

RESEARCH ARTICLE

Group Recommendation Based on Heterogeneous Graph Algorithm for EBSNs

YUE WU¹, WEI HUANG (WAYNE)^{2,3}, LEAN YU'D⁴, AND YINGJIE TIAND⁴¹National Laboratory of Mechanical Manufacture System, School of Management, Xi'an Jiaotong University, Xi'an 710049, China²DeFin, NCAMS, Shenzhen 518055, China³College of Business, Southern University of Science and Technology, Shenzhen 518055, China⁴School of Economics and Management, University of Chinese Academy of Sciences, Beijing 101408, China

Corresponding author: Wei Huang (Wayne) (huangw7@sustech.edu.cn)

The work of Yue Wu was supported in part by the Research under Grant 71731009, Grant 72061127002, and Grant 2018WZDXM020.

The work of Wei Huang (Wayne) was supported in part by the Research under Grant 71433001 and Grant 71722014, and in part by the Shenzhen Key Research Base in Arts and Social Sciences.

ABSTRACT Emerging event-based social networks (EBSNs), such as Meetup, have grown rapidly and become popular in recent years. EBSNs differ from conventional social networks such as Facebook in that they not only involve online social interactions but also include offline, in-person interactions. Thus, EBSNs are naturally heterogeneous and possess more valuable social information. Group recommendations in EBSNs are typically only based on the interest information filled in by users, or friends' group information. Both these methods may not well reflect users' real intentions. In this study, we propose a recommender system to predict groups that may interest EBSN users, based on a novel heterogeneous augmented graph method and a random walk with restart algorithm. In this approach, online and offline social interactions are combined into a single heterogeneous augmented graph capturing all useful relationships, including user-to-group relationships, user-to-event relationships, user-to-attribute relationships, and group-to-attribute relationships, and among others. To our knowledge, this work is the first attempt to apply a random walk algorithm into group recommendation in EBSNs. Extensive experiments on Meetup datasets demonstrate that our proposed recommender system achieves better results in terms of recall, precision, F-Measure and MRR metrics in comparison with the other four commonly used algorithms, including random recommendation, interest-based recommendation, interest- and neighborhood-based recommendation, and Katz Centrality. The significant recommendation performance of our approach may further enhance user satisfaction of EBSNs. Moreover, our approach to group recommendation may also be extended to other recommendation-related applications such as event or friend recommendation.

INDEX TERMS Online group, recommendation, heterogeneous graph, random walk.

I. INTRODUCTION

Human society functions on the basis of complex interdependent relations as well as a diverse array of interactions among people and between people and other entities, such as natural and built environments, communities, products, and services. Thus, social interactions are universal and pervasive. Driven by the accelerating development of ubiquitous computing systems, human social interactions in virtual, online spaces accessed via the Internet have become commonplace. People

often form friendships and naturally join family, neighborhood, or school groups as well as groups such as political organizations or interest groups. Indeed, human society may be considered to comprise a massive social network (SN), in which every person or even every animal, object, or entity such as products, communities, or pets comprise a massive set of graph nodes, while their relationships may be represented by edges or links. The rich, complex structure of this social network is created and maintained by collective human effort, which has the potential to be extremely beneficial for the society in turn. The study of social networks has attracted the attention of researchers from a diverse variety of fields,

The associate editor coordinating the review of this manuscript and approving it for publication was Walter Didimo¹.

including sociology, economics, historiography, computer science, and information systems, as this subject is understood to have a major impact on society.

Enabled by the widespread adoption of communication tools such as computer technologies and the internet, online social networks have been adopted in everyday life for most people in many, if not most, human societies. Facebook¹ and Twitter² are among the most popular online social networks, allowing users to share thoughts, make friends and communicate with each other, among many other features. Moreover, some social networks have also been created that are oriented towards specific media or subjects, such as Flickr³ for photo sharing, YouTube⁴ for video sharing, SlideShare⁵ for sharing slides, and ResearchGate⁶ for researchers. The number of online social network users has massively increased over the last two decades, such that many people tend to be more active in online social networks than in the offline, physical social network, i.e., the physical society. It is clear that online social networks have become an important part of our collective life.

In recent years, EBSNs such as Meetup⁷ and Eventbrite⁸ have also emerged [1], [2]. These services provide convenient platforms for people to create, distribute, and organize social events. Typically, users can post new events on EBSNs ranging from informal meetings to formal activities, as well as join existing events, post comments, share photos, etc. In addition to these typical online social network activities, EBSNs also promote face-to-face offline social interactions, in which some members participate in an event in the physical world. Thus, EBSNs not only provide an online virtual space where users can exchange thoughts and share experiences, but also capture offline in-person social interactions. Specifically, an EBSN consists of both online social network (i.e., group membership, event publication) and offline social network (i.e., event participation). Therefore investigating such networks may be expected to be of considerable benefit. Many EBSNs have attracted massive userbases and have been experiencing rapid business growth. Increasing numbers of people have been joining activities advertised online in the real world, and people with common interests often prefer to share their experiences and thoughts in private online groups. For example, as of 2020, Meetup had 49 million members, creating some 15,106 social events per day. As such networks continue to grow and thrive, the present work is motivated by the special feature of EBSNs, that is, their structure as combined heterogeneous social networks.

The tendency of people to come together and form groups or communities is inherent in the structure of society; and the processes by which such groups take shape and attract

members has been a theme in a considerable body of social science research [3], [4], [5]. With widespread adoption of the Internet, new forms of organizing such as online communities have become prevalent on nearly every online social media platform. Among global Internet users, 76% have participated in an online community [6]. Although their functionality varies markedly, these platforms all provide facilities enabling multi-person social interactions. Businesses use online communities to facilitate peer-to-peer customer support [7], build brand loyalty [8] and foster knowledge sharing and collaboration among their employees [9], while individuals join online groups to exchange information, to interact with likeminded others, and to organize and participate in collaborative work or collective action [10], [11]. The development of internet technologies has significantly transformed online communities as new type of entity, with the potential to create considerable benefits by promoting user interactions and information diffusion.

The proliferation of online communities, despite benefiting both social media and users in many ways, has also created challenges for Internet users, who must find online communities to join from amid an overwhelming volume and variety of such communities. Over the last few years, the growth of online users has slowed down, whereas the number of Internet offerings – public websites and communities - has grown exponentially [12]. Social media users, therefore, generally must expend considerable effort to choose the most useful online communities for a given purpose. However, ordinary users not only cannot usually express their preferences accurately when using vertical search engines to find groups, but may even prefer to simply be told directly which groups they should join [13]. Moreover, groups play a significant role in EBSNs, in which events are organized by a group, and users can join some groups relevant to their interests. It may be considered that people within a given group have weak friendly relationships with each other.

Consequently, helping users find suitable online groups is a promising area for the application of recommender systems, which has been of interest both to the industry and to academic researchers. Studies have been conducted on recommendation in other applications such as recommending friends [14], [15], events [16], driving routes [17], movies/music [18], [19], image tags [20], news [21], learning performance [22], etc. However, thus far the group recommendation has been explored only sporadically [23], [24], especially for EBSNs.

Currently, groups recommendation in EBSNs are typically only based on interest information filled in by the users, which is not very effective, or based on the idea that groups joined by the friends of a user are also likely to be joined by this user. Both these methods may be considered quite naïve and do not reflect users' real intentions. Hence, in this study, we propose a new comprehensive group recommender system based on a random walk algorithm to address this problem. To the best of our knowledge, no prior research has been conducted on group recommendation based on a random

¹www.facebook.com

²twitter.com

³www.flickr.com

⁴www.youtube.com

⁵www.slideshare.net

⁶www.researchgate.net

⁷www.meetup.com

⁸www.eventbrite.com

walk algorithm in EBSNs, which are identified as SNs comprising both online and offline social interactions. In fact, our proposed recommender framework is not restricted to group recommendation. It can also be easily extended to other kinds of recommendations, such as friend recommendation and event recommendation [25], [26]. To demonstrate the effectiveness of the proposed approach, we focus on group recommendation, i.e., personalized recommendation of event-based groups to users.

The present work includes three main contributions to the relevant literature, as given below.

(1) We propose a unified group recommendation framework. Specifically, by constructing a heterogeneous, augmented graph, the proposed approach can incorporate all the necessary information into a single social graph to make better recommendations, which should be more suitable for EBSNs requiring integration of both online and offline social networks. Furthermore, new information can be easily added to boost the effectiveness of our proposed recommendation system with little change to the algorithm.

(2) We augment the random walk algorithm to adapt to the constructed heterogeneous augmented graph containing various types of nodes and links in a single graph. In prior works, random walk algorithms have been applied in homogenous graphs containing only relationships among entities of a given type. Our approach allows the incorporation of more types of nodes into a single graph and makes the assignment of link weights more effective in a simple manner, because the proposed graph structure already captures the inherent indirect relationships among seemingly unrelated entities.

(3) We conducted extensive experiments on real datasets to compare our proposed algorithm to other commonly used recommendation algorithms. The results demonstrate the superiority of our novel group recommendation method based on a heterogeneous augmented graph structure and a random walk algorithm.

The remainder of this paper is structured as follows. We describe related works in Section II and present the heterogeneous augmented graph construction method and group recommendation method based on random walk in Section III. The dataset and evaluation metrics used are introduced in Section IV. We present experimental results verifying the efficacy of our proposed recommender system in Section V and conclude the work in Section VI.

II. RELATED WORK

The present work is closely related to three key areas of research, including communities on social network services (SNS), community recommender systems, and random walk algorithm-based recommendation.

A. COMMUNITIES/GROUPS ON SOCIAL NETWORK SERVICES

The term “community” denotes a collection of people who have joined an explicitly defined group within an organization that supports the formation of such groups and provides

facilities for the members to exchange information with each other and engage in other joint activities. In some circumstances, other terms such as “group” are commonly used in the same meaning that we apply to “community” here. Thus, we will use the terms “community” and “group” interchangeably in this paper. Traditional social communities or groups provide many offline interactions for their members, including opportunities for affiliation or companionship, social support, information exchange capabilities, and support for collective action.

With the development of technical infrastructure, the Internet has created new opportunities and platforms (such as Flickr, Facebook, etc.) for people to interact. Likewise, social network services (SNS) provide basic online communication capabilities to support the development of interpersonal relationships, feelings of companionship, and perceptions of affiliation. Communities in SNSs also encourage a variety of online interactions such as discussion and knowledge sharing, access to information and dissemination of ideas, as well as other collective activities such as software development or political action.

Unlike these traditional social structures which function either online or offline, EBSNs are distinct as an emergent type of social network identified by the co-existence of both online and offline social interactions. Reference [2] is the first work to comprehensively study properties of EBSNs including time and location patterns, network properties, and several popular issues in social networks such as community detection and information flow. Subsequently, this new type of social structure has recently attracted increasing research interest. For example, Liu et al. explored the roles of event size and interactivity in social networking behaviors [2], while Zhang et al. exploited matrix factorization to model interactions on EBSNs by considering location features, social features, and implicit patterns simultaneously in a unified model [13]. Xu and Liu proposed a semantic-enhanced and context-aware hybrid collaborative filtering for event recommendation [27]. However, prior research on this topic has only presented some algorithms for certain problems, and further research remains necessary on EBSN-based community/group recommendation.

B. ONLINE GROUP RECOMMENDER SYSTEM

In this section, we provide an overview of three categories of group recommender systems, including content-based, collaborative-filtering-based, and hybrid recommender systems [13], [27], [28], [29].

Content-based recommender systems have been designed to use the content correlation between items and past preferences of active users for recommendation. For example, Spertus et al. presented an empirical comparison of community recommendations for a given user, based on their similarity to communities to which users actually belonged [30]. Bagher et al. proposed the concept of trends to capture the interests of users in selecting items among different groups of similar items [31].

Collaborative filtering is regarded as one of the most widely-used and successfully-developed recommendation approaches [32]. Collaborative filtering methods may be further divided into two typical classes, including memory-based and model-based approaches [33]. Memory-based approaches mainly focus on finding similar users or items for recommendations. Two algorithms, namely the Pearson correlation coefficient (PCC) algorithm [34] and the vector space similarity (VSS) algorithm [35] have commonly been applied in both user-based and item-based approaches as similarity computation metrics. In contrast to the memory-based approaches, the model-based approaches to collaborative filtering fit prediction models based on training data and then apply them to predict users' preference on items [36], [37]. Algorithms in this class include clustering methods [38], [39], aspect models [40] and Bayesian networks [36], [41]. Liu et al. presented a recommendation model fusing social relations and item contents with user ratings by modifying the model using a Bayesian Probabilistic Matrix Factorization algorithm [42]. He et al. proposed the use of neural networks to improve collaborative filtering [43]. Although collaborative filtering has been shown to be effective when users expressed enough ratings to share common ratings with other users, it has tended to perform poorly for so-called cold-start users [44], [45], [46].

Hybrid recommender systems combine both content-based recommendation and collaborative filtering to make predictions [28], [47]. Generally, four classes of approaches have been used to construct hybrid recommender systems. The first such class of hybrid systems involves the combination of separate collaborative and content-based recommenders by combining outputs of these recommenders linearly, or using an evaluation metric to dynamically choose the best recommender [48], [49]. The second class adds content-based characteristics to collaborative models, which can be of benefit in addressing the cold-start user problem of collaborative filtering by employing user information in the prediction process [28]. A collaborative content-based approach is used in another class, incorporating collaborative factors into content-based models [50]. The fourth class unifies content-based and collaborative filtering [21], [51], [52]. For example, Reference [51] described an unified probabilistic model combining collaborative and content information in a coherent manner to yield better recommendations overall in terms of accuracy and cold-start items. Reference [21] proposes a personalised news recommendation framework named HYPNER to combines both collaborative filtering-based and content-based filtering methods. Reference [52] incorporates public contextual metadata and paper-citation relationship information into both content-based and collaborative filtering approaches separately to enhance the recommendation accuracy.

In contrast to the abovementioned prior research on community recommendation, event-based communities mainly consist of users, attributes, and offline events, all of which could influence users' decisions and thus need to be

considered [13]. In this study, we propose a random walk-based recommendation model to accommodate the special nature of EBSN community recommendation, using both attributes and structure properties for recommendation, and applying the ideas of user-based and entity-based collaborative filtering approaches in a coherent manner.

C. RANDOM WALK-BASED ALGORITHM

Random walk algorithm was first introduced by Pearson [53], and subsequently used in many fields such as ecology, physics, and computer science. In general, a random walk is a mathematical formalization of a path formed by a succession of random steps. Random walk algorithm is commonly used on graphs [54], [55] as compared to other algorithms [56], [57], where each edge is tagged with a specified probability of traveling via this edge in the random walk. Similarity and transitive associations between nodes can be easily computed based on random walks, enabling them to be effective in recommender systems [58], [59], [60], [61].

Bogers presented the ContextWalk algorithm, incorporating different types of contextual information in a traditional random walk algorithm [62]. This technique modeled the browsing process of a user on a video website by taking random walks over a contextual graph. In the most similar approach to that of the present work available in the relevant literature, Yin et al. proposed a unified framework for link recommendation using a random walk algorithm [63]. In their graph construction process, both users and attributes were treated as nodes, and a link between two nodes was defined to exist if the nodes represented persons who were friends or one node representing a person included an attribute denoted by the other node. Jiang et al. proposed a Bayesian personalized ranking (BPR)-based machine learning method called HeteLearn to learn the weights of links in an information network. In order to model user preferences for personalized recommendation, a generalized random walk with restart model was proposed [60]. Manju et al. sought to solve the cold start problem in research paper recommendation by integrating social network interaction factors based on a random walk [64]. Zheng et al. proposed a personalized tag recommendation for social images based on convolution features and a weighted random walk [20]. In particular, for a given image, they selected its visual weights and determined the weight of each neighbor by mining the influence of user group metadata. Afterwards, the weighted random walk algorithm was implemented on a neighbor-tag bipartite graph.

Our work differs from prior works in several key aspects. First, most recommendation studies (e.g., random walk algorithm) have been limited to a single graph, including only one or two types of nodes. In contrast, our work is oriented to EBSNs, as they possess more rich information than other applications, including both online and offline social networks. Our proposed approach constructs a heterogeneous graph containing various kinds of nodes and edges. Prior works on traditional social networks are not well applicable to EBSNs. Second, although friend/item recommendation

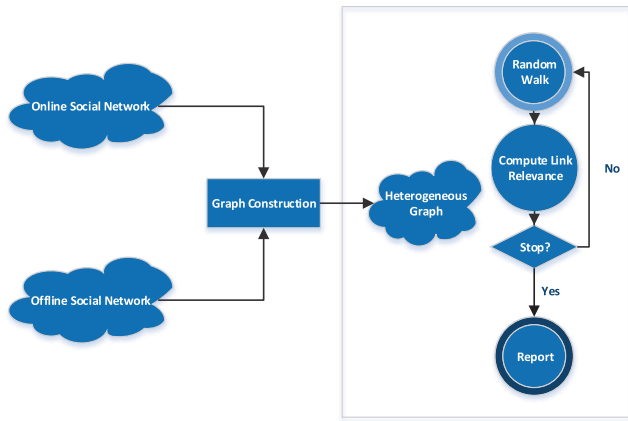


FIGURE 1. Data Flow Procedure of Proposed Recommender System.

problems for EBSNs have frequently been studied, group recommendation has not been studied previously. On some popular EBSNs, such as Meetup, groups are recommended based only on interests and common groups between friends within a group, which is naïve and not very effective, as shown in the Section V in this paper.

III. GROUP RECOMMENDATION

In this section, we present a group recommendation framework based on a random walk algorithm. We first introduce the formal definition of our group recommendation problem and then elaborate on our recommendation framework, which is divided into two parts, including graph construction and the random walk algorithm. The data flow of our recommender system is illustrated in Fig. 1. In our proposed recommender system, we first construct a heterogeneous augmented graph capturing all the available information related to the recommendation, including event participation, group membership, and interest information. Based on the relationship graph, we apply the random walk algorithm to calculate recommendation scores (i.e., the link relevance) for every group and then select the top N groups to recommend to the user. A termination condition controls the number of iterations of the random walk algorithm.

A. PROBLEM DEFINITION

In this study, we propose a recommender system to recommend groups to users in EBSNs. In this subsection, we first give a formal definition of the problem of group recommendation.

The group recommendation problem is defined as the task of recommending some groups that are most likely to interest a given user based on their activity history and personal information. That is, we want to maximize the probability that the user will join the recommended groups in the future. If we can recommend suitable groups to users, users will benefit from increased convenience in finding relevant interest groups, while groups can accept more promising members.

TABLE 1. Node Types and Denotations.

Note Type	Entity in EBSNs	Denotation
group node	group	g
user node	user	u
interest node	interest	i
event node	event	e

TABLE 2. Edge Types and Denotations.

Edge Type	Corresponding Relationship in EBSNs	Denotation
membership edge	a user joined a group	$\{u, g\}$
user-interest edge	a user has an interest	$\{u, i\}$
group-interest edge	a group is oriented towards an interest	$\{g, i\}$
event-organizer edge	an event belongs to a group	$\{e, g\}$
participation edge	a user participates an event	$\{u, e\}$

Formally, given the groups’ memberships, groups’ interests, users’ interests, events’ organizers (describing which group created a given event) and events’ participation records, we aim to recommend a list of groups that may interest the given user in order of the joining probability (i.e., recommendation score). That is, the recommended groups should be ordered in decreasing order of recommendation scores with regard to the given user and should not contain any group that the user has already joined.

In the following subsections, we introduce our proposed group recommendation framework. The group recommendation problem is modeled as link recommendation, which can be solved using a random walk algorithm on a graph constructed from social networks. Thus, our framework consists of two parts, including a graph construction process and a random walk algorithm.

B. GRAPH CONSTRUCTION

In this subsection, we describe the proposed approach to construct a heterogeneous augmented graph from given social network information having different types of nodes and edges, each of which indicates a different physical meaning. We first present the graph structure, and then describe the weight calculation of edges of various types.

1) GRAPH STRUCTURE

Based on the group memberships, group interests, user interests, event organizers and event participation records, a heterogeneous augmented graph denoted as $G(V, E)$, where V is the node set and E is the edge set, is defined in Table 1 for node types and Table 2 for edge types.

It may be observed that there exist four different types of nodes and five types of edges in the constructed graph. We provide an example as follows. Let there exist two user nodes u_1 and u_2 , three group nodes g_1, g_2 and g_3 , two event nodes e_1 and e_2 , and two interest nodes i_1 and i_2 . Group g_1 has two members u_1 and u_2 , one interest attribute i_1 , and once held an event e_1 , which was participated in by user u_2 . Group g_2 has no members, an interest i_2 and organized an event e_2 ,

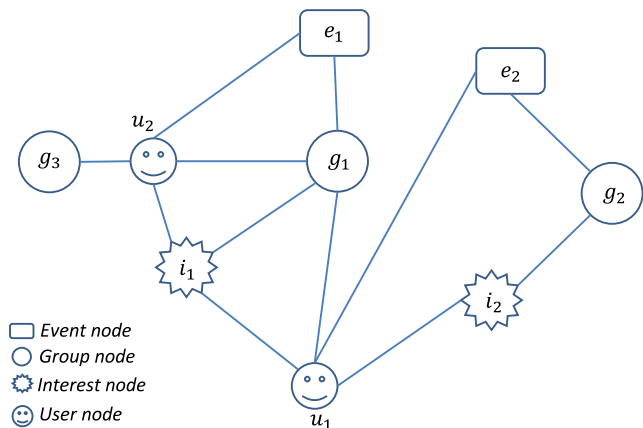


FIGURE 2. An Example Graph.

which was participated in by user u_1 . Group g_3 has only one member u_2 . Additionally, user u_1 has interests i_1 and i_2 , and user u_2 has interests i_1 . According to the above description, the graph is constructed as shown in Fig. 2.

It should be noted that the graph thus constructed is not a strict undirected graph as an edge is defined with different weights for different directions. That is, the weight of $w(a, b)$ is different from the weight of $w(b, a)$. We elaborate on this point in subsection III-B2.

The graph $G(V, E)$ is inherently a heterogeneous augmented graph as there are different types of nodes and different types of edges within a single graph. This graph captures all the information that EBSNs provide in a natural manner. We do not assume two members within a given group are friends, because, in fact, they typically only have a weak friendship with each other, and it is even possible that some members do not know of the existence of others. Instead, we retain the original indirect relationship between members within a given group, that is, they are considered to have relations because they have joined the same group. Similarly, the relationships among users who participated in the same event are expressed in terms of a participation relationship. The preservation of direct relationships enables a more natural expression of the complicated relationships in EBSNs and also enables a simple but effective weight assignment scheme, as detailed in the next subsection.

2) WEIGHT CALCULATION

Although the heterogeneous augmented graph constructed above is able to depict the structure of EBSNs in a natural and direct way, it remains insufficient to express the strength of the relationships contained in it. In other words, the importance degree of relationships (i.e., edges) has not yet been included in the graph. In this subsection, we introduce our approach to assigning importance (i.e., weight) to each relationship (i.e., edge). The weight assigned to an edge measures the importance of the end node to the start node and the strength of a relationship. We adopt a uniform weighting scheme, which can depict the weights of different edges,

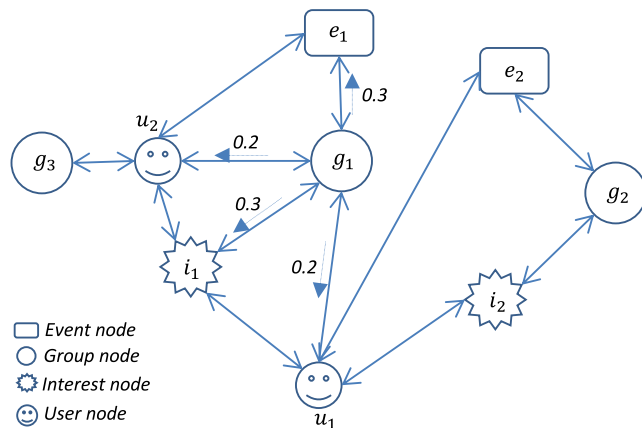


FIGURE 3. Weighted Graph.

as our heterogeneous augmented graph is already able to express the complicated relationships and importance degree differences due to indirect relationships, as explained below after the introduction of our applied weighting scheme.

We describe the weight calculation for each type of edge in Table 3. It should be clear that $w(a, b)$ denotes the weight for the edge (a, b) , and g, u, i, e denote the group, user, interest, and event node, respectively, as shown in Table 1. As mentioned above, $w(a, b)$ does not necessarily equal $w(b, a)$.

As we can see, there are ten weight functions for five types of edges. Although the graph thus constructed is more similar to an undirected graph, the weight functions for different directions are distinct, expressing different underlying meanings. Hence the graph may be considered as similar to a directed graph. Moreover, constraints are included about the damping factors in the weight functions for normalization. From the perspective of the random walk process, the weight of an edge represents the probability of walking via this edge; thus weights from all edges emitted from one node should aggregate to 1. We set the constraints as given in Equations (1)-(4), making weights originated from one node summing up to 1. For example, for all the edges originated from a group node, $\sum_{e \in N_e(g)} w(g, e) + \sum_{i \in N_i(g)} w(g, i) + \sum_{u \in N_u(g)} w(g, u) = 1$.

$$\alpha + \alpha' + \alpha'' = 1 \tag{1}$$

$$\beta + \beta' = 1 \tag{2}$$

$$\gamma + \gamma' + \gamma'' = 1 \tag{3}$$

$$\delta + \delta' = 1 \tag{4}$$

After weight calculation, parts of the edge weights are shown in Fig. 3 with reference to the example graph described in the previous subsection. Here, we assume $\alpha = \alpha' = \gamma = \gamma' = 0.3, \alpha'' = \gamma'' = 0.4, \beta = \beta' = \delta = \delta' = 0.5$.

Although defining a weight function for each type of edge may seem somewhat complicated and tedious, in this case it is quite straightforward. All functions are defined by a uniform weighting scheme in which each outgoing edge shares an equal proportion of weight, all of which aggregate to 1.

TABLE 3. Weighting Functions for Different Types of Edges.

Edge	Weight Function	Parameter	Description
(u, g)	$w(u, g) = \frac{\gamma'}{ N_g(u) }$	$N_g(u)$: the set of group nodes connected to the user node u γ' : a damping factor ($0 \leq \gamma' \leq 1$) used to control the importance of this type of relationship for the user u and for normalization	membership edge from a user node u to a group node g
(g, u)	$w(g, u) = \frac{\alpha''}{ N_u(g) }$	$N_u(g)$: the set of user nodes connected to the group node g α'' : damping factor	membership edge from a group node g to a user node u
(u, i)	$w(u, i) = \frac{\gamma''}{ N_i(u) }$	$N_i(u)$: the set of interest nodes connected to the user node u γ'' : damping factor	user-interest edge from a user node u to an interest node i
(i, u)	$w(i, u) = \frac{\delta'}{ N_u(i) }$	$N_u(i)$: the set of user nodes connected to the interest node i δ' : damping factor	user-interest edge from an interest node i to a user node u
(g, i)	$w(g, i) = \frac{\alpha'}{ N_i(g) }$	$N_i(g)$: the set of interest nodes connected to the group node g α' : damping factor	group-interest edge from a group node g to an interest node i
(i, g)	$w(i, g) = \frac{\delta}{ N_g(i) }$	$N_g(i)$: the set of group nodes connected to the interest node i δ : damping factor	group-interest edge from an interest node i to a group node g
(g, e)	$w(g, e) = \frac{\alpha}{ N_e(g) }$	$N_e(g)$: the set of event nodes connected to the group node g α : damping factor	event-organizer edge from a group node g to an event node e
(e, g)	$w(e, g) = \beta$	β : damping factor	event-organizer edge from an event node e to a group node g
(u, e)	$w(u, e) = \frac{\gamma}{ N_e(u) }$	$N_e(u)$: the set of event nodes connected to the user node u γ : damping factor	participation edge from a user node u to an event node e
(e, u)	$w(e, u) = \frac{\beta'}{ N_u(e) }$	$N_u(e)$: the set of user nodes connected to the event node e β' : damping factor	participation edge from an event node e to a user node u

Thus, the weight scheme is quite simple and easy to apply. Thanks to the informative structure of the heterogeneous graph constructed previously, the simple weight scheme is able to express the valuable relationships and importance degrees among every pair of entities (i.e., group, user, interest, and event). For example, a user u_1 who participated in quite a number of events organized by the group g_1 should be important to the group g_1 , i.e., the user u_1 and the group g_1 should have a strong relationship ($w(u_1, g_1)$ should be large). Although we simply use a uniform weighting scheme, observing that each user node in the group node g_1 shares an equal weight, there exist other paths via event nodes (that are attended by the user u_1) between the group node g_1 and the user u_1 such that the ranking of user node u_1 also gains rankings from the group node g_1 indirectly via these event nodes. In other words, the edge weight can be thought of the possibility of walking via this edge, and the closeness between two nodes can be thought of the possibility of walking from one node to the other node via multiple paths in the graph. For example, possible paths from a group node g to a user node u can be:

- group node g to the user node u directly via the group-to-user edge
- group node g to the user node u indirectly via the group-to-interest edge and interest-to-user edge
- group node g to the user node u indirectly via the group-to-event edge and event-to-user edge

We will elaborate on this idea in Section III-C. Hence, the seemingly complex indirect relationship has been captured naturally in the underlying graph structure, and the uniform weighting scheme is able to capture the strength of the inherent relationship.

C. RANDOM WALK ON HETEROGENEOUS GRAPH

In this section, we present a random walk with restart algorithm based on the newly constructed heterogeneous graph to simulate group membership behavior over time. This algorithm assumes that if a random walk on the graph from a given user node based on the edge weight (indicating the probability from the start node to the end node) finally arrives at a group node, it is considered highly possible that the given user would be interested in joining the group in the future. The assumption is straightforward and provides some intrinsic features of group membership prediction. For example, if a group shares more attributes with a given user, or the user once participated in some events organized by the group, or the group has been joined by most of the friends⁹ of this user, the probability that a walk proceeds from the user node to this group node is high, indicating that the group is considered promising for the user to join in the near future. Moreover, if the group node is close to the given user node, the random walk probability to that group node is also high, which is consistent with the fact that the nearby groups with the user in the graph are more likely to be joined by the user, as their relationship is constructed via few intermediaries.

The convergent probabilities of random walk algorithms starting from a given user node are considered as the link relevance between the user node and the respective nodes in the probability distribution, which is simply the recommendation probability discussed above. Here, we use the random walk with restart algorithm on the heterogeneous graph with user, event, interest and group nodes to calculate the link relevance

⁹The “friends” here mean the co-members of the user. That is, if two users join the same group, they can be called “friends”. Although this is not a definition of friendship, we use it here for clarity.

for a particular user node u^* . Indeed, the heterogeneous graph described above is homogenized by node coding and weight normalization in the proposed approach. Although there are different types of nodes and edges, we do not need to differentiate them in the random walk process. We can treat the heterogeneous graph as a normal homogeneous graph when we perform the random walk process. Thus, the conventional random walk with restart algorithm can be applied to our group recommender system with few modifications.

The random walk with restart function to calculate the link relevance for a particular user u^* is shown in Equation 5.

$$r_{a^+} = (1 - \theta) \sum_{a' \in N_a(a^+)} w(a', a^+) \times r_{a'} + \theta \times r_{a^+}^{(0)} \quad (5)$$

where a^+ is an arbitrary node in the graph, r_{a^+} is the link relevance of a^+ with regard to the user node u^* , i.e., the probability of walking to the node a^+ from the user node u^* , where $a \in \{u, e, g, i\}$ and u, e, g, i denote user node, event node, group node, and interest node respectively, $N_a(a^+)$ is the set of type “a” nodes connected to node a^+ , a' is a specific instance of type “a” nodes, $w(a', a^+)$ is the weight from node a' to node a^+ , $r_{a'}$ is the link relevance of node a' , $r_{a^+}^{(0)} = 1$ if $a^+ = u^*$, otherwise $r_{a^+}^{(0)} = 0$, and θ is the restart probability.

Algorithm 1 shows our proposed group recommendation algorithm based on random walk with restart. The input of the algorithm includes the EBSN graph G , the particular user u^* for whom to recommend groups, the maximum number of iterations i_{max} and a threshold t used to control the termination of the algorithm, and the number of groups n to recommend. The output is an ordered list of recommended groups in non-increasing order of link relevance. In the pseudo-code for Algorithm 1, Lines 1-5 perform initialization. Lines 6-13 comprise an iterative process to update the link relevance of each node. Two termination conditions may cease the iteration, in the case that either the number of iterations exceeds the maximum allowable, as set by the parameter i_{max} (Line 6), or that the total difference of link relevance of all nodes between the previous iteration and the current iteration is less than a threshold t (Lines 10-12). Line 14 sorts the group nodes in non-increasing order in terms of link relevance, and Line 15 returns the first n recommended groups.

IV. DATA DESCRIPTIONS AND EXPERIMENTAL DESIGN

We conducted extensive experiments on real datasets to demonstrate the effectiveness of our proposed recommender system. In this section, we describe the dataset, metrics and baseline methods used to evaluate the performance of our proposed system, while the experimental results are provided in Section V.

A. DATASET

The datasets used were originally crawled from the Meetup EBSN,¹⁰ provided by authors of a prior work [2]. The data statistics are shown in Table 4, where RSVP represents event

¹⁰https://github.com/wuyuehit/meetup_dataset

Algorithm 1 Group Recommendation Algorithm

Input: EBSN $G = \langle V, E, W \rangle$, user u^* , maximum iteration times i_{max} , a threshold t , recommended number n

Output: An ordered list of groups recommended for u^*

```

1: if  $a^+ \neq u^*$  then
2:    $r_{a^+} \leftarrow 0$ 
3: else
4:    $r_{a^+} \leftarrow 1$ 
5: end if
6: for  $i \leftarrow 1$  to  $i_{max}$  do
7:   for each  $a^+ \in V$  do
8:      $r_{a^+} = (1 - \theta) \sum_{a' \in N_a(a^+)} w(a', a^+) \times r_{a'} + \theta \times r_{a^+}^{(0)}$ 
9:   end for
10:  if  $\text{diff}(\{r_{a^+}\}_i, \{r_{a^+}\}_{i-1}) < t$  then
11:    break
12:  end if
13: end for
14: Sort the group node set in non-increasing order based on the  $r_{a^+}$ 
15: return the first  $n$  groups in the set

```

TABLE 4. Dataset Statistics.

Users	Events	RSVPs	Groups	Memberships	Interests
5,153,886	5,183,840	42,733,136	97,587	10,704,068	77,810

registration. In this dataset, the online social network is constructed by capturing the membership of online social groups (i.e., membership of a user in a group forms a user-group relationship) and the attributes of users and groups. An offline social network is generated in the same way based on the participation in social events and their organizers.

B. EVALUATION CRITERIA

To evaluate the quality of our recommender system, we adopt four standard evaluation measures, including mean reciprocal rank (MRR), precision, recall, and F-Measure, which have been commonly employed in prior recommender system research [65]. We compared our algorithm with other baseline methods in terms of these metrics. MRR is defined in Equation 6.

$$MRR = \frac{1}{|S|} \sum_{u \in S} \frac{1}{rank_u} \quad (6)$$

where S is the set of sampled user nodes and $rank_u$ is the rank of the first correctly recommended group for user u .

Precision is defined as the number of correct groups recommended divided by the total number of groups recommended, and recall is defined as the percentage between the number of correct groups recommended and the number of all true groups that the user actually joins.

Formally, precision and recall are defined by Equation 7 and Equation 8.

$$P@K = \frac{1}{|S|} \sum_{u \in S} \frac{\# \text{ of correct groups in top } K \text{ groups}}{K} \quad (7)$$

$$R@K = \frac{1}{|S|} \sum_{u \in S} \frac{\# \text{ of correct groups in top } K \text{ groups}}{\# \text{ of true groups in the ground truth}} \quad (8)$$

where $P@K$ and $R@K$ represent the precision and recall in the top K recommended groups. F-Measure is a combined metric of precision and recall, as defined by Equation 9.

$$F - \text{measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (9)$$

In the experiment, we performed 4-fold cross validation to evaluate the performance of the proposed method. In particular, we partitioned the membership dataset into four parts, three of which were used for prediction, and one of which was used for testing. To remove noisy data and guarantee the reliability of the experimental results, we randomly sampled 100 users who had joined at least 5 groups but not most of the groups. For each user, we recommended top- K groups.

C. EXPERIMENTAL DESIGN

1) BENCHMARK MODELS

We compared our proposed approach with four baseline methods. In general, two important features need to be considered in the recommendation, i.e., semantic features (i.e., the interest) and the relationship structure features. However, the baseline methods only consider one of these two features. The baseline methods used for comparison are listed as follows.

- Random recommendation (denoted by **RecByRandom**): Randomly choose some groups to recommend. The performance of this method is expected to be the worst.
- Interest-based recommendation (denoted by **RecByInterest**): This method is the most popular group recommendation method used by some EBSNs such as Meetup [66]. It only considers the interest similarity between groups and users to perform recommendation. That is, groups with similar interests will be recommended to users.
- Interest- and neighborhood-based recommendation (denoted by **RecByInt&Com**): This recommendation method is also frequently used by some EBSNs and social networks [14]. It considers two factors in the recommendation, i.e., interest similarity (same as **RecByInterest**) and collaborative friendship (i.e., group memberships in the context of EBSN). In short, groups that are joined by their friends are considered as the recommendation candidates, which are then ranked by their common interests.
- Latent factor model with location features (denoted by **PTARMIGAN**) [13]: This recommendation method is combining latent factor model with location features for EBSN. However, due to the unavailability of source code or some implementation details, we directly use their published precision results as we are using the same dataset provided in [2].
- **Katz Centrality** [67]: This recommendation method is based on the graph structure (specifically, the paths in

TABLE 5. Parameter Settings.

α	α'	α''	β	β'	γ	γ'	γ''	δ	δ'	θ
0.1	0	0.9	0.3	0.7	0.3	0.6	0.1	0.3	0.7	0.3

TABLE 6. MRR for Different Methods.

Method	MRR
Random	0.001
RecByInterest	0.027
RecByInt&Com	0.043
Katz Centrality	0.345
HeteroRandom	0.511

the social network graph), calculating the recommendation score of a group g with regard to a user u by using Equation 10.

$$\text{score}(u, g) = \sum_{l=1,2,\dots,\max} \tau^l |path_{u,g}^{<l>}| \quad (10)$$

where τ is the damping factor and $path_{u,g}^{<l>}$ is the set of all length- l paths from node u to node g . We consider the paths with lengths of no more than 4. Our method is referred to as **HeteroRandom**.

2) EXPERIMENTAL SETTINGS

To obtain good recommendation results, the parameters must be tuned in the weighting and link relevance functions, which impacts the effectiveness of the recommendation. There are seven parameters to be tuned, i.e., α , α' , β , γ , γ' , δ in Equations (1)-(4) and the restart probability θ in Equation (5), in addition to which other parameters (i.e., α'' , β' , γ'' , δ') can be derived easily according to Equations (1)-(4). We manually tune these parameters in the following steps.

- 1) Initialize all parameters to an average value, e.g., $\alpha = \alpha' = 0.33$, $\beta = 0.5$, $\gamma = \gamma' = 0.33$, $\delta = 0.5$, $\theta = 0.5$.
- 2) For each parameter, either increase or decrease this parameter with other parameters fixed at their current best values until the metric becomes worse; thus, we get the local best value for this parameter. In the tuning experiment, we use the precision metric, and we do observe the linearity between one parameter and the metric. That is, if decreasing the parameter makes the metric better, increasing it will not. Hence the exploration would only take place in one direction.
- 3) Repeat Step 2) for a fixed round (in the experiments, we set it to 10).

Through intensive manual tuning as mentioned above, we obtained the best settings of parameter values as shown in Table 5.

V. EXPERIMENTAL ANALYSIS

In this section, we evaluate our proposed recommender method and compare it with the baseline methods to demonstrate its effectiveness. We compared our proposed method with the baseline methods in terms of the evaluation metrics introduced in Section IV-B.

TABLE 7. Precision Results.

%	P@1	P@5	P@10	P@20	P@50	P@100	P@200	P@500
HeteroRandom	42.00	30.70	23.35	16.17	9.22	5.93	3.79	1.97
Katz Centrality	32.35	14.30	7.60	3.87	1.60	0.82	0.47	0.25
PTARMIGAN	16	12 (P@3)	-	-	-	-	-	-
RecByInt&Com	2.04	0.82	0.87	0.92	0.64	0.60	0.50	0.44
RecByInterest	0.98	0.59	0.59	0.39	0.26	0.21	0.16	0.13
RecByRandom	0.00	0.00	0.00	0.01	0.02	0.03	0.02	0.03

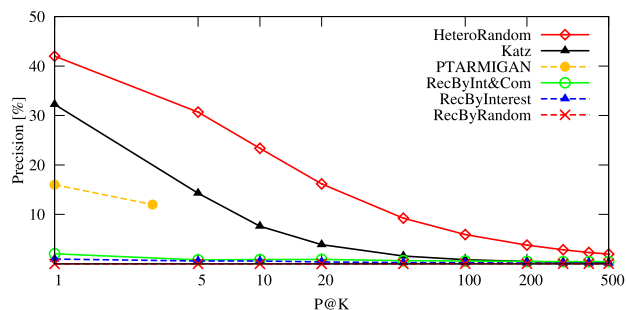


FIGURE 4. Precision Results.

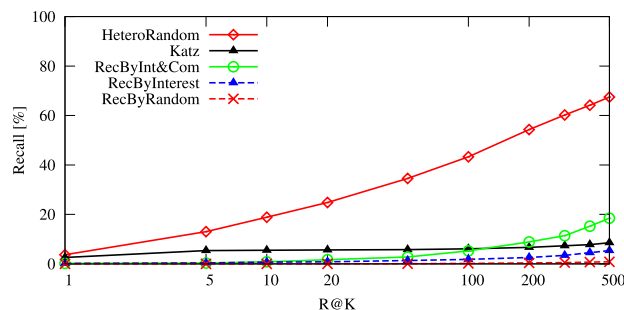


FIGURE 5. Recall Results.

A. EXPERIMENTAL RESULTS

In this subsection, we present the results of the four commonly used baseline methods and of our random walk algorithm in terms of MRR, precision, recall and F-Measure.

The results in term of the MRR metric are listed in Table 6, from which it may be observed that our method had the highest MRR value of 0.511, indicating that on average the first two groups recommended by the proposed approach accurately reflected the groups actually joined by the user subsequently. In comparison, the other three methods achieved only very poor MRR values. Katz Centrality achieve a relatively high MRR (0.345), which is mostly attributed to our informative graph structure.

Precision measures of the five methods are shown in Fig. 4 and Table 7, from which it may be observed that our method outperformed the other methods for all $P@K$ metrics. Katz Centrality achieved a relatively good precision, which, however, was primarily due to our proposed graph construction method. The latent factor model with location features achieves a relatively high precision result between 10%-20%, which is still far lower than our method. The other four methods, i.e., random recommendation, interest-based recommendation and interest- and neighborhood-based recommendation, achieved nearly the same precision values, which were significantly below 10%. Our method achieved 42% precision in the top 1 group and over 30% precision in the top 5 groups. As K increases, the precision of all methods decreased because most users subsequently joined fewer than 20 groups (as simulated by the testing partition).

We present the recall results for different methods in Fig. 5 and Table 8, from which it may be observed that our method, denoted as HeteroRandom, achieved significantly better recall than the other methods. Notably, our method

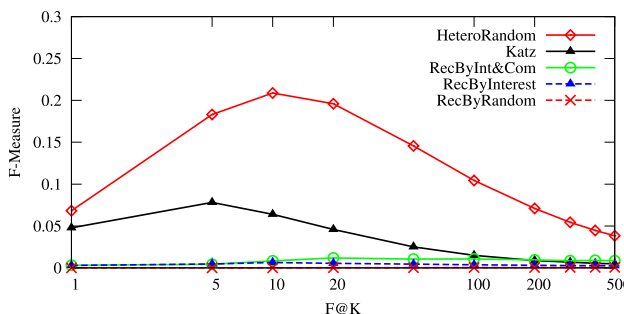


FIGURE 6. F-measure Results.

attained nearly 50% recall at $R@100$. The RecByInt&Com method also achieved a relatively higher recall than RecByInterest, Katz Centrality and Random when K exceeded 100, achieving nearly 20% at $R@500$.

F-Measure results are shown in Fig. 6 and Table 9. It may be observed that our proposed method delivered the best performance in terms of F-Measure. Moreover, we found that in order to achieve relatively good F-Measure, the number of recommended groups should be between 5 to 50, which provides information on how many groups should be recommended to users in EBSNs.

B. FURTHER DISCUSSION

From the experimental results, it is observed that our proposed approach performs significantly better than all the baseline methods, in terms of the four measures, i.e., MRR, precision, recall and F-Measure. It is interesting that Katz Centrality also achieves good performance in terms of MRR and precision. We believe that it is attributed to the heterogeneous graph structure we designed, which is informative, capturing a variety of relationships in EBSNs. The randomness of

TABLE 8. Recall Results.

%	R@1	R@5	R@10	R@20	R@50	R@100	R@200	R@500
HeteroRandom	3.71	13.05	18.87	24.85	34.57	43.31	54.32	67.49
Katz Centrality	2.58	5.38	5.51	5.62	5.77	6.12	6.68	8.54
RecByInt&Com	0.17	0.29	0.81	1.70	2.80	5.28	8.87	18.44
RecByInterest	0.17	0.38	0.69	0.86	1.34	1.83	2.57	5.37
RecByRandom	0.00	0.00	0.00	0.01	0.03	0.21	0.38	0.8

TABLE 9. F-measure Results.

%	F@1	F@5	F@10	F@20	F@50	F@100	F@200	F@500
HeteroRandom	0.0682	0.1831	0.2087	0.1959	0.1456	0.1043	0.0709	0.0383
Katz Centrality	0.0478	0.0782	0.0639	0.0458	0.0251	0.0145	0.0088	0.0049
RecByInt&Com	0.0031	0.0043	0.0084	0.0119	0.0104	0.0108	0.0095	0.0086
RecByInterest	0.0029	0.0046	0.0064	0.0054	0.0044	0.0038	0.0030	0.0025
RecByRandom	0.0000	0.0000	0.0000	0.0001	0.0002	0.0005	0.0004	0.0006

our approach gives us more generalizability and the “restart” probability avoids local optima. Lack of these properties may cause the poor recall rate of Katz Centrality. Other approaches (i.e., RecByInt&Com, RecByInterest) that only consider interest or friend relationship lose the opportunities to utilize other useful information, e.g., the link relevance.

VI. CONCLUSION

In this paper, we have proposed a group recommender system for EBSNs based on a novel heterogeneous augmented graph construction method and a random walk algorithm. The heterogeneous augmented graph constructed in the proposed method is able to capture all relationships and valuable information in EBSNs, while the structure of the heterogeneous augmented graph naturally expresses the intrinsic complex relationships and their strengths, even with a simple uniform weighting scheme. To utilize the heterogeneous augmented graph structure and weight information for group recommendation, we adopted a random walk with restart algorithm on the constructed graph, homogenized by node coding and weight normalization. The results of extensive experiments show that our recommender system achieved better recommendation results in terms of MRR, recall, precision and F-Measure, outperforming other methods by a wide margin. Furthermore, our group recommendation framework is generic and can be extended to other recommendation applications such as event recommendation and friend recommendation with relatively minor modifications.

Despite its good performance, the proposed group recommender system has rooms for improvement. First, the parameter tuning is still a bit tedious. In the near future, parameter tuning may be conducted using machine learning to obtain the optimal values automatically. Second, graph mining of large-scale graphs such as those formed by social media networks is time-consuming, especially for iterative algorithms such as random walk. Future research may further optimize the proposed group recommendation framework through parallel execution.

REFERENCES

- [1] Y. Liang, C. Huang, X. Bao, and K. Xu, “Sequential dynamic event recommendation in event-based social networks: An upper confidence bound approach,” *Inf. Sci.*, vol. 542, pp. 1–23, Jan. 2021.
- [2] X. Liu, Q. He, Y. Tian, W.-C. Lee, J. McPherson, and J. Han, “Event-based social networks: Linking the online and offline social worlds,” in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2012, pp. 1032–1040.
- [3] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan, “Group formation in large social networks: Membership, growth, and evolution,” in *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2006, pp. 44–54.
- [4] J. S. Coleman, *Foundations of Social Theory*. Cambridge, MA, USA: Harvard Univ. Press, 1994.
- [5] L. Jiang, L. Shi, L. Liu, J. Yao, B. Yuan, and Y. Zheng, “An efficient evolutionary user interest community discovery model in dynamic social networks for internet of people,” *IEEE Internet Things J.*, vol. 6, no. 6, pp. 9226–9236, Dec. 2019.
- [6] (2020). *The Rise of Online Communities*. [Online]. Available: <https://blog.gwi.com/chart-of-the-week/online-communities/>
- [7] L. B. Jeppesen and L. Frederiksen, “Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments,” *Org. Sci.*, vol. 17, no. 1, pp. 45–63, Feb. 2006.
- [8] C. E. Porter and N. Donthu, “Cultivating trust and harvesting value in virtual communities,” *Manag. Sci.*, vol. 54, no. 1, pp. 113–128, Jan. 2008.
- [9] J. Li-Ying, Z. Zhang, and Q. Long, “An alternative way to make knowledge sharing work in online communities? The effects of hidden knowledge facilitators,” *Manag. Org. Rev.*, vol. 14, no. 4, pp. 781–825, Dec. 2018.
- [10] X. Wang, B. S. Butler, and Y. Ren, “The impact of membership overlap on growth: An ecological competition view of online groups,” *Org. Sci.*, vol. 24, no. 2, pp. 414–431, Apr. 2013.
- [11] X. Wu, W. Huang, J. Jiang, G. Klein, and S. Liu, “Boundary buffering: Limiting interference and information leakage in design teams,” *Int. J. Oper. Prod. Manag.*, vol. 40, no. 3, pp. 293–313, Feb. 2020.
- [12] P. Zhang, T. Lu, B. Liu, H. Gu, and N. Gu, “Understanding social interaction across social network sites,” *Int. J. Hum.-Comput. Interact.*, vol. 36, no. 19, pp. 1818–1833, Nov. 2020.
- [13] W. Zhang, J. Wang, and W. Feng, “Combining latent factor model with location features for event-based group recommendation,” in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2013, pp. 910–918.
- [14] M. Moricz, Y. Dosbayev, and M. Berlyant, “PYMK: Friend recommendation at myspace,” in *Proc. ACM SIGMOD Int. Conf. Manag. Data*, Jun. 2010, pp. 999–1002.
- [15] Z. Zhang, Y. Liu, W. Ding, W. Huang, Q. Su, and P. Chen, “Proposing a new friend recommendation method, FRUTAI, to enhance social media providers performance,” *Decis. Support Syst.*, vol. 79, pp. 46–54, Nov. 2015.
- [16] R. Li, H. Zhu, L. Fan, and X. Song, “Hybrid deep framework for group recommendation,” *IEEE Access*, vol. 8, pp. 4775–4784, 2020.
- [17] S. Ahmad, I. Ullah, F. Mehmood, M. Fayaz, and D. Kim, “A stochastic approach towards travel route optimization and recommendation based on users constraints using Markov chain,” *IEEE Access*, vol. 7, pp. 90760–90776, 2019.
- [18] L. Xu, Y. Zheng, D. Xu, and L. Xu, “Predicting the preference for sad music: The role of gender, personality, and audio features,” *IEEE Access*, vol. 9, pp. 92952–92963, 2021.

- [19] L. Sama, H. Wang, and A. Makkar, "Movie recommendation system using deep learning," in *Proc. 9th Int. Conf. Orange Technol. (ICOT)*, Dec. 2021, pp. 1–4.
- [20] L. Zheng, Z. Tianlong, H. Huijian, and Z. Caiming, "Personalized tag recommendation based on convolution feature and weighted random walk," *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, pp. 24–35, 2020.
- [21] A. Darvishy, H. Ibrahim, F. Sidi, and A. Mustapha, "HYPNER: A hybrid approach for personalized news recommendation," *IEEE Access*, vol. 8, pp. 46877–46894, 2020.
- [22] Y. Liu, K. Xu, X. Chang, D. Di, and W. Huang, "Social network analysis based on canteen transaction records," *Statist. Interface*, vol. 11, no. 1, pp. 191–200, 2018.
- [23] H. Zhu, Y. Ni, F. Tian, P. Feng, Y. Chen, and Q. Zheng, "A group-oriented recommendation algorithm based on similarities of personal learning generative networks," *IEEE Access*, vol. 6, pp. 42729–42739, 2018.
- [24] Q. Zhu, S. Wang, B. Cheng, Q. Sun, F. Yang, and R. N. Chang, "Context-aware group recommendation for point-of-interests," *IEEE Access*, vol. 6, pp. 12129–12144, 2018.
- [25] O. Jiménez, M. Salamó, and L. Boratto, "Distance-aware event recommendation in event-based social networks," in *Artificial Intelligence Research and Development*. IOS Press, 2019, pp. 235–244.
- [26] Z. Wang, Y. Zhang, Y. Li, Q. Wang, and F. Xia, "Exploiting social influence for context-aware event recommendation in event-based social networks," in *Proc. IEEE Conf. Commun. (INFOCOM)*, May 2017, pp. 1–9.
- [27] M. Xu and S. Liu, "Semantic-enhanced and context-aware hybrid collaborative filtering for event recommendation in event-based social networks," *IEEE Access*, vol. 7, pp. 17493–17502, 2019.
- [28] Z. Zhang, D. D. Zeng, A. Abbasi, J. Peng, and X. Zheng, "A random walk model for item recommendation in social tagging systems," *ACM Trans. Manag. Inf. Syst.*, vol. 4, no. 2, pp. 1–24, Aug. 2013.
- [29] R. Mu, "A survey of recommender systems based on deep learning," *IEEE Access*, vol. 6, pp. 69009–69022, 2018.
- [30] E. Spertus, M. Sahami, and O. Buyukkokten, "Evaluating similarity measures: A large-scale study in the orkut social network," in *Proc. 11th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2005, pp. 678–684.
- [31] R. C. Bagher, H. Hassanpour, and H. Mashayekhi, "User trends modeling for a content-based recommender system," *Exp. Syst. Appl.*, vol. 87, pp. 209–219, Nov. 2017.
- [32] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King, "Recommender systems with social regularization," in *Proc. 4th ACM Int. Conf. Web Search Data Mining (WSDM)*, 2011, pp. 287–296.
- [33] D. Valcarce, A. Landin, J. Parapar, and Á. Barreiro, "Collaborative filtering embeddings for memory-based recommender systems," *Eng. Appl. Artif. Intell.*, vol. 85, pp. 347–356, Oct. 2019.
- [34] S. S. Peerzade, "Web service recommendation using PCC based collaborative filtering," in *Proc. Int. Conf. Energy, Commun., Data Analytics Soft Comput. (ICECDS)*, Aug. 2017, pp. 2920–2924.
- [35] R. Sharma, D. Gopalani, and Y. Meena, "Collaborative filtering-based recommender system: Approaches and research challenges," in *Proc. 3rd Int. Conf. Comput. Intell. Commun. Technol. (CICT)*, Feb. 2017, pp. 1–6.
- [36] P. Valdiviezo-Diaz, F. Ortega, E. Cobos, and R. Lara-Cabrera, "A collaborative filtering approach based on Naïve Bayes classifier," *IEEE Access*, vol. 7, pp. 108581–108592, 2019.
- [37] R. Chen, Q. Hua, Y.-S. Chang, B. Wei, L. Zhang, and X. Kong, "A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks," *IEEE Access*, vol. 6, pp. 64301–64320, 2018.
- [38] C. Tran, J.-Y. Kim, W.-Y. Shin, and S.-W. Kim, "Clustering-based collaborative filtering using an incentivized/penalized user model," *IEEE Access*, vol. 7, pp. 62115–62125, 2019.
- [39] G.-R. Xue, C. Lin, Q. Yang, W. Xi, H.-J. Zeng, Y. Yu, and Z. Chen, "Scalable collaborative filtering using cluster-based smoothing," in *Proc. 28th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2005, pp. 114–121.
- [40] T. Hofmann, "Latent semantic models for collaborative filtering," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 89–115, Jan. 2004.
- [41] Y. Huang and L. Bian, "A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the internet," *Exp. Syst. Appl.*, vol. 36, no. 1, pp. 933–943, Jan. 2009.
- [42] J. Liu, C. Wu, and W. Liu, "Bayesian probabilistic matrix factorization with social relations and item contents for recommendation," *Decision Support Syst.*, vol. 55, no. 3, pp. 838–850, 2013.
- [43] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proc. 26th Int. Conf. World Wide Web*, 2017, pp. 173–182.
- [44] Y. Ge, H. Xiong, A. Tuzhilin, and Q. Liu, "Cost-aware collaborative filtering for travel tour recommendations," *ACM Trans. Inf. Syst.*, vol. 32, no. 1, pp. 1–31, Jan. 2014.
- [45] Y. Jin, S. Dong, Y. Cai, and J. Hu, "RACRec: Review aware cross-domain recommendation for fully-cold-start user," *IEEE Access*, vol. 8, pp. 55032–55041, 2020.
- [46] B. Jiang, J. Yang, Y. Qin, T. Wang, M. Wang, and W. Pan, "A service recommendation algorithm based on knowledge graph and collaborative filtering," *IEEE Access*, vol. 9, pp. 50880–50892, 2021.
- [47] D. Bahl, V. Kain, A. Sharma, and M. Sharma, "A novel hybrid approach towards movie recommender systems," *J. Statist. Manag. Syst.*, vol. 23, no. 6, pp. 1049–1058, Aug. 2020.
- [48] M. Y. H. Al-Shamri and K. K. Bharadwaj, "Fuzzy-genetic approach to recommender systems based on a novel hybrid user model," *Exp. Syst. Appl.*, vol. 35, no. 3, pp. 1386–1399, Oct. 2008.
- [49] A. Bellogin, P. Castells, and I. Cantador, "Self-adjusting hybrid recommenders based on social network analysis," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf.*, New York, NY, USA, 2011, pp. 1147–1148, doi: 10.1145/2009916.2101092.
- [50] I. Soboroff and C. Nicholas, "Combining content and collaboration in text filtering," in *Proc. IJCAI*, 1999, pp. 86–91.
- [51] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, "Methods and metrics for cold-start recommendations," in *Proc. 25th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2002, pp. 253–260.
- [52] N. Sakib, R. B. Ahmad, M. Ahsan, M. A. Based, K. Haruna, J. Haider, and S. Gurusamy, "A hybrid personalized scientific paper recommendation approach integrating public contextual metadata," *IEEE Access*, vol. 9, pp. 83080–83091, 2021.
- [53] K. Pearson, "The problem of the random walk," *Nature*, vol. 72, p. 294, Jul. 1905.
- [54] F. Fouss, A. Pirotte, J.-M. Renders, and M. Saerens, "Random-walk computation of similarities between nodes of a graph with application to collaborative recommendation," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 3, pp. 355–369, Mar. 2007.
- [55] S. Tan and Q. Wu, "A random walk algorithm for automatic construction of domain-oriented sentiment lexicon," *Exp. Syst. Appl.*, vol. 38, no. 10, pp. 12094–12100, 2011.
- [56] H. Lu, X. Chen, J. Shi, J. Vaidya, V. Atluri, Y. Hong, and W. Huang, "Algorithms and applications to weighted rank-one binary matrix factorization," *ACM Trans. Manag. Inf. Syst.*, vol. 11, no. 2, pp. 1–33, Jun. 2020.
- [57] Y. Yang, S. Huang, W. Huang, and X. Chang, "Privacy-preserving cost-sensitive learning," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 5, pp. 2105–2116, May 2021.
- [58] L. Backstrom and J. Leskovec, "Supervised random walks: Predicting and recommending links in social networks," in *Proc. 4th ACM Int. Conf. Web Search Data Mining (WSDM)*, 2011, pp. 635–644.
- [59] M. Jamali and M. Ester, "TrustWalker: A random walk model for combining trust-based and item-based recommendation," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, 2009, pp. 397–406.
- [60] Z. Jiang, H. Liu, B. Fu, Z. Wu, and T. Zhang, "Recommendation in heterogeneous information networks based on generalized random walk model and Bayesian personalized ranking," in *Proc. 11th ACM Int. Conf. Web Search Data Mining*, Feb. 2018, pp. 288–296.
- [61] I. Konstas, V. Stathopoulos, and J. M. Jose, "On social networks and collaborative recommendation," in *Proc. 32nd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2009, pp. 195–202.
- [62] T. Bogers, "Movie recommendation using random walks over the contextual graph," *Tech. Rep.*, 2010.
- [63] Z. Yin, M. Gupta, T. Weninger, and J. Han, "A unified framework for link recommendation using random walks," in *Proc. Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2010, pp. 152–159.
- [64] G. Manju, P. Abhinaya, M. Hemalatha, and G. Manju, "Cold start problem alleviation in a research paper recommendation system using the random walk approach on a heterogeneous user-paper graph," *Int. J. Intell. Inf. Technol.*, vol. 16, no. 2, pp. 24–48, Apr. 2020.
- [65] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 5–53, Jan. 2004.

- [66] *How to Use Topics to Maximize Your Meetup Experience*. Accessed: Oct. Jan. 14, 2022. [Online]. Available: <https://www.meetup.com/blog/how-to-use-topics-to-maximize-your-meetup-experience/>
- [67] L. Katz, "A new status index derived from sociometric analysis," *Psychometrika*, vol. 18, no. 1, pp. 39–43, 1953.



YUE WU received the B.S. degree in computer science from the Harbin Institute of Technology, China, in 2012. She is currently pursuing the Ph.D. degree with Xi'an Jiaotong University, China. She has been involved in studying machine/deep learning in social media behavior prediction. Her major research interests include recommendation system information retrieval, social media behavior, and text mining.



WEI HUANG (WAYNE) was worked as a Tenured Professor/Visiting Scholar/a fellow in other top-tier research universities including the University of New South Wales, the University of Georgia, Ohio University and Harvard University. He is the Director of DeFin, NCAMS and Chair Professor with the College of Business, Southern University of Science and Technology (SusTech), Shenzhen, China. He has authored or coauthored more than 200 academic papers in quality international journals and conference proceedings. His research interests include digital economy, using IS/IT to support decision-making and collaboration and business analytics. He was elected as an AIS fellow in 2020. He has been the Founding Chair of AIS SIG-ISAP (IS in Asia-Pacific).



LEAN YU'D received the Ph.D. degree in management sciences and engineering from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences (CAS).

He is the Dean and the Ph.D. Supervisor at the School of Economics and Management, Beijing University of Chemical Technology. So far, he has published four monographs and over 60 SCI/SSCI articles in various journals. His research interests include business intelligence, economic forecasting, energy economics, and intelligent decision support systems. He was a winner of the National Science Fund for Distinguished Young Scholars and 100 Talents Program of Chinese Academy of Sciences. He is acted as the editor-in-chief and the managing editor or the editorial member of several international journals.



YINGJIE TIAND received the bachelor's degree in mathematics from Shandong Normal University, China, in 1994, the master's degree in applied mathematics from the Beijing Institute of Technology, Beijing, China, in 1997, and the Ph.D. degree in management science and engineering from China Agricultural University, Beijing.

He is currently a Professor with the Research Center on Fictitious Economy and Data Science, Chinese Academy of Sciences. He has published four books about data mining. His research interests include machine learning, optimization, and intelligent knowledge management.

...