Knowledge-Based Systems 228 (2021) 107248

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Predicting personalized grouping and consumption: A collaborative evolution model

Lu Yang [a], Yezheng Liu [a,b], Yuanchun Jiang [a,b], Le Wu [c,d,*], Jianshan Sun [a,b]

[a] School of Management, Hefei University of Technology, Hefei, Anhui 230009, China
[b] Key Laboratory of Process Optimization and Intelligent Decision Marketing, Ministry of Education, Hefei, Anhui 230009, China
[c] Intelligent Interconnected Systems Laboratory of Anhui Province, Hefei University of Technology, Hefei, Anhui 230601, China
[d] School of Computer Science and Information Engineering, Hefei University of Technology, Hefei, Anhui 230601, China

A R T I C L E I N F O

Article history:
Received 1 December 2020
Received in revised form 16 June 2021
Accepted 19 June 2021
Available online 29 June 2021

Keywords:
Consumption behavior
Grouping behavior
Group norms
Collaborative evolution and prediction
Temporal probabilistic matrix factorization

A B S T R A C T

With the prevalence of online social groups, the dynamic joint prediction of users' grouping and consumption behaviors on social network platforms is critical for optimizing social link suggestions and product recommendations. The group influence theory indicates that group norms affect user preference and behavior; however, the individual preferences of group members can also alter group norms. Nevertheless, the problems of how to holistically model dynamic bidirectional influence, existing between individual preferences and group norms, and then, simultaneously predict the users' grouping and consumption behaviors are still underexplored. In this study, we propose a collaborative evolution and prediction (CEP) model to address the above issues. We associate each social group with a latent group norm vector, and assign each user with a latent individual preference vector. The unobservable interplay between individual preferences and group norms is then modeled according to the underlying group influence theory. Based on these two latent vectors, we design a joint optimization function that incorporates the correlation between grouping and consumption behaviors, to enhance the prediction performance. Through extensive experiments and evolution analysis, we demonstrate the prediction effectiveness and the explanatory power of our CEP model.

1. Introduction

Online social networks (OSNs) have broken barriers that previously existed between users' social activity and consumption behavior, becoming the second most popular destination for users who wish to pay for products and services worldwide [1]. This new type of "social commerce", whereby e-commerce transactions are conducted directly through OSN platforms, has the potential to grow by more than $2 trillion by 2024 [2]. The prevalence of OSNs encourages people to join online social groups in increasing numbers, allowing users to freely join or quit groups based on personal preference. For example, over 10 million groups exist on Facebook, and 1.4 billion users interact with groups on the platform every month [3]. Online social groups generally cater to specific interests, and provide users with a virtual space to freely communicate and share ideas about their preferences. These types of social interactions can have powerful effects on the members' preferences and behaviors, through what is known as *group influence* [4,5]. Specifically, driven by individual preferences, a user voluntarily joins an online social group, which can be considered as *grouping behavior*. The grouping behavior would conversely impact individual preferences and consumption behavior owing to the group influence [6]. However, the interplay of individual preference and group influence is bidirectional, as the *group norm* [7] also evolves with changes in the majority of members' preferences. Generally, the group norm represents a consensual standard that describes group members' common preferences and guides their behaviors [8].

An example of this phenomenon is illustrated in Fig. 1. At time t, user *1* is in group *1* (with romantic tag) and has consumed products *1* and *3* (both are romance films). As user *1* has similar preferences to user *2* (both in group *1*) and user *3* (both have seen film *3*), she joins group *2* (with love and funny tags) which includes users *2* and *3*. Then, group *2* begins to influence user *1*'s preference through the norms present in that group (e.g., recommending some comedy films with love element to her, like product *2*). Thus, at time t+1, user *1*'s preference changes (i.e., she likes both romance and comedy films) and consumes product *2*. Meanwhile, as the current preferences of users *1* and *2* are consistent (i.e., both like romance and comedy films), group *1* modifies its norm (with new romantic and funny tags).

* Correspondence to: 1515 School of Management Building 1, Hefei University of Technology, Anhui 230009, China.
E-mail addresses: 2015111005@mail.hfut.edu.cn (L. Yang), liuyezheng@hfut.edu.cn (Y. Liu), ycjiang@hfut.edu.cn (Y. Jiang), lewu@hfut.edu.cn (L. Wu), sunjs9413@hfut.edu.cn (J. Sun).

https://doi.org/10.1016/j.knosys.2021.107248
0950-7051/

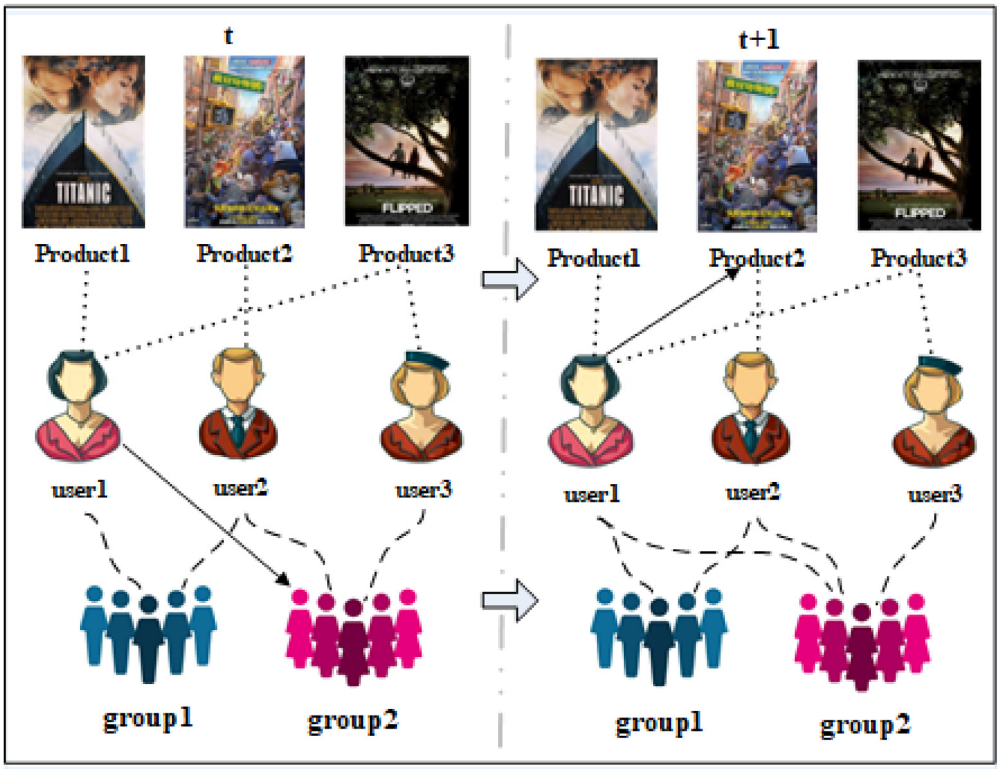


Fig. 1. Interaction of consumption and grouping behaviors.

In short, capturing abovementioned evolution on OSN platforms is significant. By understanding the dynamic influence between individual preferences and group norms, we could better perform customer relationship management and inform marketing strategies [6]. Additionally, generating accurate predictions of users' grouping and consumption behaviors is crucial for providing social link suggestion [9,10] and product recommendation [11].

In recent two decades, a significant amount of research has been devoted to better understanding social link prediction [9, 12–15] and consumption prediction [16–20]. For predicting social links, a common idea has been to compute node proximity given the topology of social networks [9,21], which is based on the principle that two users are candidates to be linked together if they are structurally close in the social graph [22]. Alternatively, for predicting product consumption, the most commonly used paradigm is collaborative filtering, wherein it is assumed that behaviorally similar users would exhibit similar preferences on products in the future [23]. However, the performance of both social link and consumption prediction models are often unsatisfactory owing to the data sparsity problem. With the increasing prevalence of social commerce and the easy access to users' social and consumption data from OSNs, extensive efforts have been devoted to incorporating one type of behavior data (social or consumption), to enhance the prediction of the other behavior [16,24–26]. Few studies have attempted to jointly model and predict users' social link and product consumption in a unified framework [27–29]. Previous works have instead tended to employ preference homophily and social relationships, to alleviate issues of data sparsity and improve prediction performance.

Despite the effectiveness of these works, we argue that these works suffer from two key limitations: failing to understand the importance of online social groups and combine both behaviors in a dynamic and dependent manner.

Previous works have mostly focused on a user's network of social friends, thereby ignoring another key component of OSNs: online social groups [25–27]. However, the dynamics of online social groups is important to the evolution of OSNs [22]. Besides, sociologist Van Vugt [30] argued that groups are ubiquitous and have powerful effects on the decision making and behavior of members.

Most previous works employed the group information to reinforce the model learning process, aiming at improving the consumption prediction [24,31]. Only limited works fused the grouping behavior but neglected the bidirectional between users and groups [10]. However, individual preferences and group norms are mutual influence, users' grouping and consumption behaviors are collaboratively evolutionary, as shown in Fig. 1.

In short, the limitations of existing studies lie in the fact that they model the users' preferences and grouping behavior from a static and independent perspective. As the grouping process and individual preference evolve over time, we aim to answer the question: how to borrow the theories related to group influence to better learn and predict the co-evolution of users' grouping and consumption behaviors?

To address these limitations, we propose a collaborative evolution and prediction (CEP) model. First, we construct a concept framework to present the evolution process: from the bidirectional influence between users and groups to users' grouping and consumption behaviors. Then, we need to realize the concept framework in an intuitive and rational model. Base on the temporal probabilistic matrix factorization algorithm, we design a latent factor based model to implement each modeling procedure of our concept framework. Specifically, we split the users' grouping and consumption behaviors into multiple evolutionary periods respectively. During each period, we associate each group with a vector of latent group norms and each user with a vector of latent individual preferences. Then, we predict the user–item interaction matrix (i.e., consumption behavior), using the latent individual preference vector revised by group influence, and the user-group interaction matrix (i.e., grouping behavior), using the latent individual preference and latent group norm vectors. As time progresses, because both latent factors are time-dependent, we can model the evolution of individual preferences and group norms by adjusting their latent factors, according to the group influence theory that governs these interactions. Finally, we designed a joint optimization function to combine the correlation between grouping behavior and consumption behavior for better overall prediction accuracy.

The contributions of this work are three-fold.

(1) We propose a CEP model to learn the collaborative evolution of the users' grouping and consumption behavior in OSNs, by modeling the dynamic bidirectional influence between individual preferences and group norms. This work not only explores the internal principle but also verifies the effectiveness, which is significant for social network analysis and real-world applications.

(2) We demonstrate the effectiveness of CEP as a social network analysis method, as the model performs well on four dimensions that is used to evaluate the capability of social network analysis method [22]. The first is information fusion, we jointly exploit two kinds of temporal behavior data, i.e., users' social and consumption behaviors. The second is pattern discovery, the CEP model realizes jointly prediction on two behaviors in dynamic OSNs, on the basis of inter-influence between users and groups. The third is scalability, when time information or one kind of behavior data is not available, the CEP model is easily adapted. The four is visualization, by analyzing explainable parameters of CEP model, the evolution characteristics of individual preferences and group norms can be presented.

(3) We validate the CEP model through a real-world dataset based on the DeviantArt[1] website, including 7358 users' grouping records amongst 1012 groups and their consumption records amongst 6188 artworks in one year. First, user-defined parameters of the proposed CEP model have been systemically investigated to obtain a reasonable combination of parameter settings. Based on the best combination of user-defined parameters, the experimental results show the effectiveness of the CEP model on dynamically predicting the users' grouping and consumption behaviors. At the same time, the experimental results on different recommended list lengths, latent feature numbers, and time periods show the robustness of the CEP model. Furthermore, based on three easily explainable and interpretable parameters of our CEP model, we can not only identify the susceptible users who are sensitive to the social influence, but also know the group evolution characteristics at different development stages.

The rest of this paper is organized as follows. Section 2 reviews the relevant social theories and modeling methods. Section 3 presents an overview of our conceptual framework and details the modeling process. Section 4 focuses on performing experiments and analyzing the results. Section 5 summarizes our study and proposes directions for future work.

2. Literature review

We first summarize social theories related to group influence and users' grouping and consumption behaviors. As these two types of behavior correspond to two research problems — social link prediction and consumption behavior prediction, we then review the literature on two issues. Finally, as we aim to present a joint prediction model, we review the literature on jointly modeling users' social and consumption behaviors.

2.1. Theoretical background

As the most "groupish" animals on the planet, humans have an innate affinity for groups [30]. In the context of OSNs, people's groupish proclivity strengthens because of the communicative convenience brought about by online social groups. Once a user joins a group, such affiliations can strongly and pervasively influence the user's decisions and behavior, through a process referred to as "group influence [4]". Many studies have provided insights into group influence on individual consumption behavior by analyzing the influence mechanism [6,32]. These studies typically argue that the most critical characteristic necessary to capture group influences on product consumption is the strength of group norms [33]. The group norm refers to unified standards that describe and prescribe which behaviors should and should not be performed in a given context [7]. With respect to consumption decisions, group norms are vital to determining product relevance, and can serve as "manuals of 'how to consume [34]'". Specifically, the group norm can revise the members' consumption preferences, and further suggest what products they should purchase [35]. Abstract group norms can be inferred from the members' overt behavior or from their expressed preferences [8].

However, one difficulty in assessing group norms is that group influence is dynamic. On the one hand, the group norm is defined and renegotiated over time, and conflicts can merge as individual members violate the norm [5]. For example, when absorbing new members with a different preference, the group generally, and automatically, updates its norms to satisfy the new members. On the other hand, several reports have verified that the strength of group norms perceived by one member might dynamically vary as the member spends more time within the group [6,36]. Thus, the dynamic group influence exerted by a group on a particular member at a particular time includes both the group norm and its strength, as perceived by the member at that time.

The above analysis demonstrates how a group dynamically influences its members' preferences and consumption behavior. Next, we explore the reverse, namely how changes in individual preferences affect the users' grouping behavior. Research on human sociality, particularly research that studies group formation, has already noted that people tend to form groups if they have similar preferences or characteristics [37]. The phenomenon – by which socially proximate individuals tend to share similar characteristics, including those that are unobserved or latent – is referred to as homophily [38]. The homophily effect plays an important role in understanding users' social behavior. Before a user enters a group, they perform self-categorization according to personal characteristics or individual preferences [39]. If the user's individual preference is similar to a group's norm, they tend to join the group. Apart from the homophily effect possibly influencing the customers' grouping behavior, sociologists have demonstrated that friends can change an individual social behavior [40]. Particularly, friends can serve as active connectors who introduce a user to their own social groups [41]. Furthermore, one's probability of joining a group depends on whether they have friends within the group [42]. The current study regards this motivation from friends within a group as the friend effect.

In OSNs, grouping behavior is a voluntary and conscious choice that is made according to the users' interests or preferences [43]. When users' individual preferences change with time, their corresponding grouping behavior also changes [35]. On the one hand, users would again find groups whose norm is similar to their new preference. One way is that users will again search for and find groups whose norms are similar to their new preference. However, the strength of the pressure from friends with different preferences can have differential effects. If a user's new preference and his/her friend's preference are similar, the friend

[1] https://www.deviantart.com/

effect becomes highly effective in modifying the user's grouping behavior.

Therefore, prior theoretical research has provided significant insights into the mutual influence between the users' grouping and consumption behaviors. Specifically, once a user joins a group, the group begins to influence this user's individual preference and consumption behavior through its group norms. The effects of group influence on users are dynamic, and depend on the strength of the group norm as perceived by its users. Conversely, individual preferences influence the users' grouping behavior by working on two motivations: namely, homophily and friend effects. Lastly, as group norms and individual preferences are dynamically changing, grouping and consumption behaviors are co-evolutionary.

2.2. Social link prediction

Social link prediction refers to attempts at predicting potential new social connections in the near future based on partially observed connections in a social network [21], which can assist people in finding new friends [44] and social groups [45]. As a key application of the social network analysis, traditional social link prediction approaches are based on similarity-based algorithms [22], such as node-based proximity or topology-based similarity, given a social network graph [21,46]. The similarity of two nodes can be inferred through their shared neighbors [46], path length [47] and the transition probability between them [48]. In recent years, various studies have utilized additional information to improve the performance of social link prediction models [9]. For example, Tang et al. [25] incorporated users' consumption records to predict whom users would trust in the future. Similarly, Beigi et al. [5] considered emotional impact when performing trust prediction. Additionally, owing to the fact that social networks are dynamic, researchers have leveraged temporal information in their social link prediction models, by analyzing the evolution patterns of social networks [12,15,49]. An intuitive and effective approach is to collapse multiple time-sliced linked data into a single matrix with weighted averaging, and then, apply static link prediction models [12]. Wu et al. [15] proposed an adaptively time series forecasting method to predict the future similarity between each node pair by using the Markov chain to represent the importance values of all nodes.

Unlike the above methods, multiple graph embedding models have emerged in recent years [50,51]. These models aim at learning the low-latent representations of nodes using the principle of preserving topological information. Most of them not only model the first-order connections but also consider the higher-order proximity. Study [52] proposed a novel knowledge graph embedding model to better learn the head and tail entity embeddings by defining each relation as a composition relation, including the entity translation and a diagonal projection matrix.

Nowadays, with the development of deep learning, a handful of deep learning-based models have been used to address the link prediction problem. For example, Wang et al. [53] utilized an autoencoder to preserve the topology information, which can be deemed as a deep neural network-based group embedding method. Later on, Liao et al. [14] leveraged the advantages of deep networks and social network embedding, which preserved both the structural proximity and attribute proximity in a social network.

2.3. Consumption behavior prediction

Users' consumption behavior prediction refers to attempts at recommending desired products to the users. Collaborative filtering has been widely used to produce personalized recommendation of items [54], wherein the latent factor model has exhibited considerable success [17,55]. Traditional collaborative filtering profiles an individual's preferences by identifying like-minded users with similar historical consumption behaviors. Specifically, it assumes that users with similar historical preferences are likely to have the same preferences in the future [17,23]. Among the various collaborative filtering techniques, matrix factorization [17] plays the central role, projecting users and items into a shared latent space, using the latent feature vector to represent users or items. Then, a user's interaction with an item is modeled as the inner products of their feature vectors. Although this method has experienced some success, when interaction records are sparse, the performance of matrix factorization suffers dramatically. Therefore, Mnih et al. [56] proposed the probabilistic matrix factorization (PMF) method, which models the predictive error of matrix factorization as a Gaussian distribution. The gradient descent algorithm is used to determine the local maxima of the posterior probability over the user and item latent vectors. The PMF method scales linearly with the number of transaction records, and more importantly, performs well on large, sparse datasets.

Users are likely to be influenced by their social relations when making consumption decisions [33]. Thus, researchers have attempted to incorporate additional information about these social relations, to mitigate problems of data sparsity and improve the performance of consumption prediction [16,26,27,57]. For example, Ma et al. [16] represented social relations as relationship matrix that was factorized with the user–item interaction matrix simultaneously. Additionally, friendship and group membership information [11] and social contextual information [57] were considered together to improve the consumption prediction accuracy. Besides, the random walk method also be used to combine the knowledge of a trust network among users [58].

With the popularity of deep learning in recent years, several deep learning-based methods have been developed to predict consumption behavior [59,60]. He et al. [19] built a neural network-based method of collaborative filtering (NCF), which passed the latent factor vectors of users and items in matrix factorization through the embedding layer of a neural network to improve modeling. Recently, the graph convolutional network (GCN) has shown a promising ability on consumption behavior prediction, by treating users and items as nodes on a user–item network graph [61]. Inspired by GCN approaches, Wang et al. [62] a neural graph collaborative filtering (NGCF) to capture the higher-order collaborative signals between users and items during the embedding learning process. Similarly, Wu et al. [26] used a GCN to extract user embeddings, containing transitive information from the social network for item recommendations. He et al. [63] found that it was necessary to use all components of the GCN when producing item recommendations, and proposed a LightGCN method, which only includes the neighborhood aggregation component for collaborative filtering and outperforms both NCF and NGCF.

However, none of the above methods can capture the dynamics of users' consumption behavior. Various researchers have investigated time-dependent methods, such as using a tensor factorization [64], session-based temporal graph model [18] to capture past temporal patterns. Nevertheless, these methods do not extrapolate future temporal dynamics, to estimate future changes in individual preferences. Recently, research has begun to model the evolution of individual preferences. For example, Zhang et al. [65] proposed a temporal matrix factorization (TMF) method, which could capture the temporal effect by incorporating single time-invariant biases. Additionally, owing to the success of recurrent neural networks on modeling sequence data, Wu et al. [66] proposed a recurrent recommender network (RRN) method to track the dynamic evolution of users and items, using a long short-term memory (LSTM) autoregressive model. Subsequently, researchers constructed a neural network-based tensor factorization (NTF) model for predictive tasks on dynamic relational data, which leveraged the LSTM to learn characterize the temporal evolution on relational data and incorporated the multi-layer perceptron structure to learn features of latent factors [67].

2.4. Joint modeling

To date, a few researchers have attempted to jointly model social link prediction and item recommendation, for example, the factor-based random walk model [27] and the neural-based joint prediction model [29] were devised to simultaneously predict friend connections and consumption preference. The key idea of these studies was to define a static user's latent factor, which was shared by the social link prediction and item recommendation. Unfortunately, this static setup fails to highlight the dynamic process of the users' friend links and item consumption. Wu et al. [28] proposed a dynamic model to realize friend link prediction and item recommendation from an evolutional perspective, but our research differs from their work in numerous ways. Firstly, the research problems are different. The focus of our research is the dynamic interaction between users and groups in OSNs, while Wu et al. [28] investigated the dynamic interplay amongst users. To the best of our knowledge, our work is the first to jointly model the evolutionary process of grouping behavior and consumption behavior in OSNs. Second, the modeling processes that were used in the studies are different. Our focus on the social network including groups is heterogeneous, compared with the users' social friend network. Additionally, the mechanism of group influence is more complicated than the social influence from friends [4].

3. CEP model

This section provides details about the construction of the CEP model. First, we define the collaborative evolution problem, and present the conceptual framework. Then, we introduce the modeling process to implement our conceptual framework. Finally, we present the parameter learning algorithm and prediction method for users' consumption and grouping behaviors.

3.1. Conceptual framework

Suppose that N users, M products and J groups exist in an OSNs, where users can join groups and consume products at any period $t, t = 1, \ldots, T$. We encode users' grouping and consumption behaviors in two three-dimensional tensors respectively, called grouping tensor $A \in \mathbb{R}^{N\times J\times T}$ and consumption tensor $C \in \mathbb{R}^{N\times M\times T}$, whose element is a binary indicator. In detail, $A^t_{ap} = 1$ if user a joins group p at period t and 0 otherwise. Similarly, $C^t_{ai} = 1$ if user a buys product i at time t, otherwise, it is 0. To avoid confusion, we use a, b, c and d to represent users; p, q to denote groups; i, j to signify products. Given the grouping tensor A and consumption tensor C from time *1* to T, our goals for this work are twofold:

(1) To model the evolution of user a's individual preference and group p's norm from period *1* to T

(2) To predict the probabilities of user a consuming product i and joining group q at period T+1

Fig. 2 illustrates the conceptual framework of our CEP model, including theoretical constructs and their corresponding notations. For ease of explanation, we list the main notations and the corresponding descriptions in Table 1.

In line with the theoretical analysis presented above, user a's grouping behavior at time $t-1$ (joined group p) triggers the influence of group p on user a's individual preference at time T (U^t_a) through group p's norm (G^{t-1}_p). The effectiveness of this influence is determined by user a's perceived strength on group p's norm at time T (I^t_{pa}).

Owing to the influence of group p, user a's individual preference evolves from time $t-1$ to time t (i.e., U^{t-1}_a changes into U^t_a). Meanwhile, the group norm that represents the common preference among members will change when the preferences of that group's members' change [7] (i.e., the solid arrow from U^t_a to G^t_p), such that group p's norm also evolves from time $t-1$ to time t (i.e., G^{t-1}_p changes into G^t_p).

Subsequently, driven by the current individual preference (U^t_a), user a would want to consume product i (C^t_{ai}) and join group q (A^t_{aq}). Particularly, user a's grouping behavior would be driven by two motivations: the homophily effect [38] ($D^t_{H\_aq}$) and friend effect [41] ($D^t_{F\_aq}$). On the one hand, user a wants to join group q at time t, owing to the fact that group q's norm is consistent with user a's individual preference at time t (large $D^t_{H\_aq}$ means high consistent). On the other hand, user a's friends who are in group q attract user a to join group q at time t (large $D^t_{F\_aq}$ indicates high attraction).

3.2. Modeling process

Following the conceptual framework of our CEP model in Fig. 2, this section initially models the collaborative evolutions of group's latent norm vector G^t_p and user's latent preference vector U^t_a. These two variables are then employed to predict the next-period consumption and grouping behaviors. To realize our CEP model, we use the temporal probabilistic matrix factorization algorithm. The key idea is at each time period (t), we map the latent vectors of users (U^t_a), groups (G^t_p) and products (V_i) to a same low-rank latent feature space, such that the consumption behavior (C^t_{ai}) and grouping behavior (A^t_{aq}) at that moment can be modeled as inner products in that space. The latent self-concept vector U^t_a is shared over these two kinds of behaviors.

3.2.1. Modeling evolution of group norms

The group norms inherently evolve and are explicitly unobservable, in which we use a latent vector $G^t_p \in \mathbb{R}^{D\times 1}$ to denote group p's norm at time t, a latent matrix $G^t \in \mathbb{R}^{J\times D}$ to indicate all groups' norms in period t, and a latent tensor $G = \left[G^1, \ldots, G^t, \ldots, G^T\right] (G \in \mathbb{R}^{T\times J\times D})$ to express all groups' norms at different periods. D is the number of latent features, J is the number of groups, and T is the amount of time periods. Similarly, we use a latent vector $U^t_a \in \mathbb{R}^{D\times 1}$ to denote user a's preference in period t, a latent matrix $U^t \in \mathbb{R}^{N\times D}$ to indicate all users' preferences in period t, and a latent tensor $U = \left[U^1, \ldots, U^t, \ldots, U^T\right] (U \in \mathbb{R}^{T\times N\times D})$ to express the individual preferences of all users at various time periods.

To capture the time-dependency between G^{t-1}_p and G^t_p, we model the evolutionary pattern of group norms as a combination of two parts, namely, the previous group norm G^{t-1}_p and the weighted preferences of all current group members $\sum_{a\in N^t_p} F^t_{ap} U^t_a$. F^t_{ap}, indicating the weight of member a contributing to p's norm at period t. N^t_p is the number of current members in group p at time t. In our work, we simply set $F^t_{ap} = \frac{1}{|N^t_p|}$, which means that group members have equal contribution on group p's norm at time t. This naïve influence weight is widely explored in many studies on social networks [68].

Therefore, we model group p's norms at period t as

$$p\left(G^t_p\right) = \mathcal{N}\left(G^t_p | \overline{G}^t_p, \sigma^2_G \mathrm{I}\right),$$

Table 1
Key constructs, notations and notation descriptions.

Constructs	Notations	Notation descriptions
Statistics	N, M, J, T	The number of users, products, groups, and time period, respectively
Individual preference	U_a^t	User a's latent preference vector at time period t
Group norm	G_p^t	Group p's latent norm vector at time period t
product property	V_i	Product i's latent feature vector
Perceived strength	I_{pa}^t	At time period t, user a's perceived strength of group p's norm
Group influence	E_{pa}^t	At time period t, user a's perceived influence from group p equals to the product of G_p^t and I_{pa}^t
Homophily effect	$D_{H\_aq}^t$	At time period t, the homophily degree between user a and group q
Friend effect	$D_{F\_aq}^t$	At time period t, user a's perceived influence from the friends in group q
Consumption behavior	C_{ai}^t	The probability of user a consuming product i at time period t, $C_{ai}^t \in \mathbb{R}^{N\times M\times T}$
Grouping behavior	A_{aq}^t	The probability of user a joining group q at time period t, $A_{aq}^t \in \mathbb{R}^{N\times J\times T}$

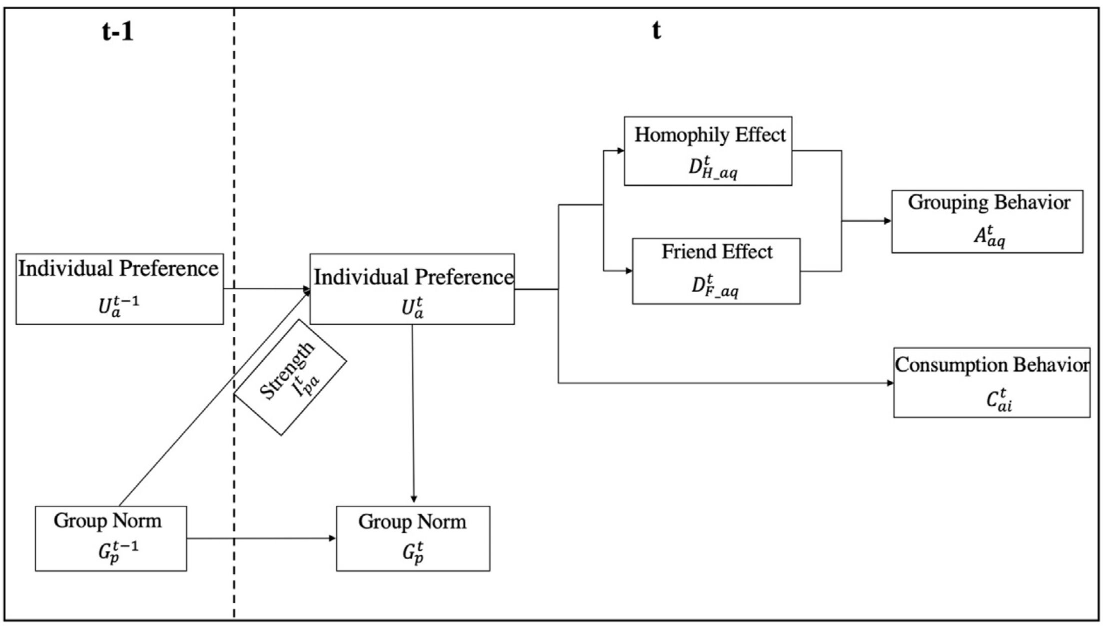


Fig. 2. Conceptual framework of CEP model.

$$\text{where } \overline{G}_p^t = \left(1-\gamma_p\right) G_p^{(t-1)} + \gamma_p \sum_{a\in N_p^t} F_{ap}^t U_a^t, \tag{1}$$

$s.t.\forall p \in G, \forall a \in U, 0 \le \gamma_p \le 1.$

In Eq. (1), we model the group norm G_p^t by a Gaussian distribution with mean $\overline{G}_p^t$ and variance $\sigma_G^2 I$. Variable γ_p is a group-specific parameter, which denotes the degree of the norms' evolution. A large γ_p indicates that group p is unstable and labile. This parameter is generally related to the group scale.

Consistent with the setting of previous studies [28,56], we place a zero-means Gaussian prior on the group norm at initial period:

$$p\left(G_p^1\right) = \mathcal{N}\left(G_p^1|0, \sigma_{G1}^2 \mathrm{I}\right). \tag{2}$$

Therefore, all groups' norms in various periods $G = \left[G^1, \ldots, G^t, \ldots, G^T\right]$ are modeled as follows:

$$p\left(G|\sigma_G^2, \sigma_{G1}^2\right) = \prod_{p=1}^{J} \mathcal{N}\left(G_p^1|0, \sigma_{G1}^2 \mathrm{I}\right) \prod_{t=2}^{T} \mathcal{N}\left(G_p^t|\overline{G}_p^t, \sigma_G^2 \mathrm{I}\right). \tag{3}$$

3.2.2. Modeling evolution of individual preference

Groups exert influence on users by altering their individual preferences [35]. Similar with the modeling pattern of group norms, we model a user's preference in period t as a function of two aspects: the user's preference at period $t-1$ and the perceived group influence at period t. First, we measure the group influence before modeling the evolution of individual preferences. As shown in Fig. 2, group influence E_{pa}^t can be measured as the multiplication of group p's norms at period $t-1$ (G_p^{t-1}) and user a's perceived strength of group p's norms at period t (I_{pa}^t):

$$E_{pa}^t = I_{pa}^t G_p^{t-1},$$
$$s.t.\ \ \forall a \in U, \forall p \in G. \tag{4}$$

In Eq. (4), perceived strength I_{pa}^t indicates that a group would exert varying influence effectiveness on its members, in which a user would perceive varying influences from different groups. Quantifying I_{pa}^t is generally a significant and challenging task because the detailed interaction processes and content among group members are rarely available in the context of online social groups. An operable alternative is measuring the similarity of the consumption behaviors between users and groups, that is computing the same consumed products at a period. This alternative method is consistent with the empirical literature on the strength of relationships [69,70].

The equation to calculate group p's influence strength on member a is as follows:

$$I_{pa}^t = \frac{2 \times R_{pa}^t}{N_p^t - 1}. \tag{5}$$

$$\text{where } R_{pa}^{t}=\sum_{c\in N_{p}^{(t-1)}\cap c\neq a}\frac{s_{a,c}^{(t-1)}}{s_{a}^{(t-1)}+s_{c}^{(t-1)}-s_{a,c}^{(t-1)}}. \tag{6}$$

where R_{pa}^{t} is the non-standardized influence strength perceived by user a from group p at time t and N_{p}^{t} is the number of members in group p at period t. $s_{a}^{(t-1)}$ ($s_{c}^{(t-1)}$) indicates the number of products purchased by user a (c) at period $t-1$, and $s_{a,c}^{(t-1)}$ is the number of products consumed by a and c at period $t-1$.

We then explicitly model user a's individual preference at period $t=2,3,\ldots,T$ by Eq. (7).

$$\begin{aligned} p\left(U_{a}^{t}\right)&=\mathcal{N}\left(U_{a}^{t}|\overline{U}_{a}^{t},\sigma_{U}^{2}\mathrm{I}\right),\\ \text{where } \overline{U}_{a}^{t}&=\left(1-\alpha_{a}\right)U_{a}^{(t-1)}+\alpha_{a}\sum_{p\in N_{a}^{(t-1)}}E_{pa}^{t}\\ &=\left(1-\alpha_{a}\right)U_{a}^{(t-1)}+\alpha_{a}\left(\sum_{p\in N_{a}^{(t-1)}}I_{pa}^{t}G_{p}^{(t-1)}\right), \end{aligned} \tag{7}$$

$s.t.\ \forall a\in U,\forall p\in G, 0\leq\alpha_{a}\leq 1.$

In Eq. (7), U_{a}^{t} is assumed to follow a Gaussian distribution with mean $\overline{U}_{a}^{t}$ and variance $\sigma_{U}^{2}\mathrm{I}$. $N_{a}^{(t-1)}$ is the set of affiliated groups of user a at period $t-1$. α_{a} is a user-specific susceptibility parameter, which indicates the extent to which the user a is subject to group influence in the individual preference. A large α_{a} means that user a in general values and accepts the suggestions of groups.

Similarly, a zero-mean Gaussian prior is designed to model individual preferences at the initial period, and all users' preferences in different periods are modeled as follows:

$$p\left(U|\sigma_{U}^{2},\sigma_{U1}^{2}\right)=\prod_{a=1}^{N}\mathcal{N}\left(U_{a}^{1}|0,\sigma_{U1}^{2}\mathrm{I}\right)\prod_{t=2}^{T}\mathcal{N}\left(U_{a}^{t}|\overline{U}_{a}^{t},\sigma_{U}^{2}\mathrm{I}\right). \tag{8}$$

3.2.3. Predicting grouping behavior

As shown in Fig. 2, the homophily effect $D_{F\_aq}^{t}$ and friend effect $D_{H\_aq}^{t}$ are two intentions of users' grouping behaviors. The literatures suggest that individuals establish links based on preferences for connecting with similar others [71]. Therefore, a user probably joins groups that include her close friends or homogeneously match his/her preference. According to the measurement of homophily in most collaborative filtering models, we use the product of latent preference vector U_{a}^{t} and latent norm vector $G_{q}^{(t-1)}$ to calculate the homophily between user a and group q at time t, i.e., $D_{H\_aq}^{t}$.

To calculate the friend effect $D_{F\_aq}^{t}$, we assume that users prefer to join groups that include her close friends [42]. We calculate the friend effect of user a by computing preference similarities between the user and her friends within group q, that is, $U_{a}^{t}\sum_{b\in N_{q}^{(t-1)}}^{N_{top}}U_{b}^{(t-1)}$. Variable b is the friend of a belonging to group q, $N_{q}^{(t-1)}$ is the member set of group q at time $t-1$, and N_{top} is the number of user a's closest friends in group q.

With the homophily effect $D_{H\_aq}^{t}$ and friend effect $D_{F\_aq}^{t}$, the probability of user a joining group q at time t can be modeled as follows:

$$\begin{aligned} \overline{A}_{aq}^{t}&=\left(1-\beta_{a}\right)D_{H\_aq}^{t}+\frac{\beta_{a}}{\left|N_{top}\right|}D_{F\_aq}^{t}\\ &=\left(1-\beta_{a}\right)U_{a}^{t}G_{q}^{(t-1)}+\frac{\beta_{a}}{\left|N_{top}\right|}U_{a}^{t}\sum_{b\in N_{q}^{(t-1)}}^{N_{top}}U_{b}^{(t-1)}, \end{aligned} \tag{9}$$

$s.t.\ \forall a\in U,\forall q\in G, 0\leq\beta_{a}\leq 1.$

where β_{a} is a user-specific balance parameter, which indicates the influence degree of homophily and friend effects on the grouping behavior. At the initial period, groups have no members, and the homophily effect is the only determinant factor for grouping behavior.

$$\overline{A}_{aq}^{1}=\left\langle U_{a}^{1},G_{q}^{1}\right\rangle. \tag{10}$$

Given the predicted grouping probability in Eq. (9), the likelihood of users' grouping behaviors could be expressed as follows:

$$p\left(A|U,G\right)=\prod_{t=1}^{T}\prod_{a=1}^{N}\prod_{q=1}^{J}\mathcal{N}\left[\left(A_{aq}^{t}|\overline{A}_{aq}^{t},\sigma_{A}^{2}\mathrm{I}\right)\right]^{Y_{aq}^{t}}. \tag{11}$$

where A_{aq}^{t} follows a Gaussian distribution with mean $\overline{A}_{aq}^{t}$ and variance $\sigma_{A}^{2}\mathrm{I}$, and Y_{aq}^{t} is an indicator tensor that equals 1 if user a joins group q at period t.

3.2.4. Predicting consumption behavior

With user's individual preference vector U_{a}^{t}, we can predict users' consumption behaviors. According to the collaborative filtering methods [17,28], the consumption probability of user a for product i at time t is modeled as follows:

$$p\left(C|U,V\right)=\prod_{t=1}^{T}\prod_{a=1}^{N}\prod_{i=1}^{M}\mathcal{N}\left[\left(C_{ai}^{t}|\left\langle U_{a}^{t},V_{i}\right\rangle,\sigma_{C}^{2}\mathrm{I}\right)\right]^{Y_{ai}^{t}}. \tag{12}$$

where Y_{ai}^{t} is an indicator tensor that equals 1 if user a purchases product i at period t, $V_{i}\in\mathbb{R}^{D\times 1}$ is the latent factor of the products in latent feature space $V\in\mathbb{R}^{M\times D}$, $\langle\rangle$ denotes the inner product of individual preference and product feature vectors. We add a zero-mean Gaussian prior to the product latent feature according to the traditional matrix factorization models [56]:

$$p\left(V|\sigma_{V}^{2}\right)=\prod_{i=1}^{M}\mathcal{N}\left[\left(V_{i}|0,\sigma_{V}^{2}\mathrm{I}\right)\right]. \tag{13}$$

3.3. Model learning and prediction

We summarize the graphical representation of the proposed model in Fig. 3, where the achromatic and chromatic variables indicate the observed and latent variables, respectively. Specifically, in Fig. 3, given the users' consumption and grouping behaviors from period $t=1$ to T, we can acquire the following parts from the CEP model: the latent product feature (i.e. orange variance V_{i}), the evolution pattern of user a's latent individual preference from period $t=1$ to T (i.e. U_{a}^{1} to U_{a}^{T} in blue sequence), the evolution pattern of group p's latent group norm from $t=1$ to T (i.e. G_{p}^{1} to G_{p}^{T} in green sequence) and the personalized parameters of the user and group (i.e. red variances α_{a}, β_{a}, and γ_{p}). On the basis of these settings, we can infer user a's individual preferences at period T+1 (U_{a}^{T+1}) and further predict user a's grouping and consumption behaviors. We calculate the grouping probability A_{aq}^{T+1} on the basis of the homophily effect ($D_{H\_aq}^{T+1}$ in the dotted vertical box) and friend effect ($D_{F\_aq}^{T+1}$ in the dotted horizontal box).

Given the sequence of consumption matrix C and grouping matrix sequence A, we aim to study the parameter set $\Phi=[U,G,V,\alpha,\beta,\gamma]$, where $\alpha=[\alpha_{a}]_{a=1}^{N}$, $\beta=[\beta_{a}]_{a=1}^{N}$, and $\gamma=\left[\gamma_{p}\right]_{p=1}^{J}$. Specifically, we have the posterior distribution over the parameters Φ as follows:

$$\begin{aligned} p\left(U,G,V,\alpha,\beta,\gamma|C,A\right)\propto\ & p\left(C|U,V\right)\times p\left(A|G,U,\beta\right)\\ &\times p\left(U|\alpha\right)\times p\left(G|\gamma\right)\times p\left(V\right). \end{aligned} \tag{14}$$

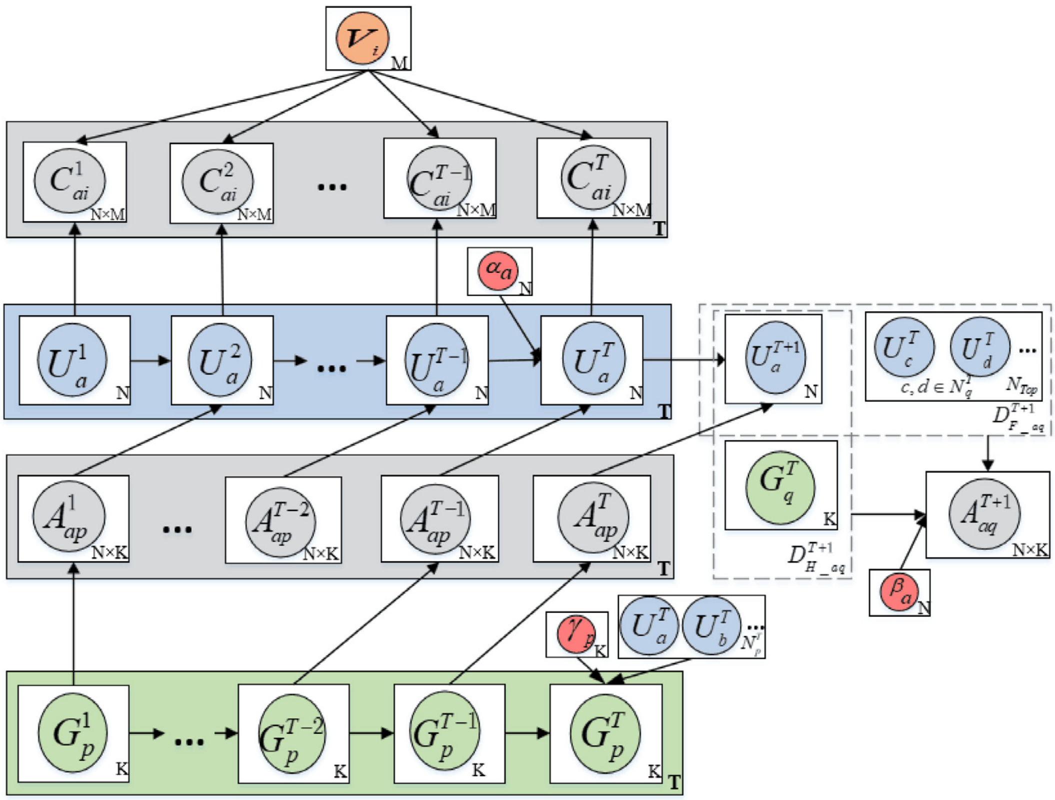


Fig. 3. Graphical representation of CEP model.

Maximizing the log posterior of the above equation is equivalent to minimizing the following pointwise loss:

$$\begin{aligned}
\min_{\Phi} \varepsilon(\Phi) &= \frac{1}{2}\sum_{t=1}^{T}\sum_{a=1}^{N}\sum_{i=1}^{M} Y_{ai}^{t}\left[C_{ai}^{t}-\overline{C}_{ai}^{t}\right]^{2} \\
&+ \frac{\lambda_A}{2}\sum_{t=1}^{T}\sum_{a=1}^{N}\sum_{q=1}^{J} Y_{ap}^{t}\left[A_{aq}^{t}-\overline{A}_{aq}^{t}\right]^{2} + \frac{\lambda_{G1}}{2}\sum_{p=1}^{J}\left\|G_p^1\right\|_F^2 \\
&+ \frac{\lambda_G}{2}\sum_{t=2}^{T}\sum_{p=1}^{J}\left\|G_p^t-\overline{G}_p^t\right\|_F^2 + \frac{\lambda_{U1}}{2}\sum_{a=1}^{N}\left\|U_a^1\right\|_F^2 \\
&+ \frac{\lambda_U}{2}\sum_{t=2}^{T}\sum_{a=1}^{N}\left\|U_a^t-\overline{U}_a^t\right\|_F^2 + \frac{\lambda_V}{2}\sum_{i=1}^{M}\left\|V_i\right\|_F^2,
\end{aligned}$$

$$s.t.\ \forall a \in U, \forall p, q \in G, \forall i \in V, \tag{15}$$

where $\lambda_A = \frac{\sigma_C^2}{\sigma_A^2}, \lambda_{G1} = \frac{\sigma_C^2}{\sigma_{G1}^2}, \lambda_G = \frac{\sigma_C^2}{\sigma_G^2}, \lambda_{U1} = \frac{\sigma_C^2}{\sigma_{U1}^2}, \lambda_U = \frac{\sigma_C^2}{\sigma_U^2}, \lambda_V = \frac{\sigma_C^2}{\sigma_V^2}$.

Variable λ_A is a trade-off coefficient between the consumption and the grouping prediction losses, and λ_U (λ_G) is a coefficient measuring how individual preferences (group norms) changes over time. λ_{U1} and λ_{G1} are regularization parameters for latent individual preference and latent group norm at initial period $t = 1$. λ_V is the regularization parameter of the latent product feature.

The coupling between U, G, V and the balance parameters (α, β, γ) does not make a convex objective function in Eq. (15). In practice, a local minimum could be achieved by iteratively performing gradient descent on each parameter. Specifically, the derivative of each parameter is as follows:

$$\begin{aligned}
\nabla_{U_a^t} &= \sum_{i=1}^{M} Y_{ai}^{t}(\overline{C}_{ai}^{t}-C_{ai}^{t})V_i + I\left[t=1\right]\lambda_{U1}U_a^1 \\
&+ I\left[t \ge 2\right]\lambda_U(\overline{U}_a^t - U_a^t) \\
&+ I\left[t < T\right]\lambda_U(1-\alpha_a)(\overline{U}_a^{(t+1)} - U_a^{(t+1)}) \\
&+ I\left[t \ge 2\right]\lambda_G \sum_{p \in N_a^t}\gamma_p F_{pa}^t(\overline{G}_p^t - G_p^t) \\
&+ I\left[t=1\right]\lambda_A \sum_{q=1}^{J} Y_{aq}^{1}(\overline{A}_{aq}^{1} - A_{aq}^{1})G_q^1 \\
&+ I\left[t \ge 2\right]\lambda_A(1-\beta_a)\sum_{q=1}^{J} Y_{aq}^{t}(\overline{A}_{aq}^{t} - A_{aq}^{t})G_q^{(t-1)} \\
&+ I\left[t \ge 2\right]\lambda_A \frac{\beta_a}{|N_{top}|}\sum_{b \in N_q^t}^{Ntop} Y_{aq}^{t}(\overline{A}_{aq}^{t} - A_{aq}^{t})U_b^{(t-1)}.
\end{aligned} \tag{16}$$

$$\begin{aligned}
\nabla_{G_p^t} &= I\left[t=1\right]\lambda_A \sum_{a=1}^{N} Y_{ap}^{1}(\overline{A}_{ap}^{1} - A_{ap}^{1})U_a^1 \\
&+ I\left[t < T\right]\lambda_A \sum_{a=1}^{N} Y_{ap}^{(t+1)}(1-\beta_a)(\overline{A}_{ap}^{(t+1)} - A_{ap}^{(t+1)})U_a^{(t+1)} \\
&+ I\left[t=1\right]\lambda_{G1}G_p^1 + I\left[t \ge 2\right]\lambda_G(\overline{G}_p^t - G_p^t) \\
&+ I\left[t < T\right]\lambda_G(1-\gamma_p)(\overline{G}_p^{(t+1)} - G_p^{(t+1)}) \\
&+ I\left[t < T\right]\lambda_U \sum_{a \in N_p^{(t+1)}} \alpha_a I_{ap}^{(t+1)}(\overline{U}_a^{(t+1)} - U_a^{(t+1)}).
\end{aligned} \tag{17}$$

$$\nabla_{V_i} = \sum_{t=1}^{T}\sum_{a=1}^{N} Y_{ai}^{t}(\overline{C}_{ai}^{t} - C_{ai}^{t})U_{ai}^{t} + \lambda_V V_i. \tag{18}$$

$$\nabla_{\alpha_a} = \lambda_U \sum_{t=2}^{T} (\overline{U}_a^t - U_a^t) \left(\sum_{p \in N_a^{(t-1)}} I_{pa}^t G_p^{(t-1)} - U_a^{(t-1)} \right). \tag{19}$$

$$\nabla_{\beta_a} = \lambda_A \sum_{t=2}^{T} \sum_{q=1}^{J} Y_{aq}^t (\overline{A}_{aq}^t - A_{aq}^t) \left(\frac{U_a^t}{N_{top}} \sum_{b \in N_p^{(t-1)}}^{N_{top}} U_b^{(t-1)} - U_a^t G_q^{(t-1)} \right). \tag{20}$$

$$\nabla_{\gamma_p} = \lambda_G \sum_{t=2}^{T} (\overline{G}_p^t - G_p^t) \left(\sum_{a \in N_p^t} F_{ap}^t U_a^t - G_p^{(t-1)} \right). \tag{21}$$

where I$[x]$ is an indicator function that equals 1 if x is true and equals 0 otherwise.

In the updating step, no constraints exist on U, G and V, thus we can directly update them using the stochastic gradient descent (SGD) approach [72]. With the bound constraints of α_a, β_a and γ_p, a local minimum can be found using the projected gradient descent (PGD) method [28]. In particular, for each $\alpha_a \in [0, 1]$, $\beta_a \in [0, 1]$, and $\gamma_p \in [0, 1]$, the PGD method updates the current solutions α_a^h, β_a^h, and γ_p^h in the hth iteration to $\alpha_a^{(h+1)}$, $\beta_a^{(h+1)}$ and $\gamma_p^{(h+1)}$ by the following rules:

$$\alpha_a^{(h+1)} = P\left[\alpha_a^h - \eta \nabla_{\alpha_a}\right], P(\alpha_a) = \begin{cases} \alpha_a & if \quad 0 \le \alpha_a \le 1, \\ 0 & if \quad \alpha_a < 0, \\ 1 & if \quad \alpha_a > 1. \end{cases} \tag{22}$$

$$\beta_a^{(h+1)} = P\left[\beta_a^h - \eta \nabla_{\beta_a}\right], P(\beta_a) = \begin{cases} \beta_a & if \quad 0 \le \beta_a \le 1, \\ 0 & if \quad \beta_a < 0, \\ 1 & if \quad \beta_a > 1. \end{cases} \tag{23}$$

$$\gamma_p^{(h+1)} = P\left[\gamma_p^h - \eta \nabla_{\gamma_p}\right], P(\gamma_p) = \begin{cases} \gamma_p & if \quad 0 \le \gamma_p \le 1, \\ 0 & if \quad \gamma_p < 0, \\ 1 & if \quad \gamma_p > 1. \end{cases} \tag{24}$$

After acquiring the related parameters $\Phi = [U, G, V, \alpha, \beta, \gamma]$, the two goals in the problem definition process can be answered: (1) the relative contribution of grouping behavior to the evolution process of each user's individual preference, and (2) the predicted consumption and grouping behaviors at T+1 given by:

$$U_a^{(T+1)} \approx (1 - \alpha_a) U_a^T + \alpha_a \sum_{p \in N_a^T} I_{pa}^{(T+1)} G_p^T. \tag{25}$$

$$\overline{A}_{aq}^{(T+1)} = (1 - \beta_a) U_a^{(T+1)} G_p^T + \frac{\beta_a}{|N_{top}|} U_a^{(T+1)} \sum_{b \in N_q^T}^{N_{top}} U_b^T. \tag{26}$$

$$\overline{C}_{ai}^{(T+1)} = U_a^{(T+1)} \times V_i \approx \left[(1 - \alpha_a) U_a^T + \alpha_a \sum_{p \in N_a^T} I_{pa}^{(T+1)} G_p^T. \right] \times V_i. \tag{27}$$

We obtain a probabilistic value for a user's grouping behavior on the basis of the homophily and friend effects by using Eq. (26). The predicted consumption probability of Eq. (27) captures the group influence on the user's preference when such user belongs to the groups.

Appendix A presents the pseudo code of our CEP model. The time complexity of the proposed CEP model lies in computing the latent representations of each user, group, and product and the balance parameters. We suppose that *c_pos* non-empty consumption records exist in consumption tensor C and that *a_pos* link records exist in grouping tensor A ($c\_pos \ll N \times M$, $a\_pos \ll N \times J$). Then, the average consumption and grouping records of a user at each period are $t_{c\_pos} = \frac{c\_pos}{N \times T}$ and $t_{a\_pos} = \frac{a\_pos}{N \times T}$, respectively. In each iteration, the time complexity is computed as $O\left(N \times T \times D \times \left(t_{c\_pos} + t_{a\_pos}\right)\right) = O\left(D \times (c\_pos + a\_pos)\right)$ for U, $O\left(D \times c\_pos\right)$ for V, $O\left(J \times T \times D \times \frac{a\_pos}{J \times T}\right) = O\left(D \times a\_pos\right)$ for G and $O\left(c\_pos + a\_pos\right)$ for the balance parameters. Therefore, the total complexity in each iteration is $O\left(D \times (c\_pos + a\_pos)\right)$, which is linear with the records at the period. Besides above theoretical analysis, we explore the computational complexity of our CEP model in Appendix B to further verify its practicability.

4. Experiments

This section implements the proposed CEP model on real-world data from DeviantArt. Firstly, Section 4.1 introduces the dataset and descriptive statistics. Next, Section 4.2 presents the evaluation protocols of our experiments. Especially, we use the orthogonal experimental design method to analysis the robustness of our CEP model and obtain the best combination of user-defined parameter settings. Then, Section 4.3 thoroughly evaluates the performance of the CEP model. Last, Section 4.4 analyzes the evolution of group norms and individual preferences.

4.1. Descriptive statistics

Our data are collected from the DeviantArt website, the world's largest online art community with over 60 million unique monthly visitors and 30 million registered users. In DeviantArt, art appreciators interact with arts by *favouriting* artworks or by *commenting* on artworks. The groups in DeviantArt are self-organized associations of users with consistent interests. Users are free to join any groups. DeviantArt records the time when users favor an artwork and join a group. Therefore, the data offer us opportunities to capture the interactions amongst users, groups and products over time.

Our experiment was conducted from May 1, 2017 to April 30, 2018. We adopt the universal but effective approach, which collapses the entire data into several discrete time intervals [10,73], to capture the time-varying nature of individual preferences and group norms. We divide the 12-month time span into 12 equal intervals. We use the data of the first 11 months for training and those of the last month for testing. We filter the grouping and consumption data to ensure each user is recorded to have joined at least two groups. Moreover, each group has at least two members, and every user has at least two artwork interaction records. Table 2 shows the statistics and sparsity of the DeviantArt dataset after data pre-processing.

Since the grouping and consumption data are limited to positive-only feedback, We adopt a uniform sampling method, which are used to solve the one-class problem of implicit feedback and have shown relatively high performance [19,28,29]. Someone may argue that we can use more advanced sampling strategies for sampling the negative samples. We agree that these non-uniform sampling techniques might further improve the performance of our CEP model. For example, the self-paced ensemble method [74], the adaptive ensemble method [75], and the weight-based method [76]. As the focus of this paper is the evolution design on users' two behaviors under any sampling techniques, we take the uniform random sampling as a solution, and leave how to design better evolution model with different sampling techniques as a future work.

To be specific, we randomly sample m times (we set the default sample ratio $m = 3$) missing data as observed pseudo negative records with a weight of $1/m$ at each iteration in the learning process, and we reselect the pseudo negative samples in each iteration. Considering that the sampling process is random, and the negative samples change each time, each missing record provides a considerably weak negative signal [28,29]. Appendix C elaborates the influence of different sampling ratios on prediction performances of grouping behavior and consumption behavior.

Table 2
Statistics of DeviantArt.

Element	Size
Users	7,358
Groups	1012
Artworks	6188
Time periods	12
Training consumption behavior	46,391
Training grouping behavior	55,577
Test consumption behavior	7,765
Test grouping behavior	6,930
Consumption behavior density	0.102%
Grouping behavior density	0.746%

4.2. Evaluation protocols

The aim of our CEP model is to jointly predict users' grouping and consumption behaviors, that is computing the probabilistic score $A_{aq}^{(T+1)}$ (or $C_{ai}^{(T+1)}$) for each candidate group (or item) and subsequently rank these scores and choose the largest ones (up to some threshold, e.g., top-K) as putative interacted groups (or items). Since considering all groups (or products) as candidates is too time-consuming, we followed the common strategy [27] that randomly samples 100 groups (or items) that are not interacted by the user until the time period of the test dataset, ranking the test groups (or items) among the 100 sampled groups (or items) and select the top-K groups (or items) as recommendation results.

4.2.1. Evaluation metrics

In fact, as the one-class nature of implicit feedback, the prediction task of our CEP model can be seen as a binary classification problem with imbalanced data [19]. In this scenario, the accuracy metric does not well reflect the model performance, thus precision and recall are commonly used to evaluate the performance on minority class [75]. Furthermore, as the harmonic mean of precision and recall, the F1-score also been widely used in classification task [76]. To judge the top-K ranking quality, these three metrics can be defined as follows:

(1) *Pre@K* also refers to as positive predictive value, indicating that of all the positive predicted conditions, a proportion of that is a true positive. By applying it to top-K recommendation, the metric is defined as:

$$Pre@K = \frac{n_{rel\_list}}{K},$$

where n_{rel_list} is the number of correctly predicted products (groups) in the ranking list. K is the length of the predicted recommendation list. High precision means that an algorithm performs well in placing truly consumed products (or truly joined groups) in a top-K recommendation list regardless of their rank.

(2) *Recall@K* indicates that of all positive conditions, the proportion of that is a true positive. In top-K recommendation, it is defined as:

$$Recall@K = \frac{N_{list\_test}}{N_{test}},$$

where N_{list_test} is the number of products (or groups) that appear not only in the predicted recommendation list but also in the testing dataset. N_{test} is the number of related products (groups) that a user really consumes (or joins) in the testing dataset. A recall takes a global view on all products (or groups) and a high recall indeed reflects users' adoption to the recommendation list.

(3) *F@K* is an integrated metric with consideration of the precision and recall of the recommendation results, defined as:

$$F@K = \frac{2 \times Prec@K \times Recall@K}{Prec@K + Recall@K}.$$

4.2.2. Baseline methods

We compare our proposed CEP model with the following four baseline methods, which can be classified into static and dynamic models.

(1) PMF [56]: This model predicts users' grouping or consumption behavior by using corresponding behavioral information from a static perspective.

(2) LightGCN [63]: This model is an advanced neural graph based recommendation model, which only consider the neighborhood aggregation component in GCN for collaborative filtering. The experimental results show that this model not only improve performance, but also is efficient.

(3) TMF [65]: This model is constructed on the basis of two assumptions: time is represented by a series of consecutive time periods, and a temporal dependence exists between two contiguous user latent factors (i.e., U_a^t and $U_a^{(t-1)}$). TMF captures temporal effects by incorporating single time-invariant biases.

(4) NTF [67]: This model firstly uses the LSTM architecture to learn dynamic interaction data, then uses the multi-layer perception structure to learn the latent feature vectors.

We also compare baselines with two variations derived from CEP, namely, evolving grouping prediction (EAP) and evolving consumption prediction (ECP), to further validate the effectiveness of collaboratively modeling users' two types of behavior with CEP. EAP predicts users' grouping behavior (i.e., removing the first term in Eq. [15]), whereas ECP predicts users' consumption behavior (i.e. $\lambda_A = 0$ in Eq. [15]).

4.2.3. Parameter settings of baselines

We tune all parameters of the methods to the best values for a fair comparison. The number of latent features is fixed to 15 for all models, other parameters used in all methods on the experiments are shown in Table 3.

4.2.4. Sensitivity analysis of user-defined parameters

In order to obtain the best performance for our CEP, its user-defined parameters are worth investigating. The six input parameters in CEP mode are as follows: λ_A, λ_U, λ_{U1}, λ_G, λ_{G1} and λ_V. Parameter λ_A is a trade-off parameter between grouping and consumption prediction losses. Parameter λ_U regularizes that the users' latent individual preferences change over time, similarly, λ_G regularizes that the latent group norm varies with time. λ_{U1} and λ_{G1} are the regularization parameters of the latent individual preference vector and latent group norm vector at $T = 1$, respectively. λ_V is the regularization parameter of the latent product factor. Besides, learning rate (lr) remarkably influences the prediction performance, which controls the training process and convergence speed. Hence, we need to determine the best setting of above seven user-defined parameters for CEP model to make sure its excellent performance and robustness.

A common method to obtain reasonable values is through the trial and error method, which is a time-consuming and expensive process. The orthogonal experimental design (OED) or Taguchi method [77], provides a mathematical tool called the orthogonal arrays which allows the analysis of the relationships between a large number of user-defined parameters within the smallest number of possible experiments [78]. For our CEP model, there are seven parameters to analyze, considering the whole combinations among parameters is a burdensome work with an inflated cost. Therefore, the OED method is used to acquire the best combination of our CEP model. The number of levels for each parameter is set as seven. A full-factorial analysis needs $7^7 = 823{,}543$ experiments. By contrast, if we use the OED method, an orthogonal array $L_{49}(7^7)$ that contains only 49 experiments is enough to obtain reasonable values for CEP model. Appendix D elaborates the OED process and result analyses. Seven parameters of CEP model are optimized by orthogonal designs and range analysis, and the best combination of parameters is obtained, summarized in Table 4.

Table 3
Parameters of baselines.

Model	Training parameter	Best value for prediction
PMF	Learning rate Regularization parameter	0.001 0.01
LightGCN	Number of layers Learning rate Regularization parameter Batch size	1 0.001 0.01 64
TMF	Learning rate Regularization parameter Regularization parameter of transition matrix	0.001 0.01 0.01
NTF	Number of hidden layers Learning rate Batch size BN scale parameter BN shift parameter	3 0.001 16 0.99 0.001

Table 4
The best combination of parameters for CEP model.

Model	Training parameter	Best value for prediction
CEP	Learning rate (lr) Trade-off parameter between prediction losses (λ_A) Regularization parameters of user evolution (λ_U) Regularization parameters of group evolution (λ_G) Regularization parameters of users' initial status (λ_{U1}) Regularization parameters of groups' initial status (λ_{G1}) Regularization parameter of product feature (λ_V)	0.01 1 3 0.1 0.01 0.01 0.01

4.3. Performance comparison

In this section, we would sequentially discuss the performances of all models on two behavior predictions when the number of latent features (D) varies, when the time periods (T) changes, and when the length of the recommendation lists (K) varies.

4.3.1. Overall performance

We first compare predictive performances of the CEP model against other methods under the best setting. Figs. 4 and 5 show that under the setting of $T = 11$, the prediction results of the consumption and grouping behaviors produced by each model with respect to three evaluation metrics when the number of latent features varies from 5 to 20 (i.e., $D = [5, 10, 15, 20]$). We only show the predictive results when for the top five (i.e., $K = 5$) recommended projects due to the space constraints. The sensitivity analysis of K is specified in the Section 4.3.3.

In Figs. 4 and 5, we can observe several results. Firstly, the proposed CEP model shows the best prediction performance on grouping and consumption behaviors amongst three metrics, outperforming state-of-the-art methods, such as NTF and Light GCN. Secondly, most models considering time information (i.e., CEP, EAP, ECP, NTF) perform better than the static baselines without temporal consideration (i.e., PMF and LightGCN). Such a result proves that time information plays an important role in the prediction of users' behaviors. It is worth noting that the LigthGCN outperforms TMF for grouping behavior prediction (when $D = 15$), showing the advantage of graph-based neighborhood aggregation on learning latent feature factors. Thirdly, the CEP model has a significantly better predictive power in grouping behaviors than all baselines, followed by EAP. This result indicates that leveraging the users' consumption history is effective in grouping prediction. Fig. 5 shows similar results for consumption behavior prediction, that is, CEP outperforms all baselines, followed by ECP. Such an outcome indicates the effectiveness of incorporating the grouping information in the individual preference prediction. Fourthly, the predicted results of CEP are better than EAP and ECP on two predictive tasks, which indicates the benefit of collaborative predictions. This result is obtained because users' grouping and consumption behaviors are correlative. The performances of all models significantly improve from $D = 5$ to 15 with the increase in the number of latent features but decrease when D further increases (except for the *Pre*@5 of CEP model in Fig. 5). The reason for this phenomenon is that additional local optima will be generated as D excessively increases, leading to poor performance [27]. Therefore, we set $D = 15$ in the following experiments.

4.3.2. Rolling prediction

We further measure the rolling prediction performances using a moving window approach to verify the robustness of CEP model. Specifically, the preceding time periods of time T (i.e., $T =$ 1 to T-1) are used as a training set. Time T (i.e., $T = 8, 9, 10, 11$) is used as the test set. We only compare the predictive results generated by all models on the comprehensive evaluation metric F@5 due to the space constraints, shown in Table 5.

The boldface and italic highlight the best and second-best performers, respectively. With regard to the four testing windows of grouping behavior prediction in Table 5(a), CEP always performs best, followed by EAP. Similar results can be found in Table 5(b). A conspicuous result shows that, when $T = 10$, the performance of three methods (i.e., PMF, LightGCN, and TMF) do not further improve and even decrease. This situation happens because users or groups may significantly shift their preferences or norms at time period $T = 10$ and consequently limit these models' performances. On the one hand, the static models (PMF and LightGCN) cannot model the dynamic changes of users and groups. On the other hand, the dynamic model TMF cannot accurately capture the striking temporal changes by a single time-invariant biases. Instead, our CEP model achieves the best performance by collaboratively modeling the evolution of user preference and group norms.

4.3.3. Impact of the length of recommendation list (K)

This subsection discusses the effect of the length of recommendation list (K) on the CEP model's performance.

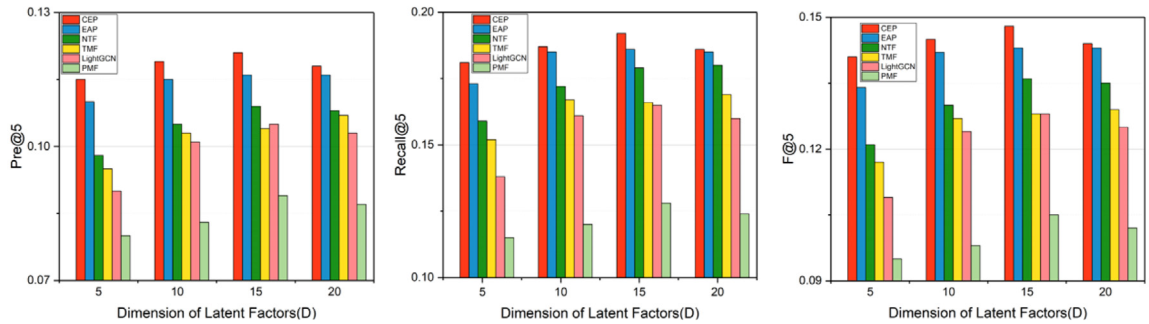


Fig. 4. Prediction of grouping behavior.

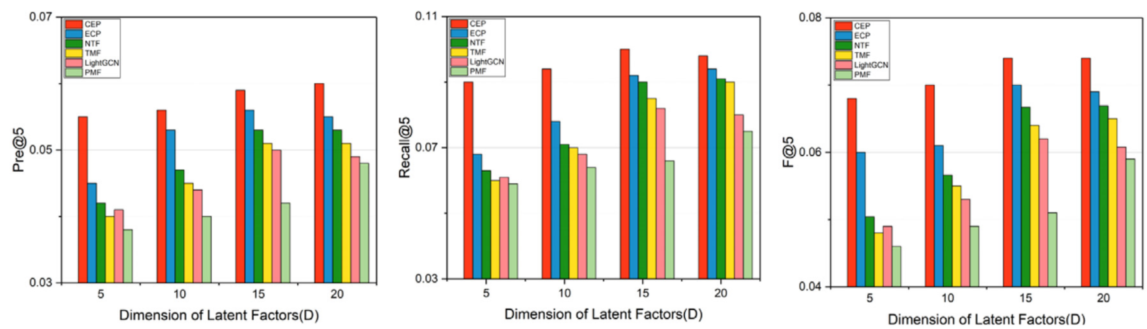


Fig. 5. Prediction of consumption behavior.

Table 5
Prediction performances of two behaviors on T ($D = 15$).

(a) *F@5* of grouping behavior

Model	$T = 8$	$T = 9$	$T = 10$	$T = 11$
PMF	0.104	0.104	0.102	0.105
LightGCN	0.126	0.127	0.126	0.128
TMF	0.126	0.126	0.125	0.128
NTF	0.133	0.134	0.134	0.135
EAP	*0.139*	*0.140*	*0.141*	*0.143*
CEP	**0.142**	**0.143**	**0.145**	**0.148**

(b) *F@5* of consumption behavior

Model	$T = 8$	$T = 9$	$T = 10$	$T = 11$
PMF	0.047	0.051	0.048	0.051
LightGCN	0.059	0.060	0.060	0.062
TMF	0.061	0.063	0.063	0.064
NTF	0.066	*0.067*	*0.067*	0.067
ECP	*0.065*	*0.067*	*0.067*	*0.070*
CEP	**0.068**	**0.070**	**0.071**	**0.074**

(a) shows the *F@5* of grouping behavior, (b) shows the *F@5* of consumption behavior. The best performer is in **boldface** and the second best performer is in *italic*.

Fig. 6 draws the performance of the CEP model on grouping and consumption behaviors, with the list length K ranging from 5 to 20. The CEP model outperforms all baseline methods for all K, followed by two variants (EAP and ECP). The CEP model achieves a more significant improvement in consumption prediction, compared with the grouping prediction. The grouping data are substantially denser than the consumption data (0.746% vs. 0.102% in Table 1). Thus, our CEP model can utilize dense grouping data to reduce the sparsity issue of consumption prediction and produce better results.

4.4. Evolution analysis

In the proposed CEP model, we define two latent representations (i.e. U_a^t and G_p^t) and three personalized parameters (i.e., α_a, β_a, and γ_p) for users and groups. With these variables, we can recognize the evolution patterns of group norms and individual preferences. In this section, we perform the evolution analysis on group norms and individual preferences in combination with the three parameters.

4.4.1. Evolution results for group norms

As mentioned in Section 3.2.1, the evolution of a group's norm from $t-1$ to t (G_p^t) can be regarded as the integration of historical status (G_p^{t-1}) and members' latest preferences ($\sum_{a\in N_p^t} F_{ap}^t U_a^t$). Such integration is balanced by group-specific parameter γ_p. Groups with small γ_p means that group norms slowly evolve. In this subsection, we first show the statistical characteristics of group norm evolution, and then explore evolution rules of group norms.

Fig. 7 shows that average change rates of all the group norms at contiguous time periods. The red curve (denotes mean values) is smoother than the blue curve (denotes std values). Specifically, all the mean values of change rates over 11 periods are lower than 40% but higher than 20%. Specifically, from an average perspective, the groups' norms evolve with a moderate speed. However, the std curve of change rates is swing, especially the std values at $T = 3, 8$ and 11. This phenomenon is because of the group heterogeneity, which some groups' norms sharply change, while other groups slightly vary. In Fig. 7, although all of the groups dynamically change with time, their change rates are different, thus their evolutions are not synchronous.

The left (right) y-axis denotes the mean (std) of change rate over all groups' norms, and the x-axis indicates the time periods.

Group-specific parameter γ_p reflects the evolution stability of group norms. Fig. 8 reports the statistics of γ_p values for all the groups and find that most groups' norms slowly evolve and develop with a stable pattern (the mean is 0.27 and the standard deviation is 0.01) because more than 80% of γ_p values are ranged in [0.1, 0.4).

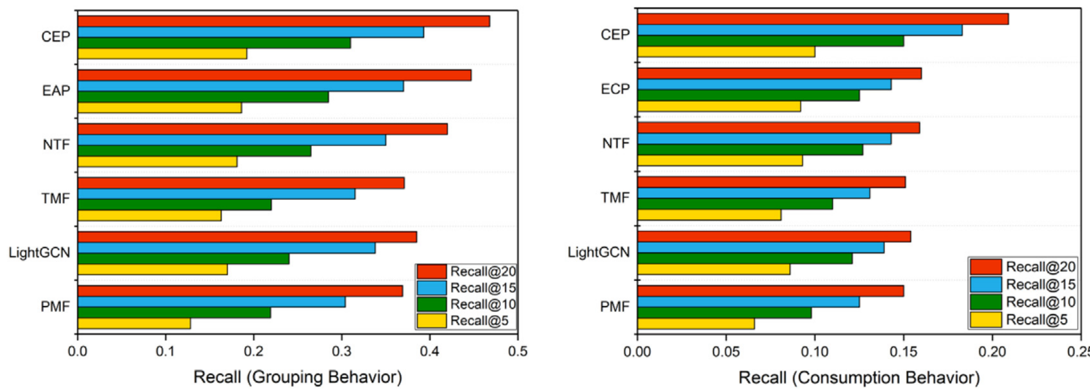


Fig. 6. Recall among top-K.

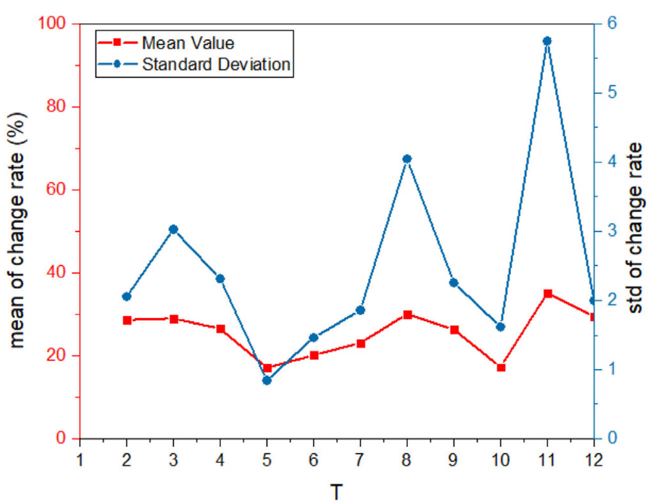


Fig. 7. The statistic of change rate over group norm.

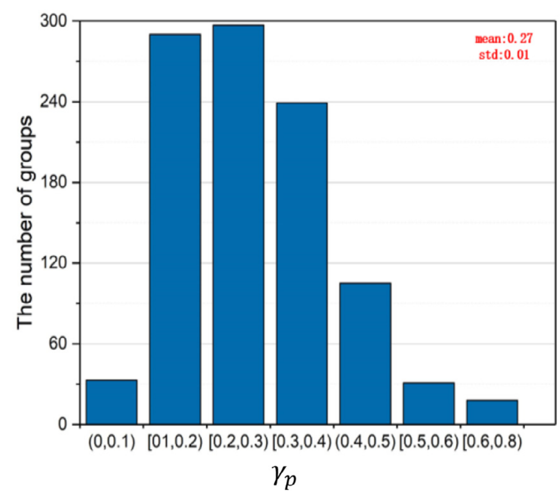


Fig. 8. The statistic of γ_p.

Groups constantly develop in their life time, called the group evolution. Dr. Bruce W. Tuckman [79] indicates that group development includes five stages: forming, storming, norming, performing and adjourning. In online social groups, the *growing scale* (the number of a group added members in one year) is a mirror of group developing characteristics on different stages with the dynamic joining process. Here, we analyze the relation between *growing scale* and group-specific parameter γ_p to show the existing dependency between group development and group evolution and summarize the evolution rules of group norms, shown as in Fig. 9.

The distribution of γ_p approximates an inverted *U*-shape with the increase in the *growing scale*, except for when the *growing scale* is extremely small. We would explain the relation between γ_p and *growing scale* according to group developmental stages [79]. At the forming stage (with the *growing scale* less than 15), most members discover the group boundaries by supporting the initial group norms, thus their groups slightly change their norms (the majority of γ_p values are lower than 0.3). Conversely, few groups apparently change norms (with γ_p between 0.4 and 0.6) because some members favor a more active approach to discover the group boundaries, such as subvert current norms. At the storming stage (the *growing scale* rises from 15 to 100), intragroup conflict emerges and increases — the more member, the more intense (the values of γ_p rise from 0.2 to 0.6 until reach the peak with the increase in the *growing scale*). At the norming stage (the *growing scale* increases from 100 to 170), members use cohesion to solve the conflict and establish standardized group norms to highlight the group cohesion (the values of γ_p decrease from 0.6 to 0.2 when the *growing scale* increases). At the performing stage (the *growing scale* further rises from 170 to 250), groups evolve as an effective and inclusive organization whose norms accept differences and keep up to date (most values of γ_p stay at the range [0.2, 0.4]).

Two groups with large *growing scale* (approximately 150) and γ_p (near 0.8) are tagged with red circles in Fig. 9. In comparison with other groups' conflict during the storming stage, the two groups face an intense situation because their *growing scale* is particularly large. This notion means that many people join a group in an extraordinarily short time aggravates friction and arguments. Thus, initial norm collapse occurs.

This correlation between growing scale and group-specific parameter γ_p reveals the evolution rules of group norms at different development stages.

4.4.2. Evolution results for individual preferences

This subsection attempts to study the evolving pattern of individual preferences. As discussed in Section 3.2.2, individual preferences at period t (U_a^t) not only rely on historical preferences (U_a^{t-1}) but also are revised by group influence (E_{pa}^t), with a user-specific parameter α_a for balance. A large α_a denotes that the influence from group p on user a's preference (E_{pa}^t) is great.

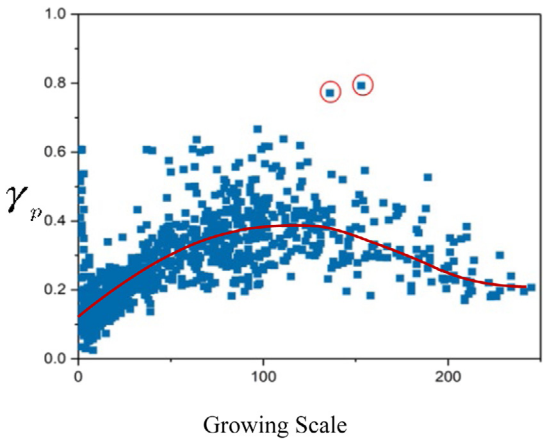


Fig. 9. The relation between growing scale and γ_p.

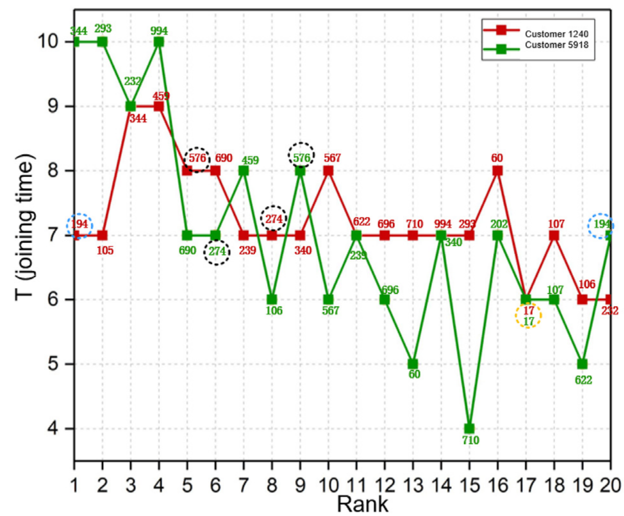


Fig. 10. The influence of 20 groups on two users at $T = 12$ period.

Given that another user-specific parameter β_a could reflect a user's preferences on grouping behaviors (preference homophily or friend effect), we would explore individual preferences and its evolution characteristics by analyzing E^t_{pa}, α_a, and β_a.

Firstly, we explicate the dynamic group influence E^t_{pa} on the individual preference, including the influences from the same group on different users and the influences of different groups on the same user. To this end, we first select two typical users (user ids 1240 and 5918) who joined 20 common groups from $T = 1$ to $T = 11$. Then, we show the influences from the 20 groups on two users at $T = 12$. In Fig. 10, the bottom shows the 20 groups ranked by their influence effects from left to right on the x-axis, the y-axis reports the time periods of grouping behaviors, and the figures above the nodes denote the group ids.

In Fig. 10, the influences from a same group on different members can be significantly different even at the same time. For example, group 194 (marked with a blue circle) is ranked first for user 1240, whereas it is positioned the last for user 5918, although both of them joined group 194 at $T = 7$. This phenomenon is because of the markedly different perceived influence strengths by the two users. By contrast, group 17 (tagged with a yellow circle) exerts the same influence on the two users. Apart from the above two extremes, a group generally has a similar influence on members who joined the group at the same time (e.g. groups 576 and 274 marked with black circles).

The influences from different groups on the same user are different. From an overall perspective, the ranking of group influence is high when the grouping behavior is later (a large T), which is consistent with many studies on social influence [36]. Specifically, once familiar with the group and its norms, the person receives sufficient information. Thus, the value of group norms and its influence on individual preferences decreases.

Next, we analyze two user-specific parameters α_a and β_a to further reveal the evolution characteristics of individual preferences. α_a captures the effectiveness of group influence on user a' preference evolution, while β_a reflects which effect (preference or friend) determines a user's grouping behavior.

Figs. 11 and 12 report the statistical information of α_a and β_a, respectively. In Fig. 11, we obtain that most users are influenced by their affiliated groups at varying degrees (a mean of 0.41 and a standard deviation of 0.03). Fig. 12 shows that about 84% of β_a concentrated in the range (0.4, 0.6), a mean of 0.49 and a standard deviation of 0.01. The distribution of β_a is more concentrated than those of α_a and γ_p. This situation means that most users' grouping behaviors are simultaneously driven by the homophily and friend effects.

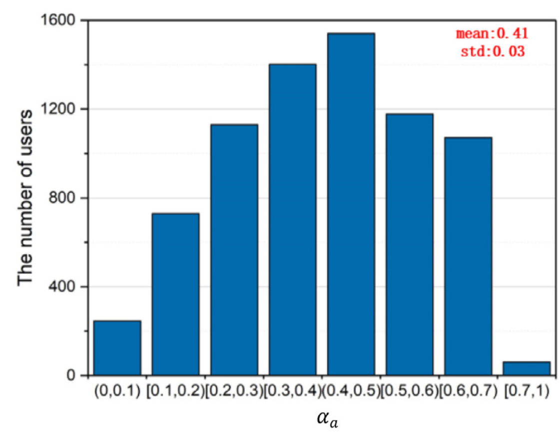


Fig. 11. The Statistics of α_a.

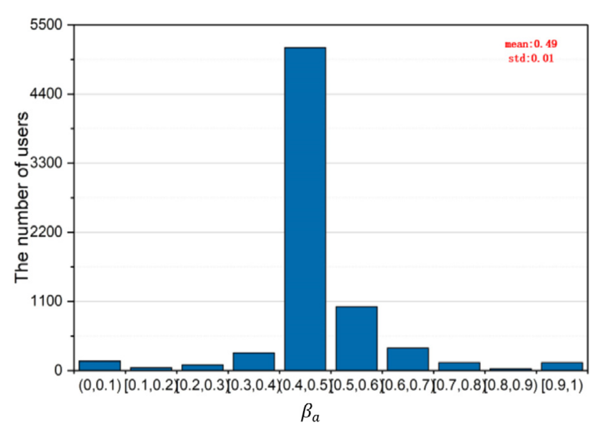


Fig. 12. The Statistic of β_a.

This section discusses the effects of parameters on the performance of the CEP model in terms of the user-specific parameter β_a.

To verify the reasonability of the setting of β_a, which is used to balance the homophily effect ($D^t_{H\_aq}$) and friend effect ($D^t_{F\_aq}$) in predicting grouping behavior, we separately analyze the prediction accuracy under friend and homophily effects (setting $\beta_a = 0$ and $\beta_a = 1$, respectively), as shown in Fig. 13. The 'adaptive

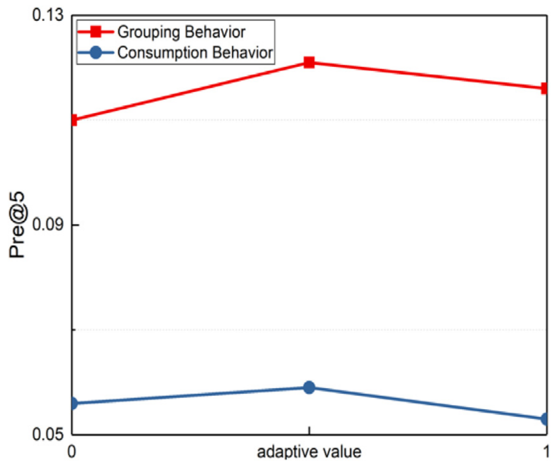


Fig. 13. Impact of β_a.

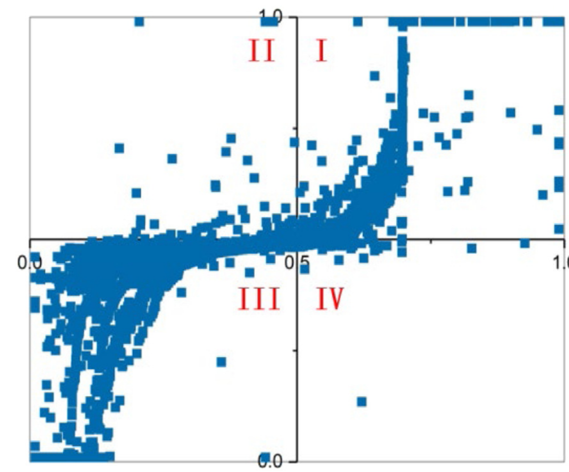


Fig. 14. The Relation between α_a and β_a.

value' in the x-axis represents β_a learned from the CEP model. When β_a is set to 0 or 1, the *Pre*@5 values on two behavior predictions are lower than the results under 'adaptive value'. Thus, the homophily and friend effects are crucial for users' grouping behaviors [80]. Besides, the consumption prediction under the homophily effect ($\beta_a = 0$ is better than the result of friend effect ($\beta_a = 1$), while the grouping prediction shows an opposite result. Ignoring either the homophily effect or the friend effect leads to the overestimation of the other effect [80].

Fig. 14 draws the correlation between α_a and β_a. Parameter α_a shows a significant positive correlation with β_a. This situation is because α_a and β_a reflect a user's susceptibility degree on social relationship (group membership and friendship). A user with large α_a and β_a (the first quadrant of Fig. 14) is a susceptible person who is easily influenced by the two relationships. By contrast, the third quadrant represents those people who are not impressionable. Besides, some users lie in the second and fourth quadrants. The two reasons for this result are as follows: these users are only sensitive to one social relationship and the random error.

Overall, individual preferences evolve with time under the dynamic group influence, and individual grouping behavior also changes with groups under the friend effect. The effectiveness of group influence and friend effect is determined by the user's susceptibility degree to social relationship. We can realize the user classification and discover the susceptible users (the first quadrant) by distinguishing the values of α_a and β_a.

5. Conclusion and discussion

This study provides a CEP model to jointly predict users' grouping and consumption behaviors in the dynamic OSNs. Driven by the importance of online social groups and the limitation of current studies on modeling the collaborative evolution between users and groups, we use a temporal probabilistic matrix factorization algorithm to construct our CEP model. First, this algorithm can project users' two kinds of temporal implicit feedback into a shared latent space. Then, at each time period, grounded in the bidirectional influence mechanism between users and groups, the collaboratively modeling can be easily realized by directly operating latent vectors. Next, by modeling the time-dependency between different time periods, the collaborative evolution of users' grouping and consumption behaviors can be learned. After that, we designed a joint optimization function to combine the correlation between these two behaviors for better prediction.

We conduct comprehensive experiment on an actual dataset to prove the effectiveness of our CEP model. Both prediction results on two behaviors show that the CEP model outperforms other state-of-the-art methods, such as LightGCN and NTF. It is worth noting that the CEP model shows more excellent results on consumption prediction when data is sparser. Meanwhile, the experimental results on different recommended list lengths, latent feature numbers, and time periods show the robustness of the CEP model. Moreover, the findings discovered by evolution analysis show the explanatory power of our CEP model. For example, by focusing on two user-specific parameters, we can identify the susceptible users who are sensitive to the social influence (including group influence and friend effect). The inverted U-distribution of group-specific parameter and group growing scale reveals group evolution characteristics at different development stages. The above findings provide empirical evidence for the effectiveness and rationality of our CEP model.

However, there are some limitations of CEP model. First, we only use the pointwise loss as our optimization function and adopt a uniform under-sampling procedure to learn the implicit feedback. In fact, both loss functions and sampling strategies are powerful on prediction performance. Thus, exploring the influence of different loss optimizations and non-uniform sampling strategies on our CEP model is one exploration in the future. Second, the excellent performance of LightGCN method have verified that the simplified graph convolution operation outperforms matrix-based methods on modeling interaction data, which inspires us to explore how to migrate the CEP model to a graph-based deep learning framework in the future work.

CRediT authorship contribution statement

Lu Yang: Investigation, Software, Validation, Formal analysis, Writing - original draft, Writing - review & editing. **Yezheng Liu:** Funding acquisition, Validation, Writing - review & editing. **Yuanchun Jiang:** Formal analysis, Writing - review & editing. **Le Wu:** Conceptualization, Investigation, Validation, Formal analysis, Writing - review & editing. **Jianshan Sun:** Validation, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work is supported by the Major Program of the National Natural Science Foundation of China (91846201), the Fundamental Research Funds for the Central Universities, China (JZ2020HGPA0114), the National Natural Science Foundation of China (71722010, 71872060), the Haier Rainforest Project and Zhejiang Lab Project, China (2019KE0AB04), the China Postdoctoral Science Foundation (2021M690852).

Appendix A. The algorithm of CEP model

Algorithm 1: Parameter Learning of the CEP model
Input: The consumption sequence C and the affiliation sequence A
Output: The predicted grouping behaviors and consumption behaviors at time T+1

Initialize U, G, V, α, β and γ
While the loss function of Eq. (15) does not converge **do**
  for user $a = 1$; $a \leq N$; a++ **do**
    for time $t = 1$; $t \leq T$; t++ **do**
      Fix $G, V, \alpha, \beta, \gamma$ update U_a^t using SGD
    end for
    Fix G, U, V, β, γ update α_a using PGD
  end for
  for group $p = 1$; $p \leq J$; p++ **do**
    for time $t = 1$; $t \leq T$; t++ **do**
      Fix $U, V, \alpha, \beta, \gamma$ update G_p^t using SGD
    end for
    Fix G, U, α, β update γ_p using PGD
  end for
  for product $i = 1$; $i \leq M$; i++ **do**
    Fix $U, G, \alpha, \beta, \gamma$ update V_i using SGD
  end for
end while
Predict users' preferences at $T + 1$ according to Eq. (25)
for all $(a, q, T + 1)$ records in affiliation test dataset **do**
  Predict users' grouping behavior at $T + 1$ based on Eq. (26)
end for
for all $(a, i, T + 1)$ records in consumption test dataset **do**
  Predict users' consumption behavior at T+1 based on Eq. (27)
end for

Appendix B. Actual runtime comparison

In Section 3.3, we analyze the runtime complexity of the CEP model. Here, we compare all models' actual runtimes and conduct experiments on a Core8Duo 3.6 GHz machine with Windows 10 and 16 GB of memory. Table A.1 reports the actual runtime of a single iteration of the training phase for every model. Four baselines are independent prediction models, that is, they need to be conducted twice to predict the grouping and consumption behaviors. We only report the runtime of four baselines on the consumption prediction as they show the same computing performance on the grouping behavior.

The PMF costs the least time as its static modeling. Although the LightGCN is also a static model, the neighborhood aggregation of every node on graph is time-consuming. Compared with PMF, the dynamic models (TMF and NTF) require additional runtime due to the addition of time information. Specifically, we compare the best baselines of static and dynamic and find that the NTF could improve 1.05-time (for grouping prediction) and 1.08-time (for consumption prediction) on the *F@5* metric with half the time compared with LightGCN. Furthermore, we find that CEP could improve 1.09-time (for grouping prediction) and 1.11-time (for consumption prediction) than NTF on the *F@5* metric at the 2.12-time cost on each iteration.[2] This analysis demonstrates that the time cost is in a reasonable range despite the predictive benefit of our CEP model at a cost of runtime.

The CEP model costs the most runtime because of its dynamical and collaborative modeling process. The CEP needs to compute each user's latent preference vector and each group's latent norm vector at each time slice. In the modeling of each latent individual preference vector, the CEP must calculate the influence strength function with consumption data, which adds to the time cost. In real-world applications, we can train the CEP offline and store the predictive results of consumption and grouping behaviors based on the output of CEP in the server. In the online stage, users can obtain real-time predictions by retrieving the predictions from the server, which is time efficient and can be easily applied to real-world social platforms and e-commerce sites.

Table A.1
The runtime of all models (min.).

Time	PMF	LightGCN	TMF	NTF	CEP
Each iteration	0.12	0.62	0.25	0.33	1.4

Table A.2
The impact of negative sampling ratio m.

(a) Grouping behavior				(b) Consumption behavior			
	Pre@5	Recall@5	F@5		Pre@5	Recall@5	F@5
$m = 1$	0.118	0.186	0.144	$m = 1$	0.053	0.085	0.065
$m = 2$	0.120	0.187	0.146	$m = 2$	0.055	0.093	0.069
$m = 3$	0.121	0.192	0.148	$m = 3$	0.056	0.094	0.070
$m = 4$	0.121	0.191	0.147	$m = 4$	0.059	0.100	0.074
$m = 5$	0.119	0.187	0.145	$m = 5$	0.060	0.098	0.074
$m = 6$	0.116	0.181	0.142	$m = 6$	0.059	0.098	0.073

Table A.3
Factors and levels of OED.

Levels	Factors						
	A λ_A	B λ_U	C λ_G	D λ_{U1}	E λ_{G1}	F λ_V	G lr
1	0.01	0.01	0.001	0.001	0.001	0.001	0.0001
2	0.1	0.1	0.005	0.005	0.005	0.005	0.001
3	0.5	0.5	0.01	0.01	0.01	0.01	0.005
4	1	1	0.1	0.1	0.1	0.1	0.01
5	3	3	3	0.5	0.5	0.5	0.05
6	5	5	5	1	1	1	0.1
7	7	7	7	5	5	5	0.25

Appendix C. Sampling ratio analysis

The advantage of the flexible negative sampling has been shown in many works [19,29]. Here, we flexibly control the sampling ratio m to explore its influence on prediction performance. We empirically fix $m = 3$ on grouping part/consumption part while we debug the consumption part/ grouping part. Then, we show the performance of CEP model with different negative sampling ratio (range from 1 to 6) on two behavioral predictions in Table A.2.

[2] **Note that**, NTF needs to be conducted twice for predicting grouping and consumption behaviors respectively, while the CEP can simultaneously predict the grouping and consumption behaviors by conducting once. Thus, the final runtime of NTF needs to multiply by 2, i.e., its each runtime is 0.66 min.

Table A.4
Parameters sensitivity results of CEP model based on $L_{49}(7^7)$ orthogonal array.

No.	λ_A	λ_U	λ_G	λ_{U1}	λ_{G1}	λ_V	lr	Grouping behavior			Consumption behavior		
								Pre@5	Recall@5	F@5	Pre@5	Recall@5	F@5
1	1	1	1	1	1	1	1	0.109	0.167	0.131	0.040	0.068	0.050
2	1	2	3	4	5	6	7	0.066	0.105	0.081	0.035	0.053	0.042
3	1	3	5	7	2	4	6	0.105	0.173	0.130	0.049	0.083	0.061
4	1	4	7	3	6	2	5	0.109	0.174	0.134	0.052	0.085	0.064
5	1	5	2	6	3	7	4	0.112	0.176	0.137	0.053	0.084	0.065
6	1	6	4	2	7	5	3	0.111	0.173	0.135	0.052	0.084	0.064
7	1	7	6	5	4	3	2	0.109	0.176	0.134	0.052	0.084	0.064
8	2	1	7	6	5	4	3	0.087	0.152	0.110	0.040	0.059	0.048
9	2	2	2	2	2	2	2	0.110	0.168	0.132	0.050	0.084	0.062
10	2	3	4	5	6	7	1	0.110	0.188	0.138	0.055	0.088	0.068
11	2	4	6	1	3	5	7	0.098	0.137	0.114	0.041	0.056	0.047
12	2	5	1	4	7	3	6	0.115	0.182	0.140	0.056	0.092	0.069
13	2	6	3	7	4	1	5	0.109	0.186	0.140	0.056	0.088	0.068
14	2	7	5	3	1	6	4	0.112	0.187	0.140	0.054	0.093	0.068
15	3	1	6	4	2	7	5	0.110	0.177	0.135	0.049	0.081	0.061
16	3	2	1	7	6	5	4	0.109	0.167	0.132	0.049	0.081	0.061
17	3	3	3	3	3	3	3	0.119	0.189	0.146	0.058	0.090	0.070
18	3	4	5	6	7	1	2	0.112	0.178	0.137	0.053	0.086	0.066
19	3	5	7	2	4	6	1	0.112	0.182	0.139	0.051	0.087	0.064
20	3	6	2	5	1	4	7	0.096	0.135	0.112	0.043	0.060	0.050
21	3	7	4	1	5	2	6	0.115	0.177	0.139	0.056	0.091	0.069
22	4	1	5	2	6	3	7	0.087	0.131	0.105	0.026	0.049	0.034
23	4	2	7	5	3	1	6	0.112	0.178	0.137	0.053	0.089	0.067
24	4	3	2	1	7	6	5	0.112	0.177	0.137	0.052	0.077	0.062
25	**4**	**4**	**4**	**4**	**4**	**4**	**4**	**0.120**	**0.192**	**0.147**	**0.058**	**0.097**	**0.073**
26	4	5	6	7	1	2	3	0.112	0.184	0.139	0.058	0.087	0.069
27	4	6	1	3	5	7	2	0.115	0.181	0.141	0.057	0.091	0.070
28	4	7	3	6	2	5	1	0.117	0.187	0.143	0.056	0.090	0.069
29	5	1	4	7	3	6	2	0.109	0.175	0.134	0.049	0.081	0.061
30	5	2	6	3	7	4	1	0.111	0.174	0.135	0.053	0.082	0.064
31	5	3	1	6	4	2	7	0.100	0.137	0.115	0.038	0.057	0.045
32	5	4	3	2	1	7	6	0.114	0.178	0.138	0.055	0.089	0.068
33	5	5	5	5	5	5	5	0.114	0.171	0.137	0.055	0.087	0.067
34	5	6	7	1	2	3	4	0.113	0.178	0.138	0.057	0.082	0.067
35	5	7	2	4	6	1	3	0.113	0.180	0.139	0.055	0.089	0.068
36	6	1	3	5	7	2	4	0.112	0.170	0.135	0.051	0.077	0.061
37	6	2	5	1	4	7	3	0.112	0.176	0.136	0.051	0.075	0.061
38	6	3	7	4	1	5	2	0.107	0.174	0.133	0.051	0.082	0.063
39	6	4	2	7	5	3	1	0.114	0.184	0.14	0.056	0.091	0.069
40	6	5	4	3	2	1	7	0.099	0.140	0.116	0.045	0.056	0.050
41	6	6	6	6	6	6	6	0.094	0.121	0.106	0.046	0.059	0.051
42	6	7	1	2	3	4	5	0.110	0.178	0.136	0.052	0.081	0.063
43	7	1	2	3	4	5	6	0.105	0.171	0.130	0.050	0.081	0.061
44	7	2	4	6	1	3	5	0.106	0.172	0.131	0.051	0.082	0.063
45	7	3	6	2	5	1	4	0.108	0.173	0.132	0.050	0.083	0.062
46	7	4	1	5	2	6	3	0.110	0.176	0.135	0.051	0.087	0.064
47	7	5	3	1	6	4	2	0.113	0.178	0.138	0.057	0.080	0.067
48	7	6	5	4	3	2	1	0.112	0.176	0.136	0.054	0.085	0.066
49	7	7	7	7	7	7	7	0.071	0.097	0.081	0.033	0.043	0.034

With the increase of sampling ratio, the performance of grouping behavior prediction first increases and then decreases when m>3. It shows that too large sampling ratio may adversely hurt the performance [19]. The optimal sampling ratio is around 3 to 5, and the performance of the consumption behavior part has the same tendency with the social part.

Appendix D. Orthogonal experimental design

The OED method or Taguchi method is a variable reduction technique which can improve the algorithm performance at a minimum cost [77,78]. A key process to the OED method is the selection of orthogonal arrays. For our CEP model, we consider the following user-defined parameters: λ_A, λ_U, λ_{U1}, λ_G, λ_{G1}, λ_V, and learning rate (lr). The number of levels for each factor (i.e., each parameter) is set seven, summarized in Table A.3.

For seven levels of seven factors, the orthogonal array $L_{49}(7^7)$ that contains only 49 experiments is adopted in our OED. Table A.4 summarizes the experimental results within 25 runs of our CEP model for grouping and consumption behavior predictions, as each experiment would converge within 25 iterations. From Table A.4, we know that each factor has seven levels, and each level is repeated seven times.

By analyzing the Table A.4, we could obtain a reasonable (may not the best) combination of parameter settings, corresponding the experiment of No. 25. It indicates that when each factor is set as the fourth level (i.e., $\lambda_A = \lambda_U = 1$, $\lambda_G = \lambda_{U1} = \lambda_{G1} = \lambda_V = 0.1$, $lr = 0.01$), the CEP model shows the best performance on two behavior prediction tasks among 49 experiments. However, compared with the full-factorial analysis that needs $7^7 = 823{,}543$ experiments, the orthogonal array $L_{49}(7^7)$ that only contains 49 experiments may not contains the best combination of parameter settings. Generally, the orthogonal experiment is followed by a range analysis, which is helpful to determine the best combination of parameter settings. Therefore, we further conduct the range analysis based on the experimental results of $L_{49}(7^7)$ orthogonal array, as shown in Tables A.5 and A.6. As F@5 is an integrated metric that balances precision and recall of the recommendation results, we report statistical results on the F@5 metric when seven factors at different levels. Specifically, Table A.5 shows the F@5 results of grouping behavior prediction,

Table A.5
Range analysis on grouping behavior based on $L_{49}(7^7)$ orthogonal array.

Grouping behavior								
Measurements	Levels	λ_A	λ_U	λ_G	λ_{U1}	λ_{G1}	λ_V	lr
K value	1	0.882	0.880	0.930	0.933	0.924	0.932	0.962
	2	0.914	0.884	0.927	0.917	0.929	0.930	0.949
	3	0.940	0.931	0.921	**0.942**	**0.940**	**0.934**	0.940
	4	**0.949**	0.945	**0.940**	0.911	**0.940**	0.908	**0.961**
	5	0.936	**0.946**	0.921	0.928	0.880	0.924	0.950
	6	0.902	0.908	0.895	0.879	0.892	0.872	0.920
	7	0.883	0.912	0.872	0.896	0.900	0.906	0.724
K-avg value	1	0.126	0.126	0.133	0.133	0.132	0.133	0.137
	2	0.131	0.126	0.132	0.131	0.133	0.133	0.136
	3	0.134	0.133	0.132	**0.135**	**0.134**	**0.133**	0.134
	4	**0.136**	0.135	**0.134**	0.130	0.134	0.130	**0.137**
	5	0.134	**0.135**	0.132	0.133	0.126	0.132	0.136
	6	0.129	0.130	0.128	0.126	0.127	0.125	0.131
	7	0.126	0.130	0.125	0.128	0.129	0.129	0.103
Best level		4	5	4	3	3/4	3	4
R		0.010	0.009	0.009	0.009	0.008	0.008	0.034

Table A.6
Range analysis on consumption behavior based on $L_{49}(7^7)$ orthogonal array.

Consumption behavior								
Measurements	Levels	λ_A	λ_U	λ_G	λ_{U1}	λ_{G1}	λ_V	lr
K value	1	0.410	0.376	0.422	0.423	0.431	0.431	0.450
	2	0.430	0.420	0.437	0.417	0.434	**0.436**	0.453
	3	0.441	0.431	0.445	**0.447**	**0.439**	**0.436**	0.444
	4	**0.444**	**0.451**	**0.448**	0.442	0.436	0.426	**0.457**
	5	0.440	**0.451**	0.423	0.441	0.427	0.432	0.448
	6	0.418	0.436	0.418	0.407	0.413	0.412	0.446
	7	0.417	0.435	0.407	0.423	0.420	0.427	0.302
K-avg value	1	0.059	0.054	0.060	0.060	0.062	0.062	0.064
	2	0.061	0.060	0.062	0.060	0.062	0.062	0.065
	3	0.063	0.062	0.064	**0.064**	**0.063**	**0.062**	0.063
	4	**0.063**	0.064	**0.064**	0.063	0.062	0.061	**0.065**
	5	0.063	**0.064**	0.060	0.063	0.061	0.062	0.064
	6	0.060	0.062	0.060	0.058	0.059	0.059	0.064
	7	0.060	0.062	0.058	0.060	0.060	0.061	0.043
Best level		4	4/5	4	3	3	2/3	4
R		0.004	0.010	0.006	0.006	0.004	0.003	0.022

where the K value refers to the sum of F@5 values of grouping behavior on the corresponding level when each factor is set as different levels. The K-avg value in Table A.5 equals the corresponding K value divided by the number of levels (i.e., 7). The R value indicates the range of K-avg value. The boldface indicates the best performance and highlights the best level.

According to the K value and K-avg value in Table A.5, we obtain the best level of each factor on grouping behavior prediction. Though referring to Table A.3, we know the best combination of CEP model on grouping behavior prediction is as follows: $\lambda_A = 1$, $\lambda_U = 3$, $\lambda_G = 0.1$, $\lambda_{U1} = 0.01$, $\lambda_{G1} = 0.01/0.1$, $\lambda_V = 0.01$, $lr = 0.01$. Similarly, though Table A.6, we also acquire the best level of seven factors on consumption behavior prediction, that is, $\lambda_A = 1$, $\lambda_U = 1/3$, $\lambda_G = 0.1$, $\lambda_{U1} = 0.01$, $\lambda_{G1} = 0.01$, $\lambda_V = 0.005/0.01$, $lr = 0.01$. As our CEP model needs to simultaneously optimize two prediction functions and different levels of the same factor can be optimized for different objective functions. Therefore, to balance the performances of two prediction tasks, the best levels of seven parameters (i.e., $\lambda_A, \lambda_U, \lambda_G, \lambda_{U1}, \lambda_{G1}, \lambda_V, lr$) are as follows: 4, 5, 4, 3, 3, 3, 4. We could find that Table A.4 does not contain this combination of levels. The corresponding results of above parameter combination on Pre@5, Recall@5, and F@5 for grouping behavior prediction are 0.121, 0.192, 0.148 (0.059, 0.100, 0.074 on consumption behavior prediction). Compared with the best performance in Table A.4, the results are better on two prediction tasks. Therefore, the best combination of parameter settings of CEP model is as follows: $\lambda_A = 1$, $\lambda_U = 3$, $\lambda_G = 0.1$, $\lambda_{U1} = 0.01$, $\lambda_{G1} = 0.01$, $\lambda_V = 0.01$, $lr = 0.01$.

Appendix E. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.knosys.2021.107248.

References

[1] S. Kemp, More than half of the people on earth now user social media, 2020, https://datareportal.com/reports/more-than-half-the-world-now-uses-social-media/. (Accessed 21 July 2020).
[2] Technavio, Social commerce market by device and geography - forecast and analysis 2020–2024, 2020, https://www.technavio.com/report/social-commerce-market-industry-analysis&utm_source=pressrelease&utm_medium=bw&utm_campaign=t25_v2_wk26_rfs(middle)&utm_content=IRTNTR43965/. (Accessed 29 June 2020).
[3] M. Iqbal, Facebook revenue and usage statistics 2020, 2020, https://www.businessofapps.com/data/facebook-statistics/#1/. (Accessed 30 July 2020).
[4] K. Lewin, Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change, Hum. Relat. 1 (1) (1947) 5–41, http://dx.doi.org/10.1177/001872674700100103.
[5] D.R. Forsyth, Group Dynamics, seventh ed., Cengage Learning, Boston, 2018.
[6] C.M. Harmeling, R.W. Palmatier, E. Fang, D. Wang, Group marketing: Theory, mechanisms, and dynamics, J. Mark. 81 (4) (2017) 1–24, http://dx.doi.org/10.1509/jm.15.0495.
[7] D.J. Terry, M.A. Hogg, Group norms and the attitude-behavior relationship: A role for group identification, Pers. Soc. Psychol. Bull. 22 (8) (1996) 776–793, http://dx.doi.org/10.1177/0146167296228002.
[8] M.A. Hogg, S.A. Reid, Social identity, self-categorization, and the communication of group norms, Commun. Theory 16 (1) (2006) 7–30, http://dx.doi.org/10.1111/j.1468-2885.2006.00003.x.
[9] P. Wang, B.W. Xu, Y.R. Wu, X.Y. Zhou, Link prediction in social networks: the state-of-the-art, Sci. China-Inf. Sci. 58 (1) (2015) 1–38, http://dx.doi.org/10.1007/s11432-014-5237-y.
[10] X. Wang, R. Donaldson, C. Nell, P. Gorniak, M. Ester, J. Bu, Recommending groups to users using user-group engagement and time-dependent matrix factorization, in: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, AAAI Press, 2016, pp. 1331–1337, http://dx.doi.org/10.5555/3015812.3016008.
[11] Q. Yuan, L. Chen, S. Zhao, Factorization vs. regularization: fusing heterogeneous social relationships in top-n recommendation, in: Proceedings of the 5th ACM Conference on Recommender Systems, ACM, 2011, pp. 245–252, http://dx.doi.org/10.1145/2043932.2043975.
[12] D.M. Dunlavy, T.G. Kolda, E. Acar, Temporal link prediction using matrix and tensor factorizations, ACM Trans. Knowl. Discov. Data 5 (2) (2011) 1–27, http://dx.doi.org/10.1145/1921632.1921636.
[13] N.Z. Gong, A. Talwalkar, L. Mackey, L. Huang, E.C.R. Shin, E. Stefanov, E.R. Shi, D. Song, Joint link prediction and attribute inference using a social-attribute network, ACM Trans. Intell. Syst. Technol. 5 (2) (2014) 1–20, http://dx.doi.org/10.1145/2594455.
[14] L. Liao, X. He, H. Zhang, T.S. Chua, Attributed social network embedding, IEEE Trans. Knowl. Data Eng. 30 (12) (2018) 2257–2270, http://dx.doi.org/10.1109/TKDE.2018.2819980.
[15] X. Wu, J. Wu, Y. Li, Q. Zhang, Link prediction of time-evolving network based on node ranking, Knowl. Based Syst. 195 (2020) 105740, http://dx.doi.org/10.1016/j.knosys.2020.105740.
[16] H. Ma, H. Yang, M.R. Lyu, I. King, SoRec: Social recommendation using probabilistic matrix factorization, in: Proceedings of the 17th ACM Conference on Information and Knowledge Management, ACM, 2008, pp. 931–940, http://dx.doi.org/10.1145/1458082.1458205.
[17] Y. Koren, R. Bell, C. Volinsky, Matrix factorization techniques for recommender systems, Computer 42 (8) (2009) 30–37, http://dx.doi.org/10.1109/MC.2009.263.
[18] L. Xiang, Q. Yuan, S. Zhao, L. Chen, X. Zhang, Q. Yang, J. Sun, Temporal recommendation on graphs via long- and short-term preference fusion, in: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2010, pp. 723–732, http://dx.doi.org/10.1145/1835804.1835896.
[19] X. He, L. Liao, H. Zhang, L. Nie, T.S. Chua, Neural collaborative filtering, in: Proceedings of the 26th International Conference on World Wide Web, ACM, 2017, pp. 173–182, http://dx.doi.org/10.1145/3038912.3052569.
[20] L. Wu, Y. Yang, K. Zhang, R. Hong, Y. Fu, M. Wang, Joint item recommendation and attribute inference: An adaptive graph convolutional network approach, in: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, 2020, pp. 679–688, http://dx.doi.org/10.1145/3397271.3401144.

[21] D. Liben-Nowell, J. Kleinberg, The link-prediction problem for social networks, J. Am. Soc. Inf. Sci. Technol. 58 (7) (2007) 1019–1031, http://dx.doi.org/10.1002/asi.20591.

[22] D. Camacho, N. Panizo-Lledot, G. Bello-Orgaz, A. Gonzalez-Pardo, E. Cambria, The four dimensions of social network analysis: An overview of research methods, applications, and software tools, Inf. Fusion 63 (2020) 88–120, http://dx.doi.org/10.1016/j.inffus.2020.05.009.

[23] G. Adomavicius, A. Tuzhilin, Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions, IEEE Trans. Knowl. Data Eng. 17 (6) (2005) 734–749, http://dx.doi.org/10.1109/TKDE.2005.99.

[24] W. Pan, C. Li, GBPR: group preference based Bayesian personalized ranking for one-class collaborative filtering, in: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, ACM, 2013, pp. 2691–2697, http://dx.doi.org/10.5555/2540128.2540516.

[25] J. Tang, H. Gao, X. Hu, H. Liu, Exploiting homophily effect for trust prediction, in: Proceedings of the 6th ACM International Conference on Web Search and Data Mining, ACM, 2013, pp. 53–62, http://dx.doi.org/10.1145/2433396.2433405.

[26] L. Wu, J. Li, P. Sun, R. Hong, Y. Ge, M. Wang, DiffNet++: A neural influence and interest diffusion network for social recommendation, IEEE Trans. Knowl. Data Eng. (2020) 1–14, http://dx.doi.org/10.1109/TKDE.2020.3048414.

[27] S.-H. Yang, B. Long, A. Smola, N. Sadagopan, Z. Zheng, H. Zha, Like like alike: joint friendship and interest propagation in social networks, in: Proceedings of the 20th International Conference on World Wide Web, ACM, 2011, pp. 537–546, http://dx.doi.org/10.1145/1963405.1963481.

[28] L. Wu, Y. Ge, Q. Liu, E. Chen, R. Hong, J. Du, M. Wang, Modeling the evolution of users' preferences and social links in social networking services, IEEE Trans. Knowl. Data Eng. 29 (6) (2017) 1240–1253, http://dx.doi.org/10.1109/TKDE.2017.2663422.

[29] J. Li, L. Wu, R. Hong, K. Zhang, Y. Ge, Y. Li, A joint neural model for user behavior prediction on social networking platforms, ACM Trans. Intell. Syst. Technol. 11 (6) (2020) http://dx.doi.org/10.1145/3406540, Article 72.

[30] M. Van Vugt, T. Kameda, Evolution and groups, in: J.M. Levine (Ed.), Group Processes, Psychology Press, New York and London, 2013, pp. 297–322.

[31] L. Gao, J. Wu, Z. Qiao, C. Zhou, H. Yang, Y. Hu, Collaborative social group influence for event recommendation, in: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management, ACM, 2016, pp. 1941–1944, http://dx.doi.org/10.1145/2983323.2983879.

[32] K. De Valck, G.H. Van Bruggen, B. Wierenga, Virtual communities: A marketing perspective, Decis. Support Syst. 47 (3) (2009) 185–203, http://dx.doi.org/10.1016/j.dss.2009.02.008.

[33] R.B. Cialdini, N.J. Goldstein, Social influence: Compliance and conformity, Annu. Rev. Psychol. 55 (1) (2004) 591–621, http://dx.doi.org/10.1146/annurev.psych.55.090902.142015.

[34] H.J. Schau, A.M. Muñiz Jr., E.J. Arnould, How brand community practices create value, J. Mark. 73 (5) (2009) 30–51, http://dx.doi.org/10.1509/jmkg.73.5.30.

[35] X. Zhang, S. Li, R.R. Burke, Modeling the effects of dynamic group influence on shopper zone choice, purchase conversion, and spending, J. Acad. Mark. Sci. 46 (6) (2018) 1089–1107, http://dx.doi.org/10.1007/s11747-018-0590-9.

[36] H. Risselada, P.C. Verhoef, T.H.A. Bijmolt, Dynamic effects of social influence and direct marketing on the adoption of high-technology products, J. Mark. 78 (2) (2014) 52–68, http://dx.doi.org/10.1509/jm.11.0592.

[37] B. Wellman, The school child's choice of companions, J. Educ. Res. 14 (2) (1926) 126–132, http://dx.doi.org/10.1080/00220671.1926.10879675.

[38] C.A. Wang, X.M. Zhang, I.H. Hann, Socially nudged: A quasi-experimental study of friends' social influence in online product ratings, Inf. Syst. Res. 29 (3) (2018) 641–655, http://dx.doi.org/10.1287/isre.2017.0741.

[39] R.E. Burnkrant, C. Alain, Informational and normative social influence in buyer behavior, J. Consum. Res. 2 (3) (1975) 206–215, http://dx.doi.org/10.1086/208633.

[40] S.L. Feld, The focused organization of social ties, Am. J. Sociol. 86 (5) (1981) 1015–1035, http://dx.doi.org/10.1086/227352.

[41] S. Hasan, S. Bagde, Peers and network growth: evidence from a natural experiment, Manage. Sci. 61 (10) (2015) 2536–2547, http://dx.doi.org/10.1287/mnsc.2014.2109.

[42] L. Backstrom, D.P. Huttenlocher, J.M. Kleinberg, X. Lan, Group formation in large social networks: Membership, growth, and evolution, in: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2006, pp. 44–54, http://dx.doi.org/10.1145/1150402.1150412.

[43] R.P. Bagozzi, U.M. Dholakia, Intentional social action in virtual communities, J. Interact. Mark. 16 (2) (2002) 2–21, http://dx.doi.org/10.1002/dir.10006.

[44] L.M. Aiello, A. Barrat, R. Schifanella, C. Cattuto, F. Menczer, Friendship prediction and homophily in social media, ACM Trans. Web. 6 (2) (2012) 1–33, http://dx.doi.org/10.1145/2180861.2180866.

[45] M. Girvan, M.E.J. Newman, Community structure in social and biological networks, Proc. Natl. Acad. Sci. USA 99 (12) (2002) 7821–7826, http://dx.doi.org/10.1073/pnas.122653799.

[46] C.G. Akcora, B. Carminati, E. Ferrari, User similarities on social networks, Soc. Netw. Anal. Min. 3 (3) (2013) 475–495, http://dx.doi.org/10.1007/s13278-012-0090-8.

[47] L. Lü, C.H. Jin, T. Zhou, Similarity index based on local paths for link prediction of complex networks, Phys. Rev. E 80 (4 Pt 2) (2009) 046122, http://dx.doi.org/10.1103/PhysRevE.80.046122.

[48] G. Jeh, J. Widom, SimRank : A measure of structural-context similarity, in: Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2002, pp. 538–543, http://dx.doi.org/10.1145/775047.775126.

[49] Z. Zhang, J. Wen, L. Sun, Q. De Ng, S. Su, P. Yao, Efficient incremental dynamic link prediction algorithms in social network, Knowl. Based Syst. 132 (sep.15) (2017) 226–235, http://dx.doi.org/10.1016/j.knosys.2017.06.035.

[50] S. Cavallari, E. Cambria, H. Cai, C.C. Chang, V.W. Zheng, Embedding both finite and infinite communities on graphs, IEEE Comput. Intell. Mag. 14 (3) (2019) 39–50, http://dx.doi.org/10.1109/MCI.2019.2919396.

[51] C. Peng, X. Wang, J. Pei, W. Zhu, A survey on network embedding, IEEE Trans. Knowl. Data Eng. 31 (5) (2019) 833–852, http://dx.doi.org/10.1109/TKDE.2018.2849727.

[52] Q. Zhang, R. Wang, J. Yang, L. Xue, Knowledge graph embedding by translating in time domain space for link prediction, Knowl. Based Syst. 212 (2020) 106564, http://dx.doi.org/10.1016/j.knosys.2020.106564.

[53] D. Wang, C. Peng, W. Zhu, Structural deep network embedding, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2016, pp. 1225–1234, http://dx.doi.org/10.1145/2939672.2939753.

[54] Y. Koren, R. Bell, Advances in collaborative filtering, in: F. Ricci, L. Rokach, B. Shapira, P.B. Kantor (Eds.), Recommender Systems Handbook, Springer US, Boston, 2011, pp. 145–186.

[55] Y. Hu, Y. Koren, C. Volinsky, Collaborative filtering for implicit feedback datasets, in: Proceedings of the 8th IEEE International Conference on Data Mining, IEEE, 2008, pp. 263–272, http://dx.doi.org/10.1109/ICDM.2008.22.

[56] R.R. Salakhutdinov, A. Mnih, Probabilistic matrix factorization, in: Proceedings of the 20th International Conference on Neural Information Processing Systems, Curran Associates Inc., 2008, pp. 1257–1264, http://dx.doi.org/10.3233/IFS-141462.

[57] M. Jiang, P. Cui, F. Wang, W. Zhu, S. Yang, Scalable recommendation with social contextual information, IEEE Trans. Knowl. Data Eng. 26 (11) (2014) 2789–2802, http://dx.doi.org/10.1109/TKDE.2014.2300487.

[58] M. Jamali, M. Ester, Trustwalker: A random walk model for combining trust-based and item-based recommendation, in: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2009, pp. 397–406, http://dx.doi.org/10.1145/1557019.1557067.

[59] L. Chen, L. Wu, R. Hong, K. Zhang, M. Wang, Revisiting graph based collaborative filtering: A linear residual graph convolutional network approach, in: Proceedings of the AAAI Conference on Artificial Intelligence, AAAI, 2020, pp. 27–34, http://dx.doi.org/10.1609/aaai.v34i01.5330.

[60] L. Wu, X. He, X. Wang, K. Zhang, M. Wang, A survey on neural recommendation: from collaborative filtering to content and context enriched recommendation, IEEE Trans. Knowl. Data Eng..

[61] R. Ying, R. He, K. Chen, P. Eksombatchai, W.L. Hamilton, J. Leskovec, Graph convolutional neural networks for web-scale recommender systems, in: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, ACM, 2018, pp. 974–983, http://dx.doi.org/10.1145/3219819.3219890.

[62] X. Wang, X. He, M. Wang, F. Feng, T.S. Chua, Neural graph collaborative filtering, in: Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, 2019, pp. 165–174, http://dx.doi.org/10.1145/3331184.3331267.

[63] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, M. Wang, LightGCN: Simplifying and powering graph convolution network for recommendation, in: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 639–648, http://dx.doi.org/10.1145/3397271.3401063.

[64] L. Xiong, X. Chen, T.K. Huang, J.G. Schneider, J.G. Carbonell, Temporal collaborative filtering with bayesian probabilistic tensor factorization, in: Proceedings of the 2010 SIAM International Conference on Data Mining, SIAM, 2010, pp. 211–222, http://dx.doi.org/10.1137/1.9781611972801.19.

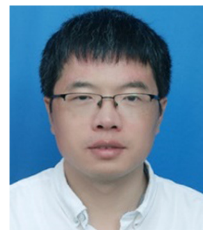
[65] C. Zhang, K. Wang, H. Yu, J. Sun, E.-P. Lim, Latent factor transition for dynamic collaborative filtering, in: Proceedings of the 2014 SIAM International Conference on Data Mining, SIAM, 2014, pp. 452–460, http://dx.doi.org/10.1137/1.9781611973440.52.

[66] C.Y. Wu, A. Ahmed, A. Beutel, A.J. Smola, H. Jing, Recurrent recommender networks, in: Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, ACM, 2017, pp. 495–503, http://dx.doi.org/10.1145/3018661.3018689.

[67] X. Wu, B. Shi, Y. Dong, C. Huang, N. Chawla, Neural tensor factorization for temporal interaction learning, in: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, ACM, 2019, pp. 537–545, http://dx.doi.org/10.1145/3289600.3290998.
[68] M. Jamali, M. Ester, A matrix factorization technique with trust propagation for recommendation in social networks, in: Proceedings of the 4th ACM Conference on Recommender Systems, ACM, 2010, pp. 135–142, http://dx.doi.org/10.1145/1864708.1864736.
[69] D.Z. Levin, R. Cross, The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer, Manage. Sci. 50 (11) (2004) 1477–1490, http://dx.doi.org/10.1287/mnsc.1030.0136.
[70] M. Gartrell, X. Xing, Q. Lv, A. Beach, R. Han, S. Mishra, K. Seada, Enhancing group recommendation by incorporating social relationship interactions, in: Proceedings of the 16th ACM International Conference on Supporting Group Work, ACM, 2010, pp. 97–106, http://dx.doi.org/10.1145/1880071.1880087.
[71] G. Kossinets, D.J. Watts, Empirical analysis of an evolving social network, Science 311 (5757) (2005) 88–90, http://dx.doi.org/10.1126/science.1116869.
[72] L. Bottou, Large-scale machine learning with stochastic gradient descent, in: Proceedings of the 19th International Conference on Computational Statistics, PHYSICA-VERLAG GMBH & CO, 2010, pp. 177–186, http://dx.doi.org/10.1007/978-3-7908-2604-3_16.
[73] S. Gao, L. Denoyer, P. Gallinari, Temporal link prediction by integrating content and structure information, in: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, ACM, 2011, pp. 1169–1174, http://dx.doi.org/10.1145/2063576.2063744.
[74] Z. Liu, W. Cao, Z. Gao, J. Bian, H. Chen, Y. Chang, T.Y. Liu, Self-paced ensemble for highly imbalanced massive data classification, in: 2020 IEEE 36th International Conference on Data Engineering, ICDE, IEEE, 2020, pp. 841–852, http://dx.doi.org/10.1109/ICDE48307.2020.00078.
[75] C. Wang, C. Deng, Z. Yu, D. Hui, R. Luo, Adaptive ensemble of classifiers with regularization for imbalanced data classification, Inf. Fusion 69 (2021) 81–102, http://dx.doi.org/10.1016/j.inffus.2020.10.017.
[76] G. Du, J. Zhang, S. Li, C. Li, Learning from class-imbalance and heterogeneous data for 30-day hospital readmission, Neurocomputing 420 (2021) 27–35, http://dx.doi.org/10.1016/j.neucom.2020.08.064.
[77] J. Khaw, B.S. Lim, L. Lim, Optimal design of neural networks using the Taguchi method, Neurocomputing 7 (3) (1995) 225–245, http://dx.doi.org/10.1016/0925-2312(94)00013-I.
[78] Shangce Gao, MengChu Zhou, Yirui Wang, Jiujun Cheng, Hanaki Yachi, Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction, IEEE Trans. Neural Networks Learn. Syst. 30 (2) (2019) 601–614, http://dx.doi.org/10.1109/TNNLS.2018.2846646.
[79] B.W. Tuckman, Developmental sequence in small groups, Psychol. Bull. 63 (63) (2001) 384–399, http://dx.doi.org/10.1037/h0022100.
[80] L. Ma, R. Krishnan, A.L. Montgomery, Latent homophily or social influence? An empirical analysis of purchase within a social network, Manage. Sci. 61 (2) (2015) 454–473, http://dx.doi.org/10.1287/mnsc.2014.1928.

Lu Yang is working toward the Ph.D. degree in Management Science and Engineering from Hefei University of Technology, China. She received her BE degree in E-commerce from Hefei University of Technology, Hefei, China. Her research interests include recommender systems, group dynamics and data mining, especially customer preference modeling and group interaction.

Yezheng Liu is a professor of Electronic Commerce in the School of Management at the Hefei University of Technology, China. He received his Ph.D. in Management Science and Engineering from Hefei University of Technology in 2001. His main research interests include decision science, electronic commerce, intelligent decision support systems and data mining. His current research focuses on online social network, personalized recommendation system and big data analytics. He is the author and coauthor of numerous papers in scholarly journals, including Marketing Science, Decision Support Systems, International Journal of Production Economics, Knowledge-Based Systems, Journal of Management Sciences in China. He is a national member of New Century Talents Project.



Yuanchun Jiang is a professor of Electronic Commerce in the School of Management at the Hefei University of Technology, China. He received his Ph.D. in Management Science and Engineering from Hefei University of Technology in 2011. He teaches electronic commerce, business intelligence and business research methods. His research interests include online marketing, electronic commerce and data mining. He has published papers in journals such as Marketing Science, Decision Support Systems, IEEE Transaction on Software Engineering, International Journal of Production Economics, Journal of Systems and Software, Knowledge-Based Systems, International Journal of Information Technology and Decision Making, amongst others.

Le Wu (corresponding author) is an associate professor of Lab of Media Computing in the School of Computer and Information at the Hefei University of Technology, China. She received the Ph.D. degree in Computer Science from University of Science and Technology of China (USTC) in 2016. Her general area of research is data mining, recommender system and social network analysis. She has published 40+ papers in referred journals and conferences, such as IEEE Trans. on Knowledge and Data Engineering, IEEE Trans. on Systems, Man, and Cybernetics: Systems, ACM Trans. on Intelligent System and Technology. Dr. Le Wu is the recipient of the Best of SDM 2015 Award, and recipient of the Distinguished Dissertation Award from China Association for Artificial Intelligence (CAAI).

Jianshan Sun is an associate professor of Electronic Commerce in the School of Management at the Hefei University of Technology, China. He received his Ph.D. in Information Systems from City University of Hong Kong in 2014. His research interests include personalized recommender systems, social network analysis and big data analytics. His research has been presented at international conferences such as ICIS, HICSS and PACIS and published in several journals including: Tourism Management, Journal of the Association for Information Science and Technology Decision Support Systems, International Journal of Production Economics, Information Processing and Management, Future Generation Computer Systems, amongst others.