by the three methods, as shown in Table 1. To visualize the results, we plotted Figs. 7 and 8 based on Table 1, where $F(m)=\dfrac{\log 4}{\log m}$, m = number of artwork groups, $F(n)=\dfrac{\log 4}{\log n}$, and n = number of artworks. Therefore, in Table 1, the fewer the number of artwork groups and the number of artworks, the better, as too many items require manual filtering by the user. And in Figs. 7 and 8, the larger the converted value, the better. As shown in Fig. 7, the y-value gradually increases as the number of input keywords increases. This is because as more constraints are imposed, the number of artwork groups that satisfy the constraints decreases. Regardless of the number of input keywords, the number of artwork groups obtained by KE-ArtR is far less than the other two methods, which greatly reduces the burden of manual screening on the user. Also in Fig. 8, the number of artworks increases as more input keywords are typed. This is because more artworks need to be retrieved to

Table 1 Result comparison of three methods.

Item	Number of input keywords	KE-ArtR	Greedy	Random
Precision	2	0.4106	0.5865	0.5871
	3	0.5852	0.3778	0.5183
	4	0.6184	0.3986	0.4843
	5	0.7015	0.3546	0.5546
	6	0.7853	0.3453	0.5915
Diversity (%)	2	88.2220	65.9399	62.7845
	3	72.0209	48.9024	39.7522
	4	68.5318	57.1120	44.5177
	5	55.8381	44.6152	37.0846
	6	47.5331	41.9518	35.3864
Coverage	2	0.0019	0.0003	0.0003
	3	0.0056	0.0005	0.0005
	4	0.0058	0.0003	0.0004
	5	0.0086	0.0003	0.0006
	6	0.0117	0.0003	0.0008
Number of artwork groups	2	197.78	1530.14	1454.32
	3	25.82	1267.04	672.38
	4	27.98	2316.58	969.02
	5	11.88	1877.36	548.74
	6	5.78	2040.60	317.88
Number of artworks	2	5.1382	9.0308	8.6142
	3	5.7091	86.0388	19.3308
	4	6.7253	51.0626	20.7895
	5	6.5703	111.2651	21.4298
	6	6.7809	21.4298	20.8821
Time cost (s)	2	0.0800	1.6321	0.2005
	3	0.0248	133.7888	0.9038
	4	0.0411	64.6932	1.7858
	5	0.0336	298.5450	1.5700
	6	0.0348	256.4022	1.8115

satisfy so many constraints. Also for any value of the parameter, KE-ArtR recommends fewer artworks than the other two methods. The number of results obtained by KE-ArtR is much lower than that of other methods, which greatly reduces the difficulty of manual screening by the user.

Profile 4: Time cost

In this profile, we compare the efficiency of the three methods as shown in Table 1. KE-ArtR can obtain items that satisfy the constraints in 0.1 s and recommend interested artworks to the user. Other methods take at most nearly 300 s and at least more than 0.2 s to obtain items that satisfy the constraints. To visualize the results, we plot Fig. 9 based on Table 1, where $F(x) = -\log x$ and $x = \text{time cost}$. In

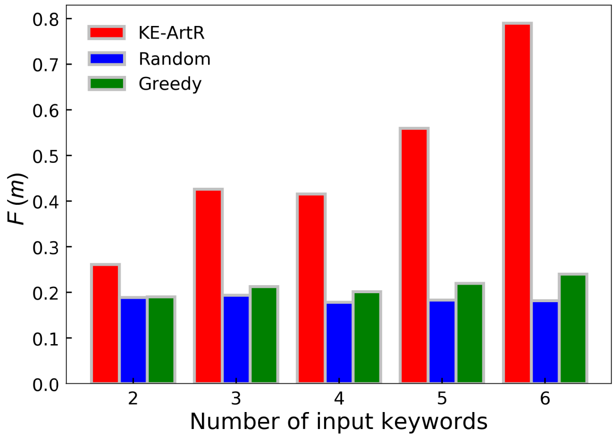

Fig. 7 ***F*(*m*) of KE-ArtR, random, and greedy.**

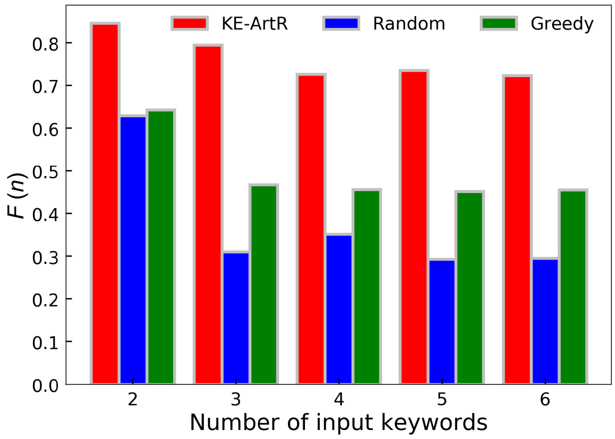

Fig. 8 ***F*(*n*) of KE-ArtR, random, and greedy.**

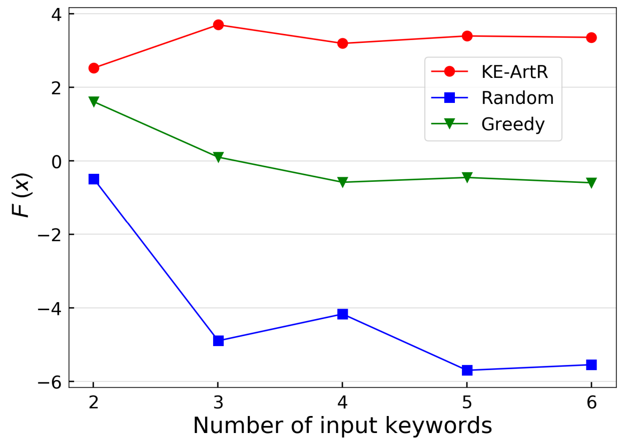

Fig. 9 ***F*(*x*) comparison of KE-ArtR, random, and greedy.**

Fig. 9, the higher y-value indicates that the model is more efficient. In particular, when the y-value is above the 0-axis, it indicates that the model took less than 1 s to obtain the result, and conversely, when the y-value is below the 0-axis, it indicates that the model took more than 1 s to obtain the result.

6 Conclusion

To address the limitations of existing artwork recommendation methods, we propose a new embedding-based keyword-driven approach in this paper. The primary challenges addressed by our approach are the unequal weightage of users' interests and the often-overlooked implicit interests. By acknowledging the hierarchical nature of users' preferences and incorporating mechanisms to identify and cater to unstated interests, our approach aims to offer more nuanced, accurate, and relevant artwork recommendations. We compare and analyze our method with two other baseline methods in terms of precise recommendations, diverse recommendations, and recommendation efficiency. Experimental results demonstrate that our approach is able to recommend multiple diverse artwork groups matched with user's interests in a much shorter period of time.

While our research has made notable strides in enhancing the artwork recommendation process, we acknowledge that there is still room for improvement and exploration. Future work may consider the integration of more complex user behavior data and the inclusion of evolving trends in the art world. Furthermore, refining the self-attention model to better capture the intricate nuances of user preferences could also enhance the system's performance. We hope our research provides a solid foundation for further exploration and development in the field of artwork recommendation, contributing to a more personalized and satisfying user experience in art exploration.

References

[1] M. Yang, C. Ge, X. Zhao, and H. Kou, FSPLO: A fast sensor placement location optimization method for cloud-aided inspection of smart buildings, *J. Cloud Comput. Heidelberg Ger*., vol. 12, no. 1, p. 31, 2023.

[2] C. Peng, C. Zhang, X. Xue, J. Gao, H. Liang, and Z. Niu, Cross-modal complementary network with hierarchical fusion for multimodal sentiment classification, *Tsinghua Science and Technology*, vol. 27, no. 4, pp. 664–679, 2022.

[3] Q. Hua, L. Chen, P. Li, S. Zhao, and Y. Li, A pixel–channel hybrid attention model for image processing, *Tsinghua Science and Technology*, vol. 27, no. 5, pp. 804–816, 2022.

[4] Y. Yan, B. Lu, and T. Xu, Exploring the relationship between host self-description and consumer purchase behavior using a self-presentation strategy, *Systems*, vol. 11, no. 8, p. 430, 2023.

[5] P. Messina, V. Dominguez, D. Parra, C. Trattner, and A. Soto, Content-based artwork recommendation: Integrating painting metadata with neural and manually-engineered visual features, *User Model. User Adapt. Interact.*, vol. 29, no. 2, pp. 251–290, 2019.

[6] Z. Tian, Y. Liu, J. Sun, X. Zhang, and Y. Jiang, Why a user prefers an artwork: A deep attention model for artwork recommendation, *J. Inf. Sci*., doi: 10.1177/01655515221116511.

[7] B. A. Yilma and L. A. Leiva, The elements of visual art recommendation: Learning latent semantic representations of paintings, in *Proc. 2023 CHI Conf. Human Factors in Computing Systems*, Hamburg, Germany, 2023, pp. 1–17.

[8] L. Deladiennee and Y. Naudet, A graph-based semantic recommender system for a reflective and personalised museum visit: Extended abstract, in *Proc. 2017 12th Int. Workshop on Semantic and Social Media Adaptation and Personalization* (*SMAP*), Bratislava, Slovakia, 2017, pp. 88–89.

[9] K. Kontiza, O. Loboda, L. Deladiennee, S. Castagnos, and Y. Naudet, A museum app to trigger users' reflection, presented at Int. Workshop on Mobile Access to Cultural Heritage (MobileCH2018), Barcelona, Spain, 2018.

[10] D. Yan, Y. Zhao, Z. Yang, Y. Jin, and Y. Zhang, FedCDR: Privacy-preserving federated cross-domain recommendation, *Digit. Commun. Netw.*, vol. 8, no. 4, pp. 552–560, 2022.

[11] M. Hong, J. J. Jung, F. Piccialli, and A. Chianese, Social recommendation service for cultural heritage, *Pers. Ubiquitous Comput.*, vol. 21, no. 2, pp. 191–201, 2017.

[12] J. Zhang, Y. Wang, Z. Yuan, and Q. Jin, Personalized real-time movie recommendation system: Practical prototype and evaluation, *Tsinghua Science and Technology*, vol. 25, no. 2, pp. 180–191, 2020.

[13] P. Bedi, S. B. Goyal, A. S. Rajawat, R. N. Shaw, and A. Ghosh, A framework for personalizing atypical web search sessions with concept-based user profiles using selective machine learning techniques, in *Advanced Computing and Intelligent Technologies*: *Proceedings of ICACIT 2021*, M. Bianchini, V. Piuri, S. Das, and R. N. Shaw, eds. Singapore: Springer, 2021, pp. 279–291.

[14] H. Ko, S. Lee, Y. Park, and A. Choi, A survey of recommendation systems: Recommendation models, techniques, and application fields, *Electronics*, vol. 11, no. 1, p. 141, 2022.

[15] S. Ahmadian, M. Ahmadian, and M. Jalili, A deep learning based trust- and tag-aware recommender system, *Neurocomputing*, vol. 488, pp. 557–571, 2022.

[16] F. Fkih, Similarity measures for collaborative filtering-based recommender systems: Review and experimental comparison, *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7645–7669, 2022.

[17] Y. Miao, X. Bai, Y. Cao, Y. Liu, F. Dai, F. Wang, L. Qi, and W. Dou, A novel short-term traffic prediction model based on SVD and ARIMA with blockchain in industrial Internet of Things, *IEEE Internet Things J.*, vol. 10, no. 24, pp. 21217–21226, 2023.

[18] H. Papadakis, A. Papagrigoriou, C. Panagiotakis, E. Kosmas, and P. Fragopoulou, Collaborative filtering recommender systems taxonomy, *Knowl. Inf. Syst.*, vol. 64, no. 1, pp. 35–74, 2022.

[19] S. Wu, F. Sun, W. Zhang, X. Xie, and B. Cui, Graph neural networks in recommender systems: A survey, *ACM Comput. Surv.*, vol. 55, no. 5, p. 97, 2022.
[20] F. Wang, H. Zhu, G. Srivastava, S. Li, M. R. Khosravi, and L. Qi, Robust collaborative filtering recommendation with user-item-trust records, *IEEE Trans. Comput. Soc. Syst.*, vol. 9, no. 4, pp. 986–996, 2022.
[21] L. Qi, W. Lin, X. Zhang, W. Dou, X. Xu, and J. Chen, A correlation graph based approach for personalized and compatible web APIs recommendation in mobile APP development, *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 6, pp. 5444–5457, 2023.
[22] S. Ahmadian, N. Joorabloo, M. Jalili, and M. Ahmadian, Alleviating data sparsity problem in time-aware recommender systems using a reliable rating profile enrichment approach, *Expert Syst. Appl.*, vol. 187, p. 115849, 2022.
[23] Y. Yang, X. Yang, M. Heidari, M. A. Khan, G. Srivastava, M. R. Khosravi, and L. Qi, ASTREAM: Data-stream-driven scalable anomaly detection with accuracy guarantee in IIoT environment, *IEEE Trans. Netw. Sci. Eng.*, vol. 10, no. 5, pp. 3007–3016, 2023.
[24] O. A. Wahab, G. Rjoub, J. Bentahar, and R. Cohen, Federated against the cold: A trust-based federated learning approach to counter the cold start problem in recommendation systems, *Inf. Sci.*, vol. 601, pp. 189–206, 2022.
[25] J. Chen, W. Wu, L. Shi, W. Zheng, and L. He, Long-tail session-based recommendation from calibration, *Appl. Intell.*, vol. 53, no. 4, pp. 4685–4702, 2023.
[26] Z. Liu, L. Yang, Z. Fan, H. Peng, and P. S. Yu, Federated social recommendation with graph neural network, *ACM Trans. Intell. Syst. Technol.*, vol. 13, no. 4, p. 55, 2022.
[27] F. Wang, G. Li, Y. Wang, W. Rafique, M. R. Khosravi, G. Liu, Y. Liu, and L. Qi, Privacy-aware traffic flow prediction based on multi-party sensor data with zero trust in smart city, *ACM Trans. Internet Technol.*, vol. 23, no. 3, pp. 1–19, 2023.
[28] A. Fosset, M. El-Mennaoui, A. Rebei, P. Calligaro, E. F. Di Maria, H. Nguyen-Ban, F. Rea, M. C. Vallade, E. Vitullo, C. Zhang, et al., Docent: A content-based recommendation system to discover contemporary art, arXiv preprint arXiv: 2207.05648, 2022.
[29] A. Lieto, G. L. Pozzato, M. Striani, S. Zoia, and R. Damiano, DEGARI 2.0: A diversity-seeking, explainable, and affective art recommender for social inclusion, *Cogn. Syst. Res.*, vol. 77, pp. 1–17, 2023.
[30] M. C. Chiu, G. J. Hwang, and L. H. Hsia, Promoting students' artwork appreciation: An experiential learning-based virtual reality approach, *Br. J. Educ. Technol.*, vol. 54, no. 2, pp. 603–621, 2023.
[31] K. M. Darda and A. Chatterjee, The impact of contextual information on aesthetic engagement of artworks, *Sci. Rep.*, vol. 13, no. 1, p. 4273, 2023.
[32] P. Messina, M. Cartagena, P. Cerda-Mardini, F. del Rio, and D. Parra, CuratorNet: Visually-aware recommendation of art images, arXiv preprint arXiv: 2009.04426, 2020.
[33] N. Gelvez, Multi-agent system used for recommendation of historical and cultural memories, *TECCIENCIA*, vol. 14, no. 26, pp. 52–62, 2019.
[34] L. Qi, J. Li, M. Elahi, K. Sood, Y. Yuan, and M. Khosravi, Guest editorial: Special issue on security and privacy for AI-powered smart IoT applications, *Digit. Commun. Netw.*, vol. 8, no. 4, pp. 411–414, 2022.
[35] X. Sun, L. Liu, A. Ayorinde, and J. Panneerselvam, ED-SWE: Event detection based on scoring and word embedding in online social networks for the internet of people, *Digit. Commun. Netw.*, vol. 7, no. 4, pp. 559–569, 2021.
[36] J. Li, X. Luo, Y. Zhang, P. Zhang, C. Yang, and F. Liu, Extracting embedded messages using adaptive steganography based on optimal syndrome-trellis decoding paths, *Digit. Commun. Netw.*, vol. 8, no. 4, pp. 455–465, 2022.
[37] L. Qi, X. Xu, X. Wu, Q. Ni, Y. Yuan, and X. Zhang, Digital-twin-enabled 6G mobile network video streaming using mobile crowdsourcing, *IEEE J. Select. Areas Commun.*, vol. 41, no. 10, pp. 3161–3174, 2023.
[38] J. Chen, S. Luo, L. Zhang, C. Zhang, and B. Cao, iPAS: A deep Monte Carlo tree search-based intelligent pilot-power allocation scheme for massive MIMO system, *Digit. Commun. Netw.*, vol. 7, no. 3, pp. 362–372, 2021.
[39] S. Wu, L. Qi, Y. Liu, Y. Yang, Y. Miao, and F. Dai, Diverse web APIs recommendation with privacy-preservation for Mashup development, in *Proc. 4th Int. Conf. Machine Learning for Cyber Security*, Guangzhou, China, 2022, pp. 472–485.
[40] L. Qi, H. Song, X. Zhang, G. Srivastava, X. Xu, and S. Yu, Compatibility-aware web API recommendation for mashup creation via textual description mining, *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 17, no. 1s, p. 20, 2021.

Hae Yoon Kim received the bachelor degree from School of the Art Institute of Chicago, Chicago, IL, USA in 1995 and the MFA degree from School of the Art Institute of Chicago, USA in 1997. He is currently a professor at Dongseo University, Busan, Republic of Korea, and has published more than 20 research papers in reputable journals. His research interests mainly focus on new media advertising design and service design.

Jie Yuan received the bachelor degree in art design from Xuzhou Industrial University, Xuzhou, China in 2012, and the master degree in design from Qingdao University, Qingdao, China in 2018. She is currently pursuing the PhD degree in design at Dongseo University, Busan, Republic of Korea. Her primary research interests include service design and digital media art design.

Fangru Lin received the bachelor degree in film studies from Beijing Film Academy, Beijing, China in 2015, the master degree in design from Dongseo University, Busan, Republic of Korea in 2020, and the PhD degree in design from Dongseo University, Busan, Republic of Korea in 2023. Her primary research interests include digital media art design, film design, and service design.