

Review Article

Research on the Prediction System of Event Attendance in an Event-Based Social Network

Tianming Lan ^{1,2}, Lei Guo,¹ Xiaofei Li,¹ and Guangfu Chen¹

¹Fujian Key Laboratory of Big Data Application and Intellectualization for Tea Industry, College of Mathematics and Computer, Wuyi University, Nanping 354300, China

²School of Information Management, Jiangxi University of Finance and Economics, Nanchang 330013, China

Correspondence should be addressed to Tianming Lan; 67622035@qq.com

Received 28 March 2022; Revised 19 July 2022; Accepted 25 July 2022; Published 5 August 2022

Academic Editor: Dharmendra Singh Rajput

Copyright © 2022 Tianming Lan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, an event-based social network recommendation system has attracted more and more researchers' attention. Most EBSN recommendation systems mainly focus on recommending events to users. However, in many daily activities, it is necessary to accurately estimate the number of event participants for EBSN event organizers. As an effective means to solve the problem of event attendance prediction, the EBSN event attendance prediction system needs to mine the context information in EBSN fully and use the information to alleviate the problems of data sparsity and cold start. It brings some new challenges to the research of EBSN event attendance prediction systems. According to user characteristics and context factors, the main task of the EBSN event attendance prediction system is to obtain accurate user preferences, adopt efficient prediction algorithms to improve prediction performance, and avoid losses. This paper summarizes the research progress of the EBSN event attendance prediction system in recent years. Firstly, this paper analyzes the recent research on event attendance prediction in EBSN; secondly, we summarize the role, significance, and challenges of EBSN event attendance prediction; third, we summarize the critical technologies of EBSN event attendance prediction; the contents include mining the contextual information that affects the user's participation in the event, user preference acquisition, the method of event attendance prediction, the data set of event attendance prediction, the evaluation indicators of event attendance prediction, etc.; fourth, we look forward to the future development directions of event attendance prediction from six aspects: the methods of integrating contextual factors, the user preference acquisition methods, the prediction algorithms, the utility evaluation of event attendance prediction, the user information security, and privacy protection, and the cold start issues; finally, we conclude this paper.

1. Introduction

Over the past few years, the event-based social network (EBSN) has become a new way for users to organize and participate in various social events [1]. EBSN connects online communities and offline events, making it easy for users to plan, organize, manage, or attend events. The typical EBSN platforms include Meetup, Douban events, Eventbrite, Facebook events, Foursquare, etc.

EBSN consists of the online network and the offline network. The online network is made up of users participating in online groups, and the offline network is made up of users

participating in events organized in reality. Figure 1 shows the common EBSN network structure framework. The framework includes objects such as users, interest groups, events, and event locations [2].

With the popularity of EBSN, the EBSN recommendation system has attracted more and more researchers' attention. Most EBSN recommendation systems mainly focus on recommending events to users. However, a large number of new events are released in EBSN every day, making it difficult for users to find favourite events to participate in among the massive events. It also makes the user's decision to participate in the event full of uncertainty. The uncertainty will



FIGURE 1: EBSN network structure.

result in many issues when the event organizer hosts the event, such as site preparation and material preparation. So, it is necessary to provide accurate predictions of the participating users of the event for the event organizer. The event attendance prediction system is an effective means to solve the problem of event attendance prediction. Therefore, research on event attendance prediction has a significant practical meaning and reference value.

Event attendance prediction is common in EBSN and location-based social networks (LBSN). For example, Koolwal and Mohbey [3] gave an extensive overview of location prediction in LBSN. Much research about the EBSN recommendation system has also emerged. According to the recommended content, Liao et al. [2] divided EBSN recommendation system into event recommendation [4], event arrangement [5, 6], group recommendation, joint recommendation, friend recommendation, venue recommendation, EBSN event attendance prediction, etc.

Among them, EBSN event attendance prediction drew the attention of many researchers and produced many research results. Relevant findings can be seen at authoritative international and national academic conferences and journals, such as SIGKDD, SIGMOD, Ubicomp, Recsys, CIKM, ICDM, SDM, JSA, DASFAA, and TWEB.

The scope of this paper is limited to event attendance prediction in EBSN. This paper mainly discusses the following three questions:

Q1: what are the concept, framework, and role of EBSN event attendance prediction?

Q2: what are the critical technologies for EBSN event attendance prediction?

Q3: what are the future directions of EBSN event attendance prediction research?

The purpose of this paper is to comprehensively review the progress of the EBSN event attendance prediction system. Readers can quickly understand and enter the EBSN event attendance prediction field by reading this article. This paper provides a reference for promoting the innovation of the EBSN event attendance prediction system and mining the richness of this research field. It provides services for researchers, practitioners, and educators interested in the EBSN event attendance prediction system. We hope that it can be helpful for them to choose EBSN event attendance prediction tasks.

In conclusion, the main contributions of this paper are in three aspects:

- (1) We summarized the EBSN event attendance prediction and prediction models
- (2) We provide an overview and summary of the critical technologies of EBSN event attendance prediction
- (3) We discussed the difficulties and challenges of EBSN event attendance prediction research and determined the new trend and future development direction to share and expand the vision of EBSN event attendance prediction system research

This paper summarizes the prediction of participation in the event from recent research. The rest of the paper is organized as follows: In Section 1, we summarize the EBSN event attendance prediction. In Section 2, we provide the basic framework for predicting participation in the event. In Section 3, we summarize and compare the key technologies of EBSN event attendance prediction. In Section 4, we analyze the future development direction in EBSN event attendance prediction. In Section 5, we present our concluding remarks.

2. Overview of EBSN Event Attendance Prediction

Event attendance prediction refers to the prediction of users who will participate in the event, which can provide precise target users for the event organizer and resolve issues users and organizers face.

Liu et al. [1] first identified the issue of prediction of event attendance in the EBSN as follows: Given an event e , predict the users who will participate in the event e during a t period before the event begins.

2.1. Role and Significance of EBSN Event Attendance Prediction. When organizing and planning events, event organizers will be challenged by the uncertainty of user participation rate and the popularity of events [7]. Accurately predicting the participants in an event can help the event organizer to organize the event successfully. Thus, the event organizer can carry out accurate event planning and advertising. The event organizer needs to plan resources according to the event attendance prediction. When the predicted number of participants is seriously inconsistent with the actual number, the organizer will suffer significant economic losses. Therefore, it is essential to predict the event participants accurately, which can reduce the losses of the event organizers and enable the participants to obtain high satisfaction. Problems caused by inaccurate predictions of event participants include the following:

- (1) Material preparation: when the event organizer overestimates the number of participants in the event, it may lead to an oversupply of event-related materials, causing unnecessary waste. On the contrary, some users will have no material available, which will worsen the experience of attending the event
- (2) Venue preparation: if the capacity of the event venue is too large, the event organizer will suffer economic

losses. If the capacity of the event venue is too small, it will affect the event, and the user's participation experience will be affected

- (3) Keeping order: when the number of attendees far exceeds the estimate, event organizers may be ill-considered in staffing arrangements, which will cause order chaos and bring great difficulties to keeping order. For example, there may be a scene of congestion and chaos when people check in or enter the venue
- (4) Bad reputation: participants with poor experience participating in events will give negative comments, which will bring bad reputations to the event organizer. Thus, it will affect the event organizer acquire new users, enhance the activity and sense of belongingness of existing users, and expand brand awareness

Consequently, accurate prediction of user participation events is the key to successful organizational events. Understanding event attendance will help the organizers make appropriate plans for accurate publicity and advertising.

This paper profoundly studies critical technologies for predicting event participants in the EBSN platform. It will help event organizers estimate the number of event participants more accurately, avoid losses, and solve the problem of accurately selecting users for the EBSN recommendation system. It will improve the influence of the EBSN platform and improve users' satisfaction with the platform, attract more users to the EBSN platform, and make the platform development in a virtuous circle.

2.2. Challenges of EBSN Event Attendance Prediction. Event organizers face three main challenges in planning events for EBSN event attendance prediction.

- (1) Mining critical contextual information affects the user's decision to participate in events. Much contextual information about users and events in EBSN affects users' decision to participate in events to varying degrees. The contextual information includes the user's friend information, the user's favourite organizer information, the title, content, location, time, organizer of the event, etc. How to fully mine and fuse the contextual information to obtain accurate participating users is a challenge for event attendance prediction
- (2) Get user preferences. Users in EBSN rarely score events after participating in events, making users' explicit preferences sparse. However, much contextual information affects users' decision to participate in events in EBSN, which is relatively easy to obtain. In a multidimensional environment, user data will be sparser, and user preference acquisition will be more complex. Users have different preferences in different contextual environments. It is also challenging to integrate these contextual preferences to get user preferences. Therefore, how to acquire or learn user

preferences using data mining, machine learning, and deep learning is another challenge for event attendance prediction

- (3) Event cold start. The new events initiated by the event organizer are all upcoming events with timeliness. Predicting event participants and recommending events to target users are limited between event creation and the event starts. Before the event, there was no information about users' participation and evaluation in EBSN. As a result, an apparent cold start problem is another serious challenge for event attendance prediction

From the above, event attendance prediction is a challenging task. It is necessary to find new methods and strategies for EBSN event attendance prediction systems, improve the performance of event attendance prediction, and solve the problems event organizers face.

In order to solve the problems and challenges of event attendance prediction, many researchers have proposed various solutions. This paper summarizes the prediction framework of most of the studies as follows.

3. Framework for the EBSN Event Attendance Prediction System

The EBSN event attendance prediction system is a typical application of EBSN, similar to the EBSN recommendation system. However, due to the lack of explicit preference, the existing research focuses on how to introduce contextual information into the prediction process.

We analyzed existing research on event attendance prediction and summarized the process for the EBSN event attendance prediction system in three phases. The framework for the EBSN event attendance prediction system is shown in Figure 2.

- (1) Data collection: collect relevant data such as user, event, contextual information, records of users' participation in events, and records of user participation in interest groups
- (2) Preference extraction: analyze the contextual information that affects user participation events, extract various user preferences using contextual information, and obtain more precise user preferences based on the degree of influence of the context
- (3) Prediction generation: based on some known user preferences and key contextual information, predict the potential preferences between users and events. Use effective methods to produce prediction results

In the EBSN event attendance prediction framework, data collection, preference acquisition, and prediction generation are essential steps to produce prediction results successfully. Some key technologies are used in these steps to ensure accurate final prediction results.

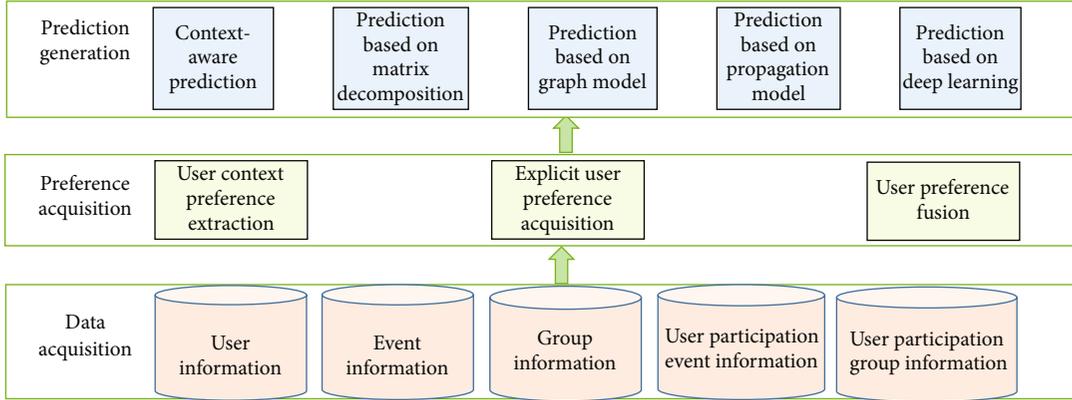


FIGURE 2: Framework for the EBSN event attendance prediction system.

4. Key Technologies of EBSN Event Attendance Prediction

4.1. Mining Contextual Information Affecting User Participation in Events. With the progress of technology, the acquisition of contextual information becomes simple. Mining users' contextual preferences are the primary work to understand users' behaviour and to provide users with better context-aware services and experience. Contextual information can be explicit or implicit and can be obtained differently. EBSN event attendance prediction considering contextual information requires users to complete extensive investigation, constantly update contextual information, and integrate contextual information. It is vital to capture contextual information and to understand users' interests and intentions. Contextual information and user behaviour complement each other and jointly determine users' interests and preferences. User behaviour reflects users' long-term preferences, and contextual information helps to find users' short-term preferences. User preferences may vary with context. EBSN event attendance prediction needs to consider different contextual information to capture accurate user preferences. Context-aware event attendance prediction provides users with more relevant and more accurate prediction results.

Many studies have mined various contextual information affecting users' participation in events. The common ones are users' social relations and event description information, such as event theme, event holding time, event holding place, event popularity, social influence of event organizers, and user tweets. Table 1 illustrates the types of contextual information that affect user engagement events and the corresponding references.

Some studies mine the contextual factors that affect users' participation in events from a unique perspective.

Karanikolaou et al. [10] analyzed the network relationship of events participated by group members and took the information extracted from the social network as the factor affecting users' participation in events. When the event is held within 10-500 kilometres from the user's location, users tend to participate in events and usually visit only a few places. At the same time, the event type, the level of trust

TABLE 1: Types of contextual information and references.

Types of contextual information	References
Users' social relations	[8–21]
Event description information	[10, 12, 14, 15, 18, 22–25]
Event popularity	[12, 16]
Social influence of event organizers	[19, 26]
User tweets	[15, 27]

and acceptance among friends, and the event attendance experience significantly impact the event participants.

Lu et al. [28] found that minimizing the uncertainty of events and participants and creating a sense of inclusiveness will encourage new users to participate in events.

Investigating event attendance from a cross-cultural perspective, Yan et al. [29] concluded that travel motivation and cultural differences are also important factors affecting user participation in events.

Zhang et al. [30] understood how weather and event duration affect the event attendance rate.

Chen et al. [31] found through interview research that the participation of EBSN members varies according to their roles. EBSN members do not all have the same participation mode but play different roles, including leaders, active contributors, fans, and peripheral members. The study found that participants with different roles may have different relationship patterns and degrees of participation with other EBSN members, switching between different roles in different situations.

Milohnić et al. [25] analyzed by collecting two questionnaires and found that the motivation of users to participate in events depends on the type of event, the motivation of different event participants is different, the motivation of male and female event participants is different, and different event attendance motivations will not affect the event experience.

Georgiev et al. [16] put forward some assumptions about the motivation of user participation and confirmed that social factors play a significant role in determining the possibility of user participation in events. Combining multiple factors is more potent than a single feature in determining the preference for event attendance.

Ding et al. [18] studied Meetup users to understand the main factors affecting their event attendance in decision-making, including social relations, time and location preferences, and user activity levels.

Cloquet and Blondel [32] combined SMS, voice call, Foursquare check-in, Twitter message, and other information to predict events.

Cesario et al. [33] studied and understood users' behaviour and mobile mode for large-scale public activities. In the paper, the prediction model using geographically tagged Twitter information is proven to be effective and reliable.

The model based on event hot words proposed by Zhang et al. [34] predicted the popularity of events on different media channels by using event streaming data.

To sum up, mining contextual information that affects user participation in events has become a research hotspot in event attendance prediction. Because the context is multi-dimensional, the user preference for events must be related to the category of contextual information. Adding various contextual information will lead to the expansion of dimension and increase sparsity. Sparsity increases with the number of contexts considered. It is almost infeasible to collect the labelled context data and users' preferences in each context to carry out a supervised learning and training system. Therefore, it is difficult to add context to event attendance prediction. However, event attendance prediction based on contextual information can provide more accurate and personalized prediction information, making it more and more popular among users. Users may need different events when considering different contextual information (e.g., time, location, and emotion). Therefore, event attendance prediction should consider the user's contextual information to make a more reasonable and appropriate prediction for users. However, it is challenging to obtain contextual information and use it in event attendance prediction in time and effectively because contextual information is dynamic and diverse and may reveal users' privacy. Research on EBSN event attendance prediction has not entirely solved these problems.

4.2. User Preference Acquisition. The accuracy of prediction largely depends on the quality and quantity of user preferences. However, users in EBSN tend to score few or no events, resulting in low prediction accuracy. A major problem in designing an effective EBSN event attendance prediction system is finding an effective and reliable method to obtain and modify user preferences. Users rarely know all their preferences clearly from the beginning. User preferences are often the result of the influence of user personalization and contextual information. The contextual information related to users and events in EBSN provides rich meaning, which can be used to understand user preferences.

Most existing studies on event attendance prediction have considered the impact of multiple contextual information on user preferences. When obtaining the comprehensive preference of users for events, a linear combination is usually used to integrate multiple user context preferences. Some studies use users' RSVP responses to events as user preferences or adopt the deep learning model, FM model,

and iterative model to learn user preferences. Table 2 shows the methods for learning user preferences and the corresponding references.

Liu et al. [1] adopted the event and community-centred diffusion model to obtain the user model, and the acquisition formula is as follows:

$$v^{k+1} = D_p \cdot v^k. \quad (1)$$

In the above formula, v^k is a column vector representing the probability of the event reaching the user after the k th diffusion step, that is, the user preference for the event. The acquisition formula of D_p is as follows:

$$D_p = \gamma D^{\text{on}} + (1 - \gamma) D^{\text{off}}. \quad (2)$$

D defines the transition matrix in heterogeneous networks.

Wu et al. [41] used the deep learning model LSTM to capture the users' evolving event preferences. This paper studies users' spatiotemporal context preferences, exclusive preferences, and order preferences, and its acquisition formulas are as follows:

$$[c_t^{\text{cp}}, h_t^{\text{cp}}] = \text{LSTM}(cp^t, c_{t-1}^{\text{cp}}, h_{t-1}^{\text{cp}}), \quad (3)$$

$$[c_t^{\text{ep}}, h_t^{\text{ep}}] = \text{LSTM}(ep^t, c_{t-1}^{\text{ep}}, h_{t-1}^{\text{ep}}), \quad (4)$$

$$[c_t^{\text{sp}}, h_t^{\text{sp}}] = \text{LSTM}([sp^t, h_t^{\text{cp}}, h_t^{\text{ep}}], c_{t-1}^{\text{sp}}, h_{t-1}^{\text{sp}}). \quad (5)$$

$c_t^{\text{cp}}, h_t^{\text{cp}}, c_t^{\text{ep}}, h_t^{\text{ep}}, c_t^{\text{sp}}, h_t^{\text{sp}}$ represent the hidden state, and cp^t, ep^t, sp^t represent the spatiotemporal context preference, exclusive preference, and order preference, respectively.

Jiang and Li [42] used a factor decomposition machine to learn user preferences, and its learning formula is as follows:

$$r_{hu} = b_u + w^T x^{hu} + \frac{1}{2} \sum_i \left(2 \cdot x_i^{hu} v_i^T z_u + \sum_{i \neq j} x_i^{hu} x_j^{hu} v_i^T v_j \right). \quad (6)$$

The formula can learn the representation of user and context features in the same hidden space. The user preference vector v_i and each context feature have the same spatial representation. Through these representations, user interaction and features can be displayed to estimate the user preference for events.

Xu et al. [17] learned the user preference by modelling the user's personal preference and the interaction between user friends through the iterative method. The formula is as follows:

$$P^t(U, E) = \sum_{u_i \in U} \sum_{e_k \in E} [s_{i,k}^0 - I(p_{i,k}^t - h_{i,k}^t)], \quad (7)$$

$$h(u_i, e_k) = h_{i,0} \cdot \prod_{j \in N_i} [1 - I(p_{j,k} - h_{j,k}) \cdot w_{ji}]. \quad (8)$$

TABLE 2: Methods and references.

Methods	References
Linear combination	[18, 23, 30, 35–40]
Users' responses	[1]
Deep learning model	[41]
FM model	[42]
Iterative model	[17]

$P^t(U, E)$ represents the user preference for events, $I(x)$ represents the discriminant function to indicate the user's friend's event attendance selection, and $h(u_i, e_k)$ represents the threshold for judging the user's participation in events.

The linear combination method can fuse multiple user context preferences and get richer user preference information than explicit feedback. However, users must actively control the weights of different context preferences, which have certain randomness and uncertainty. It cannot wholly reflect the interests and preferences of users.

User response is explicit feedback, usually users' most direct and accurate preference reflection. It is more accurate to use user replies as user preferences. The disadvantage is that explicit feedback brings extra workload to users who may not take it seriously. Therefore, the feedback information filled in by users in a hurry cannot objectively reflect user preferences to a large extent. At the same time, the information obtained through explicit feedback is also limited.

The deep learning model can automatically learn the weights of different context preferences, and the user preferences obtained are relatively accurate. Deep learning has a strong learning ability and can solve complex problems. However, deep learning is highly dependent on data. The more considerable the amount of data, the better its performance. Then, the user preference data is often small in the context environment. It will lead to the fact that the user preferences learned by deep learning are not so accurate. Moreover, the complexity of the model in deep learning will lead to a sharp increase in the algorithm's time complexity.

The FM model introduces polynomials to add the correlation of features and adopts feature combination to reduce the work of manual participation in feature combination. However, for the coefficients of the polynomial, a large number of nonzero samples are required to solve. Moreover, the feature space is often relatively sparse, so the parameter estimation becomes quite inaccurate. Moreover, the computational complexity of the FM model is relatively high.

The iterative model can capture the dynamic preferences of users. The algorithm is simple and can reliably find the optimal user preference; however, the iterative algorithm requires initialization parameters, which directly affect the convergence efficiency and whether the optimal global solution can be obtained.

Preference acquisition is one of the core issues in event attendance prediction because the purpose of event attendance prediction is to guide event organizers to find interested or valuable users in a large number of possible options in a personalized way in a specific context. They

need to capture and simulate user preferences accurately. User models can be generated through explicit or implicit data to obtain user preferences. Learning user preferences are a way to find solutions to problems. Although the semantics of preference concepts is obvious, it is challenging to obtain user preferences and use them. The complexity of the preference acquisition problem is closely related to the number of dimensions of contextual information. Therefore, it is necessary to collect and model user, event, and contextual information to generate user preferences. Then, learn the user preferences to be used in the prediction process.

4.3. Event Attendance Prediction Method in EBSN. Integrating technology and innovation in event planning to create an impressive event experience can generate strong affinity and relationships among participants to improve the popularity of events [43]. Big data technology has become a powerful tool, enabling event planners to predict event attendance, manage users, launch targeted marketing strategies, and effectively monitor event results. Machine learning technology and deep learning technology can extract information from various sources and predict the participation of events.

When obtaining prediction targets, most of the existing studies adopt two methods.

- (1) Classification. This method uses classifiers such as logistic regression, decision tree, naive Bayes, GBT-W tree, random forest, XGBoost, bagging, support vector machine, and deep learning to judge user preferences. Determine whether the user will participate in the event
- (2) Top- N method. In this method, Bayesian optimization, matrix decomposition, event, and community influence propagation models are used to sort user preferences, and N users with the most similar events and user preferences are selected to verify whether users will participate in events. Table 3 shows the methods, ways to obtain the predictive target, and the corresponding references

According to the main technical means used in the existing event attendance prediction research, we divide the event attendance prediction methods into the following seven types: context-aware prediction method, prediction method based on matrix decomposition, prediction method based on deep learning, prediction method based on the propagation model, prediction method based on the graph model, prediction method based on iteration, and prediction method based on Bayesian optimization. Table 4 shows this paper's event attendance prediction methods and the corresponding references.

The importance of contextual information urges researchers to apply contextual information to event attendance prediction. In existing event attendance prediction studies, many studies will obtain and use the contextual information that affects users' event attendance. Therefore, context-aware prediction is the focus of many related studies. Zhang et al. [36] identified eight features, including

TABLE 3: Methods, ways, and references.

Methods	Ways	References
Classification and discrimination	Logistic regression	[36, 44, 45]
	Decision tree	[30, 36, 44, 45]
	Naive Bayes	[36, 44, 45]
	GBT-W tree	[30, 31]
	Random forest	[30, 45, 46]
	XGBoost	[46]
	Bagging	[46]
	SVM	[45, 46]
Top-N method	Deep learning	[27, 41]
	Bayesian optimization	[40]
	Matrix decomposition	[42]
	Influence propagation models	[1, 47]

TABLE 4: Event attendance prediction methods and references.

Event attendance prediction methods	References
Context-aware prediction method	[30, 36, 45, 48, 49]
Prediction method based on matrix decomposition	[35, 42]
Prediction method based on deep learning	[27, 41, 50]
Prediction method based on the propagation model	[1, 24, 47]
Prediction method based on the graph model	[23]
Prediction method based on iteration	[17]
Prediction method based on Bayesian optimization	[40]

semantic, temporal, and spatial groups, by analyzing the events participated by users and then trained the supervised learning model to predict event attendance based on the extracted features. de Lira et al. [48] used the content of non-geo-tagged posts in social media to build a machine learning classifier to infer whether users participate in events. The research on participation classification is extended to [45]. The authors use the media posts shared by users before, during, and after the event to train the classifier. It can more effectively predict event attendance.

Matrix decomposition technology can improve good results for learning the implicit factor characteristics of multiple influencing factors. Jiang and Li [42] developed a feature-based matrix decomposition prediction model to obtain event participants accurately. Du et al. [35] used the personal behaviour in EBSN to find a series of factors connecting physical and cyberspace to predict the event attendance rate. The author proposed a singular value decomposition and multifactor neighbourhood (SVD-MFN) algorithm to integrate heterogeneous factors into the framework to predict the event attendance rate.

$$\tilde{r}_{ue} = p_u^T q_e + |N(u, e; k)|^{-1/2} \sum_{i \in N(u, e; k)} w_{ei} (r_{ui} - \bar{r}_u), \quad (9)$$

where \bar{r}_u represents the average score of user u on events. $N(u, e; k)$ is composed of k nearest neighbours of all events e selected as user u 's participation. w_{ei} is a free parameter, which is learned together with the parameters of the matrix decomposition model.

The deep neural network is a common forecasting technique; for example, literature [51] used the predictive model of the deep neural network. Deep learning has significant advantages in learning more accurate user preferences from multidimensional contextual information. Mehmood et al. [27] proposed a new LSTM event attendance prediction classifier based on tweet content. Wu et al. [41] proposed a three-level hierarchical LSTM architecture to simulate users' multidimensional and changing preferences and use a multilayer perceptron (MLP) to capture the complex relationship between users' preferences to obtain the probability of users' participation in events. Rizi et al. [50] proposed a neural network classifier to predict event attendance by mining text features and network topology information in user media posts.

The influence propagation model is very effective in stimulating the influence change process of events, event organizers, users, and groups. Liu et al. [1] determined the coexistence of online and offline social interaction in social networks. Considering the unique characteristics of EBSN, including network attributes, community structure, and information flow, they proposed an event-based diffusion model and a community-based diffusion model to predict the participation of events. Zhang and Lv [24] studied the potential factors of event popularity and developed a group-based social impact communication network to simulate the impact of specific groups on events. Yu et al. [47] studied how to select potential event participants in EBSN from the perspective of event organizers. The authors propose a novel credit allocation user influence preference (cd-uip) algorithm to find the most influential and popular followers as invitees. The influence of its users on other users in event e is calculated as follows:

$$\gamma_{v,u}(e) = \frac{\text{infl}(u)}{N_{\text{in}}(u, e)} \cdot \exp\left(-\frac{t(u, e) - t(v, e)}{\tau_{v,u}(e)}\right) \cdot \text{pref}_u(e) \cdot \text{pref}_v(e). \quad (10)$$

Among them, $\tau_{v,u}(e)$ is the average time taken for activity e to propagate from user v to user u . $\text{infl}(u)$ represents the influence of user u . $N_{\text{in}}(u, e)$ represents the set of potential influences of user u participating in e . $\text{pref}_u(e)$ represents the preference of user u for event e .

The graph model can well represent heterogeneous social networks in EBSN. Zhang et al. [23] proposed a group-based event attendance prediction framework based on the group background and social-related characteristics of the participation records of past events. The framework uses a personalized approach to restart random walk on hybrid EGU (event group user) networks to capture internal social relationships and predict the participation events of group members.

The iterative method can simulate the dynamic change process of user preferences. Xu et al. [17] integrated the

dynamic interdependence of potential event participants into the discrimination process of event attendance prediction through iterative methods.

Bayesian optimization framework can optimize the pairwise preference ranking of users. Li et al. [40] combined the explicit and potential characteristics of event capture from online social networks and EBSN and proposed a Bayesian optimization framework. The framework adopts paired user preferences in social relations and events, which can optimize the preference ranking of users for events.

Context-aware prediction methods combine rich contextual information to obtain more accurate user preferences, thus improving prediction accuracy. The introduction of context information can supplement user and event information in many ways, to a certain extent, making up for the lack of data sparse. At the same time, because new users and new events have basic information, the information can be used as context information so that the cold start problem can be avoided. However, it is challenging to obtain context, and its modelling and calculation complexity is high.

The prediction method based on matrix factorization is easy to implement in programming, with low implementation complexity, good prediction effect, and good expansibility. By obtaining a variety of relationships in EBSN, the matrix factorization method can effectively obtain user preferences and interpret the same data from different angles, equivalent to increasing the amount of data and making up for the lack of data sparseness. At the same time, the cold start problem is also solved to a certain extent due to the acquisition of various relationships between users and events. However, there are some disadvantages: it takes much time to train the model, and the obtained prediction results are not very interpretable.

Prediction methods based on deep learning can utilize various types of contextual information in EBSN, including text, images, audio, and even video, while other prediction methods are challenging to utilize fully. The deep learning model does not require users to manually design features. It can automatically learn potentially useful feature representations from input data and model data nonlinearly to capture intricate user and event features. Deep learning models can obtain more accurate user preferences through training on extensive data and alleviate the cold start problem. Deep neural networks have broad applicability and high flexibility and can easily and quickly build predictive models. However, there are too many parameters in deep learning, and it is difficult to explain its specific role, so the interpretability is poor. At the same time, deep learning requires a large amount of data to complete its model training. If the data is too small, the predicted effect of the learned model may be poor. Deep learning models can sometimes require extensive hyperparameter tuning.

The prediction method based on the propagation model can improve the prediction accuracy by using information such as node attributes, but it is challenging to obtain the node information. Therefore, the prediction method based on the propagation model has specific difficulties in implementation.

The prediction method based on the graph model has high flexibility and scalability. It can make high-quality predictions in the case of insufficient data and solve the problem of insufficient data and event cold start. However, the training process of this method requires a large amount of training data, which takes too long, and the prediction results are limited by historical data, and it is challenging to produce reasonable predictions for new users and events.

The iterative prediction method utilizes the dynamic interaction between users and their friends, which can simulate the impact of social events participating in the decision-making process and effectively improve event attendance prediction. However, dynamic interaction data for iterative models are challenging to obtain.

The prediction method based on Bayesian optimization uses posterior probability to optimize user preferences. It is a sorting algorithm based on matrix decomposition. It does not perform global optimization but performs sorting optimization for each user's different preferences. It is a pairwise sorting algorithm that introduces a Bayesian prior, assuming that the parameters obey a normal distribution, reducing the model's overfitting. However, the prediction model has high complexity, requires a long learning time, and cannot predict new users and new events.

To sum up, there are many event attendance prediction methods in EBSN, among which the classifier-based and context-based methods are more studied, and other methods are relatively few. EBSN is rich in contextual information. Considering the impact of more contextual information on user preferences can obtain more accurate prediction results.

4.4. Get EBSN Event Attendance Prediction Data Set. Currently, there is no standard public data set in the field of event attendance prediction in EBSN. Most researchers use the API provided by the EBSN platform to crawl the actual data of the platform as the data set to verify the model's performance. It brings significant challenges to EBSN event attendance prediction research.

Liu et al. [1] crawled data from October 2011 to January 2012 from Meetup, including 5153886 users, 5183840 events, 97587 interest groups, and 42733136 RSVP records. Du et al. [35] crawled information of the users who participated in more than three events in Beijing in 2013 in Douban events and the event attendance history of these users from February 2012 to October 2013. It includes 15050 users, 45561 events, 6570 event organizers, 313479 event organizers' concerns, and 481325 user participation event records. Feng et al. [26] crawled data of Meetup from July 2013 to October 2013 and crawled the data of the EBSN platform Plancast. Experiments on the data set proved the superiority of the solution proposed by the author.

Because EBSN event attendance prediction will take advantage of the contextual information, some researchers should crawl the data of the EBSN platform and the context-related data set. To study how weather factors affect the event attendance rate, Zhang et al. [30] not only crawled the data from Meetup but also crawled the weather data from the website <https://tianqi.2345.com>. Wang et al. [16] crawled data for Phoenix, Chicago, and San Jose in Meetup

from January 2012 to January 2014, examined the personal data of event organizers, and crawled their social account data from Twitter.

The objects in EBSN are very important. Different research uses different objects. The available objects in EBSN have their attributes. The EBSN event attendance prediction system can provide more accurate prediction services based on these attributes. Meetup and Douban events are two typical EBSN 5platforms. The data sets crawled from these two platforms have different properties of available entities. Next, we will briefly introduce these two typical data sets.

First, we introduce the properties of the available entities in Meetup:

- (1) Users: each user in the Meetup has a profile page, which includes the user's name, location, time of joining Meetup, interest groups, and user's interest tags. When registering Meetup, the user needs to determine his/her location information by specifying the postal code so that the Meetup can recommend interest groups according to the location. For similar reasons, Meetup requires users to select a group of topics of interest when registering. The total number of topics in Meetup has exceeded 100 K
- (2) Interest groups: Meetup provides 33 categories, and each interest group belongs to one of them. Users can join interest groups, and members of interest groups can act as event organizers. Each interest group specifies a group of tags from the topic tags to represent the topic interests. Meetup allows the user to disclose the information of his/her interest group or not. Interest group information includes interest group name, location, interest group members, interest group organizer, interest group description content, published events, photos, and discussion of interest group members
- (3) Event: in the interest group, only the organizer can hold the event by specifying the title, content description, location, and time. The user can express his/her intention to participate in the event through the RSVP function. An event includes five key elements:
 - (i) Event content: brief introduction of the event
 - (ii) Event organizer: the user who organizes the event
 - (iii) Participating user: the user who will participate in the event
 - (iv) Event location: the location of the event
 - (v) Event time: when the event will be held

Secondly, we introduce the properties of available entities in the second typical data set. Douban events are the most popular EBSN in China. It contains users, interest groups, event organizers, events, and other available entities. The attributes of the entities it owns are described below.

- (1) Users: users of Douban events have attributes such as user name, user introduction and user shared diary, photo album, and message
- (2) Event: event information mainly includes name, time, place, cost, type, organizer, content, the number of people interested in the event, the number of people who will participate in the event, and the event members who will participate in the event. The event content is the event's details, including three aspects: category, title, and description
- (3) Event organizer: the event organizer is the user who will host the event and is responsible for the offline event organization. The attributes it has are events that have been or will be organized

The data set plays a significant role in evaluating the event attendance prediction model. The lack of standard data sets acquires data sets more complicated. Some studies also need to crawl contextual information, making event attendance prediction data more complex.

4.5. EBSN Event Attendance Prediction Evaluation Indicators. The evaluation index plays a critical role in testing and evaluating the performance of the EBSN event attendance prediction model. Many evaluation indexes have been used in event attendance prediction algorithms, such as precision, accuracy, recall, *F*-score, AUC, RMSE, and map. Table 5 shows the evaluation metrics and the references that used the corresponding metrics.

Precision and accuracy are standard information retrieval evaluation indexes that can evaluate the prediction list's accuracy. The recall is used to measure how many positive examples are correctly classified. *F*-score is used to consider precision and recall comprehensively. AUC is the area under the ROC curve. The value of this area will not be greater than 1, and the value range is generally between 0.5 and 1. RMSE and map are evaluation indicators based on the error between the predicted and actual scores. The smaller the value, the higher the accuracy.

Generally, an algorithm is unlikely to be better than others in all evaluation indexes. Therefore, in research, multiple evaluation indexes are usually used to compare various algorithms. It leads to selecting appropriate evaluation indicators, which has a crucial impact on the event attendance prediction algorithm.

There have been many studies on the above key technologies, but these studies still have some deficiencies as follows:

- (1) The fusion of multiple contextual information is not enough. The various contextual information used in the existing research is insufficient, and the existing fusion methods do not fully express the complex relationship of the contextual information
- (2) Getting accurate user preferences can also be improved. In the existing research, it is not perfect for obtaining users' dynamic context preferences,

TABLE 5: Evaluation metrics and references.

Evaluation metrics	References
Precision	[17, 23, 30, 35, 37, 40, 42]
Accuracy	[23, 35, 39, 46]
Recall	[1, 17, 18, 23, 30, 35, 37, 46]
F-score	[23, 35, 41, 46]
AUC	[35, 36, 40, 42, 46]
RMSE	[22, 47]
Map	[22, 40]

and using different weights for different contexts is not accurate

- (3) The prediction method has limitations. With the continuous emergence of new technologies, there will be more and more prediction methods combined with new technologies
- (4) There are few utility evaluation indexes. The evaluation index of the existing EBSN event attendance prediction system is relatively simple. A single evaluation index has difficulty evaluating the effect of the algorithm accurately

Therefore, there are still many areas worthy of research in the field of event attendance prediction in the future.

5. Future Directions of EBSN Event Attendance Prediction

There have been many studies on event attendance prediction in EBSN, but the following aspects are still worthy of further research.

5.1. Method of Integrating Contextual Factors. In EBSN, the user's explicit feedback data is very sparse. So it is necessary to combine various contextual information to get implied feedback data from the user and to understand user preferences in various contexts to generate accurate predictions. Typically, these factors will be combined and applied to event attendance prediction and are rarely used alone. The decision-making of users' participation in events may be affected by many factors, such as users' personal preferences, social relations, and the interaction between users and friends. User preferences may vary from context to context. Therefore, event attendance prediction technology needs to fully use various contextual information to promote the prediction accuracy of the prediction system.

Some contextual information, such as users' social relationships, has been considered in event attendance prediction. Backstrom et al. [52] found that the probability of individual users adopting new behaviour increases with the number of friends who have participated in the behaviour, indicating that social friendship is crucial in encouraging users to participate in social activities. Ye et al. [53] proved that the influence of friends is very important for event prediction and recommendation. Since face-to-face communication is inevitable for offline events, people usually partic-

ipate in activities with friends, making EBSN more cohesive than ordinary social networks [1]. Georgiev et al. [16] found that combining multiple factors is more potent than a single feature in determining event attendance.

When fusing context to obtain user preferences, the existing methods adopt linear combination [35], LSTM model [41], two-way FM [42], linear regression [15], etc.

More and more contextual information needs to be considered in event attendance prediction and the relationship between contextual information. The existing fusion methods have increasing difficulty expressing the complex relationship of the contextual information. Multicontextual information fusion is a method to improve the quality of event attendance prediction [16]. The problem of multicontextual information fusion has not been effectively solved in event attendance prediction. Extensively investigate and analyze critical contextual information and auxiliary information data that may affect users' decision to participate in events. Integrate heterogeneous multisource data into predictive models. Use better fusion methods to express complex relationships of contextual information. Learn the influence weight of contextual information. These will become essential directions for event attendance prediction.

5.2. User Preference Acquisition Method. The realization of event attendance prediction requires user preference data. The extensiveness of user preference acquisition can significantly improve the accuracy of user preferences. Users usually do not know their preferences. They may not express their preferences accurately, and it is not easy to know which attributes of events and which values of attributes are more important. Therefore, it is necessary to obtain user preferences through data mining methods. There are various types of user preference data, which can be roughly classified into two categories: explicit feedback and implicit feedback [54]. Explicit feedback is usually considered a more informative signal of user preference. Implicit feedback is used to infer preference information. Implicit feedback data are more affluent and easier to collect, but it has some limitations.

In EBSN, to obtain the user behaviour preferences in social networks, data mining is usually carried out on the nearest neighbour users or recent projects based on the user's historical behaviour data to find similar behaviour relationships and analyze them. It can also find their internal relations through the perception of the user's relevant contextual information to provide a basis for mining the user's implicit preferences.

User preference acquisition forms the basis of event attendance prediction. When acquiring user preferences, it will combine various contextual information, which expands the amount of data to be processed, increases the complexity of data processing, and makes the data sparser. At the same time, different contexts have different effects on user preferences and influence weights. How to extract and represent user context preferences effectively, determine the different influence degree of context, and fuse multisource contextual information to obtain more accurate user preferences is worthy of further research in event attendance prediction. On the other hand, users have various context preferences that

dynamically change over time. Therefore, how to obtain users' dynamic contextual preferences quickly is also a research direction worthy of attention.

5.3. Prediction Algorithm. The core of event attendance prediction is using a prediction algorithm to generate prediction results. Therefore, prediction generation technology greatly determines the utility and performance of event attendance prediction. Currently, the event attendance prediction generation technology still faces various problems, mainly including context acquisition technology for high-dimensional data, multidimensional contextual information fusion, prediction model selection, and parameter optimization.

In recent years, researchers have proposed some model-based event attendance prediction methods, but these models have limitations and have not fully met the needs. With the development of deep learning technology, more deep learning models exist. Deep learning models can learn more accurate user preferences. Therefore, model-based event attendance prediction methods need to be further studied.

5.4. Utility Evaluation of Event Attendance Prediction. The event attendance prediction algorithm's quality can be evaluated using various evaluation indexes, such as ROC, accuracy, recall rate, and F value. These metrics help to select high-quality users from the set of available users. However, there is no standard benchmark data set and evaluation index in the event attendance prediction. Therefore, comparing the advantages and disadvantages of different baseline algorithms used in papers is difficult. The EBSN event attendance prediction system's user ranking is mainly based on the similarity measurement between events and users. The users in the prediction list are often highly similar to each other, lacking diversity and novelty. Therefore, it is necessary to introduce other evaluation indexes to evaluate the performance of the event attendance prediction algorithm. In the event attendance prediction, in addition to considering evaluation metrics that can improve forecast quality, other issues need to be considered, such as prediction coverage and diversity.

Coverage can provide a percentage of predicting users. If only a few users or even no users are interested in an event, the prediction may be infeasible. Therefore, it is necessary to expand the coverage of event attendance prediction.

Diversity is to increase the probability of novel users related to events. Prediction based on similarity measurement may make the predicted users highly similar to each other; that is, the users are similar, leading to highly similar prediction results of similar events. Therefore, users may be dissatisfied with too few types of events provided by the system.

Different evaluation indicators are suitable for different data set characteristics and task types. Choosing the appropriate evaluation index is usually very important because different evaluation indexes may be conducive to different algorithms. Many evaluation indexes have been used in the event attendance prediction algorithm, and different evalua-

tion indexes have different emphases. One algorithm is unlikely to be better than all other algorithms in all evaluation indexes. General evaluation indexes measure the performance of the algorithm. In some cases, applying incorrect evaluation indexes may lead to inappropriate algorithms.

The utility evaluation of event attendance prediction is one of the crucial issues in the research of event attendance prediction. Due to the diversity of contextual information related to event attendance prediction, the utility evaluation of event attendance prediction becomes more difficult. Currently, the evaluation indexes of EBSN event attendance prediction systems mainly focus on accuracy but pay less attention to prediction coverage and diversity. The utility evaluation of event attendance prediction has become a significant challenge in the research of event attendance prediction.

5.5. User Information Security and Privacy Protection. In recent years, with the increasing abundance of user information, social network analysis methods are becoming more and more powerful, and user information and privacy are becoming more accessible and easier to obtain. Therefore, user information security and privacy protection have become the most concerning issues. EBSN platform provides people with a convenient and colourful life, which is very popular. Some EBSN platforms provide API access interfaces for adding functions and allow third-party personnel to access the platform's user data. Users' personal information can be obtained through these API access interfaces, such as users' location, preferences, RSVP, and event comments. The introduction of open API into EBSN creates an access method that bypasses permission settings, exposing user data to third-party personnel, bringing a severe risk of privacy disclosure, and causing serious privacy problems.

The problem of privacy disclosure faced by EBSN has attracted the attention of relevant researchers. Some researchers have investigated the problem of privacy disclosure in EBSN, and some have proposed privacy protection methods in EBSN. Chung et al. [55] tried to investigate the problem of privacy disclosure in Meetup and analyze the causes and possible damage of privacy disclosure in EBSN. The authors crawled 240000 interest groups, 8.9 million users, 27 million interest groups' participation information, and 78 million topic interest information from Meetup. Analyzing the data set shows that users' network information is highly related to real life. After learning from the data set using the logistic regression model, the user's LGBT (lesbian, gay, bisexual, and transgender) status can be predicted with 93% accuracy, which is one of the most sensitive privacy information.

Dong and Zhou [56] conducted a comprehensive study on the privacy problem when users' online social and offline social are closely coupled in Meetup and used a reasoning analysis strategy to simulate the privacy destruction process in Meetup. The results show that the user's privacy can be inferred with high accuracy using several simple and effective privacy reasoning models, which shows that the privacy threat in EBSN is severe and brings destructive consequences to users.

Dou et al. [57] proposed a privacy protection method based on weighted noise injection technology for recommendation in the social network environment. Experimental results show that this method has better privacy protection performance than the traditional noise injection method.

There are many privacy disclosure problems in EBSN. For example, in the EBSN event attendance prediction, the prediction system provides accurate prediction, which depends on the user's personal information and needs to analyze the user behaviour. Therefore, much user data is usually required, and more contextual information is used for prediction. However, the massive data information used for prediction often contains much personal privacy information. Once attackers obtain the information, they can infer the user's privacy from their history. Through additional information, attackers can even identify users in real life. In this way, the user's privacy is disclosed directly or indirectly.

Understanding the balance between protecting users' privacy and providing high-quality event attendance prediction is very important for designing a secure EBSN. Therefore, privacy protection is a challenging problem in EBSN event attendance prediction.

5.6. Cold Start Problem. The cold start problem is common and complex in event prediction. The problem will be more serious when the predicted users and events are new. New users and events will continue to appear in EBSN, so the cold start problem will never completely disappear. If the event prediction is inaccurate, new users may no longer pay attention to the event organizer.

The cold start problem is related to the sparsity of available information, such as user information and event information. The similarity between existing events and new events is inferred through the relevant information of event content, and the implicit relationship between user event preference and events is combined to put forward better prediction and alleviate the cold start problem. The prediction basis cannot be formed if no user preference information is available. User contextual and event contextual information can be used to obtain user preferences, such as additional contextual information about users (user attributes, such as gender, age, geographical location, occupation, and influence) and events (event attributes, such as event subject, description content, holding time, and place).

Studies have used the influence of event organizers and groups to alleviate the cold start problem. Liu et al. [1] studied how the cold start phenomenon affects the prediction performance. When the user scale is small, the prediction performance of the cold start problem (for example, the event creator is the only seed of the event) is slightly worse than a random start. However, the recall achieved by diffusion of a single user is still quite good, which shows that it is satisfactory to use diffusion to predict event attendance in EBSN even in extreme cold start situations.

Xu et al. [17] carefully observed two types of potential participants, namely, event organizers ("seed users" with significant social influence) and new entrants. The study found that almost all groups face serious "cold start" problems. Old

members often quit, and new members join. Usually, a high proportion of new users will lead to prediction problems. Through in-depth study of the data, it is found that significant social influence helps to overcome the "cold start" problem.

The event attendance prediction problem is unique because the event's life cycle is short. The event attendance prediction must be effective after the event is created and before the event starts. It leads to a more serious cold start problem. Providing high-quality prediction for event organizers in the case of a "cold start problem" and reliable prediction in the case of highly scarce available data deserve in-depth research.

The six directions mentioned above are the problems that need to be solved and studied urgently in the EBSN event attendance prediction system. Better event attendance prediction effects can be obtained if these problems can be solved.

6. Summary

This paper analyzes recent research on event attendance prediction in EBSN, aimed at addressing the challenges faced by the event organizer, providing a reference for the daily operation of the EBSN platform, providing better planning and coordination methods for the organization of events, and providing possible research directions for event attendance prediction in EBSN.

This paper starts with the concept of event attendance. It summarizes the role and significance of event attendance prediction, the challenges faced by event attendance prediction, the framework for EBSN event attendance prediction, and a brief overview of domestic and foreign research. It summarizes the key technologies of EBSN event attendance prediction, such as the contextual information which influences the user to participate in events, the user preference acquisition, the methods of event attendance prediction, the event attendance prediction-related data sets, and the event attendance prediction evaluation indexes. Furthermore, it also looks forward to the difficulties and development direction of EBSN event attendance prediction from eight aspects, including the method of fusing contextual factors, the method of obtaining user preferences, the prediction algorithm, the utility evaluation of EBSN event attendance prediction, the user information security, the privacy protection, the cold start problem, and the event attendance interpretation.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

This work was financially supported by the Nanping Science and Technology Plan Project (N2021J007), Natural Science Foundation of Fujian Province (2021J011147, 2021J011142), and National Natural Science Foundation of China (61772245).

References

- [1] X. Liu, Q. He, Y. Tian, W. C. Lee, J. McPherson, and J. Han, "Event-based social networks: linking the online and offline social worlds," in *In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1032–1040, New York, NY, USA, 2012, August.
- [2] G. Liao, T. Lan, X. Huang et al., "Survey on recommendation systems in event-based social networks," *Journal of Software*, vol. 32, no. 2, pp. 424–444, 2021.
- [3] V. Koolwal and K. K. Mohbey, "A comprehensive survey on trajectory-based location prediction," *Journal of Computer Science*, vol. 3, no. 2, pp. 65–91, 2020.
- [4] B. Li, G. Wang, Y. Cheng, Y. Sun, and X. Bi, "An event recommendation model using ELM in event-based social network," *Neural Computing and Applications*, vol. 32, no. 18, pp. 14375–14384, 2020.
- [5] Y. Cheng, Y. Yuan, L. Chen, C. Giraud-Carrier, G. Wang, and B. Li, "Event-participant and incremental planning over event-based social networks," *Engineering*, vol. 33, no. 2, pp. 1–488, 2019.
- [6] F. Kou, Z. Zhou, H. Cheng, J. Du, Y. Shi, and P. Xu, "Interaction-aware arrangement for event-based social networks," in *In Proceedings of the 35th International Conference on Data Engineering (ICDE)*, pp. 1638–1641, Macau, 2019 April.
- [7] O. J. Leong and R. Mafas, "A review on the predictive algorithms for event attendance prediction in event-based social networks," *Journal of Critical Reviews*, vol. 7, no. 9, pp. 1654–1659, 2020.
- [8] Y. Liao, X. Lin, and W. Lam, "Exploring influence among participants for event recommendation," in *in proceedings of the 8th IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM)*, pp. 417–420, San Francisco, 2016.
- [9] H. Itoga and G. T. Lin, "Using facebook for event promotion-implementing change," *African Journal of Business Management*, vol. 7, no. 28, pp. 2788–2793, 2013.
- [10] S. Karanikolaou, I. Boutsis, and V. Kalogeraki, "Understanding event attendance through analysis of human crowd behavior in social networks," in *In Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems*, pp. 322–325, Mumbai, 2014.
- [11] M. Bogaert, M. Ballings, and D. Van den Poel, "The added value of Facebook friends data in event attendance prediction," *Decision Support Systems*, vol. 82, pp. 26–34, 2016.
- [12] G. Li, Y. Liu, B. Ribeiro, and H. Ding, "On group popularity prediction in event-based social networks," in *In Twelfth International AAAI Conference on Web and Social Media*, pp. 644–647, Stanford, 2018, June.
- [13] W. Tu, D. W. Cheung, N. Mamoulis, M. Yang, and Z. Lu, "Activity-partner recommendation," in *In Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 591–604, Springer, Cham. Ho Chi Minh City, 2015, May.
- [14] J. Y. Jiang and C. T. Li, "Analyzing social event participants for a single organizer," in *In Tenth international AAAI conference on web and social media*, pp. 599–602, Cologne, 2016, March.
- [15] Y. Hu, S. Farnham, and K. Talamadupula, "Predicting user engagement on Twitter with real-world events," in *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 9, 2015no. 1, pp. 168–177, 2015.
- [16] P. I. Georgiev, A. Noulas, and C. Mascolo, "The call of the crowd: event participation in location-based social services," in *In Eighth International AAAI Conference on Weblogs and Social Media*, pp. 141–150, 2014, May.
- [17] T. Xu, H. Zhong, H. Zhu, H. Xiong, E. Chen, and G. Liu, "Exploring the impact of dynamic mutual influence on social event participation," in *in Proceedings of the 11th International Conference on Data Mining*, pp. 262–270, Las Vegas: CSREA, 2015.
- [18] H. Ding, C. Yu, G. Li, and Y. Liu, "Event participation recommendation in event-based social networks," in *In Proceedings of the 8th International conference on social informatics*, pp. 361–375, Bellevue, 2016.
- [19] Z. Wang, Y. Zhang, Y. Li, Q. Wang, and F. Xia, "Exploiting social influence for context-aware event recommendation in event-based social networks," in *Proceedings of the 36th IEEE Conference on Computer Communications*, 2017, pp. 1–9, Atlanta, 2017.
- [20] R. Du, Z. W. Yu, Z. L. Liu, and B. Guo, "Social influence of online and offline based on events," *Journal of Computers*, vol. 37, no. 1, pp. 238–245, 2014.
- [21] G. Liao, Y. Zhao, S. Xie, and P. S. Yu, "An effective latent networks fusion based model for event recommendation in offline ephemeral social networks," in *in Proceedings of the 22nd ACM International Conference on Conference on Information & Knowledge Management*, pp. 1655–1660, San Francisco: ACM, 2013.
- [22] J. S. Zhang and Q. Lv, "Understanding event organization at scale in event-based social networks," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 10, no. 2, pp. 1–23, 2019.
- [23] S. Zhang and Q. Lv, "Hybrid EGU-based group event participation prediction in event-based social networks," *Knowledge-Based Systems*, vol. 143, pp. 19–29, 2018.
- [24] S. Zhang and Q. Lv, "Event organization 101: understanding latent factors of event popularity," in *Proceedings of the 11th International Conference on Web and Social*, vol. 11, 2017no. 1, pp. 716–719, Media, Montreal: AAAI, 2017.
- [25] I. Milohnić, K. Lesić, and T. Slamar, "Understanding the motivation for event participating-a prerequisite for sustainable event planning," in *In Proceedings of the 23rd Biennial International Congress, Tourism & Hospitality Industry 2016 (THI2016)*, pp. 204–218, Trends and Challenges, Opatija: Croatia, 2016.
- [26] K. Feng, G. Cong, S. S. Bhowmick, and S. Ma, "In search of influential event organizers in online social networks," in *In Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pp. 63–74, New York, NY, USA, 2014, June.
- [27] U. Mehmood, I. Moser, and N. Ronald, "Event attendance prediction using social media," in *In Proceedings of the Australasian Computer Science Week Multiconference*, pp. 1–7, Melbourne: ACM, 2020.
- [28] D. Lu, R. Farzan, and C. Lopez, "To go or not to go! What influences newcomers of hybrid communities to participate offline," in *Proceedings of the 8th International Conference on Communities and Technologies*, 2017, pp. 159–168, Troyes: ACM, 2017.
- [29] N. Yan and E. Halpenny, "The role of cultural difference and travel motivation in event participation," *International Journal of Event and Festival Management*, vol. 10, no. 2, pp. 155–173, 2019.

- [30] J. Zhang, W. Jiang, J. Zhang, J. Wu, and G. Wang, "Exploring weather data to predict activity attendance in event-based social network," *ACM Transactions on the Web (TWEB)*, vol. 15, no. 2, pp. 1–25, 2021.
- [31] J. Chen, X. Wang, J. Beck, C. Wu, and J. M. Carroll, "Beyond leaders and followers: understanding participation dynamics in event-based social networks," *International Journal of Human-Computer Interaction*, vol. 35, no. 20, pp. 1892–1905, 2019.
- [32] C. Cloquet and V. D. Blondel, "Forecasting event attendance with anonymized mobile phone data," *Big Data Research*, pp. 1–26, 2014.
- [33] E. Cesario, F. Marozzo, D. Talia, and P. Trunfio, "SMA4TD: a social media analysis methodology for trajectory discovery in large-scale events," *Online Social Networks and Media*, vol. 3-4, no. 4, pp. 49–62, 2017.
- [34] Q. Zhang, Y. Tang, Z. Guo, T. Wu, L. Xu, and X. Gao, "Adaptive general event popularity analysis on streaming data," in *In 2018 IEEE/ACM 5th International Conference on Big Data Computing Applications and Technologies (BDCAT)*, pp. 116–125, Zurich:IEEE, 2018, December.
- [35] R. Du, Z. Yu, T. Mei, Z. Wang, Z. Wang, and B. Guo, "Predicting activity attendance in event-based social networks: content, context and social influence," in *In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 425–434, Seattle:ACM, 2014.
- [36] X. Zhang, J. Zhao, and G. Cao, "Who will attend? – predicting event attendance in event-based social network," in *In Proceedings of the 16th IEEE International Conference on Mobile Data Management*, pp. 74–83, Pittsburgh, 2015.
- [37] T. Trinh, N. T. Nguyen, D. Wu, J. Z. Huang, and T. Z. Emara, "A new location-based topic model for event attendees recommendation," in *In Proceedings of the 2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1–6, Danang:IEEE, 2019.
- [38] R. Marinescu-Istodor, A. S. Sayem, and P. Fränti, "Activity event recommendation and attendance prediction," *Journal of location based services*, vol. 13, no. 4, pp. 293–319, 2019.
- [39] G. Nkuna and M. Coetzee, "Social event invitation and recommendation for event based social networks," in *In Proceedings of the 2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, pp. 1–5, Durban:IEEE, 2020.
- [40] S. Li, X. Cheng, S. Su et al., "Followee recommendation in event-based social networks," in *In Proceedings of the 21st International Conference on Database Systems for Advanced Applications*, pp. 27–42, Dallas:Springer, Cham, 2016.
- [41] X. Wu, Y. Dong, B. Shi, A. Swami, and N. V. Chawla, "Who will attend this event together? Event attendance prediction via deep LSTM networks," in *In Proceedings of the 2018 SIAM International Conference on Data Mining*, pp. 180–188, San Diego: SIAM, 2018.
- [42] J. Y. Jiang and C. T. Li, "Who should I invite: predicting event participants for a host user," *Knowledge and Information Systems*, vol. 59, no. 3, pp. 629–650, 2019.
- [43] P. Crowther, "Marketing event outcomes: from tactical to strategic," *International Journal of Event and Festival Management*, vol. 2, no. 1, pp. 68–82, 2011.
- [44] X. Zhang and G. Cao, *Event Attendance Prediction in Social Networks*, SpringerBriefs in Statistics, Springer, Cham, 2021.
- [45] V. M. de Lira, C. Macdonald, I. Ounis, R. Perego, C. Renso, and V. C. Times, "Event attendance classification in social media," *Information Processing & Management*, vol. 56, no. 3, pp. 687–703, 2019.
- [46] J. K. Nguyen, A. Karg, A. Valadkhani, and H. McDonald, "Predicting individual event attendance with machine learning: a 'step-forward' approach," *Applied Economics*, pp. 1–16, 2021.
- [47] Z. Yu, R. Du, B. Guo et al., "Who should I invite for my party?: combining user preference and influence maximization for social events," in *proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing*, 2015, pp. 879–883, Osaka:ACM, 2015.
- [48] V. M. de Lira, C. Macdonald, I. Ounis, R. Perego, C. Renso, and V. C. Times, "Exploring social media for event attendance," in *in proceedings of the 9th IEEE/ACM international conference on advances in social networks analysis and mining*, pp. 447–450, Sydney:ACM, 2017.
- [49] X. Wang, L. Nie, X. Song, D. Zhang, and T. S. Chua, "Unifying virtual and physical worlds: learning toward local and global consistency," *ACM Transactions on Information Systems (TOIS)*, vol. 36, no. 1, pp. 1–26, 2018.
- [50] F. S. Rizi and M. Granitzer, "Predicting event attendance exploring social influence," in *in proceedings of the 34th ACM/SIGAPP symposium on applied computing*, pp. 2131–2134, Limassol:ACM, 2019.
- [51] S. R. K. Somayaji, M. Alazab, M. K. Manoj, A. Bucchiarone, C. L. Chowdhary, and T. R. Gadekallu, "A framework for prediction and storage of battery life in IoT devices using DNN and blockchain," in *In 2020 IEEE Globecom Workshops*, pp. 1–6, Taibei, 2020, December.
- [52] L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan, "Group formation in large social networks: membership, growth, and evolution," in *In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 44–54, Philadelphia:ACM, 2006.
- [53] M. Ye, X. Liu, and W.-C. Lee, "Exploring social influence for recommendation: a generative model approach," in *In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pp. 671–680, Portland:ACM, 2012.
- [54] M. Elahi, M. Braunhofer, T. Gurbanov, and F. Ricci, "User preference elicitation, rating sparsity and cold start," in *Collaborative Recommendations: Algorithms, Practical Challenges and Applications*, pp. 253–294, World Scientific Publishing Company, Singapore, 2019.
- [55] T. Chung, J. Han, D. Choi, T. T. Kwon, J. Y. Rha, and H. Kim, "Privacy leakage in event-based social networks: a meetup case study," *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), pp. 1–22, 2017.
- [56] C. Dong and B. Zhou, "Privacy inference analysis on event-based social networks," in *In Proceedings of the 3rd Multidisciplinary International Social Networks Conference on Social Informatics*, pp. 421–438, New Jersey :ACM, 2016.
- [57] K. Dou, B. Guo, and L. Kuang, "A privacy-preserving multimedia recommendation in the context of social network based on weighted noise injection," *Multimedia Tools and Applications*, vol. 78, no. 19, pp. 26907–26926, 2019.