IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 26, 2017, accepted November 24, 2017, date of publication November 29, 2017,
date of current version December 22, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2778293

Personalized Attraction Recommendation System for Tourists Through Check-In Data

K. KESORN, W. JURAPHANTHONG, AND A. SALAIWARAKUL
Computer Science and Information Technology Department, Science Faculty, Naresuan University, Phitsanulok 65000, Thailand

Corresponding author: K. Kesorn (kraisakk@nu.ac.th)

This research was supported in part by Naresuan University, Thailand, under Grant R2559C228.

ABSTRACT Online social networks now play a prominent role in our daily lives and our decisions and behaviors in many areas. Of particular interest here is the application of social network data to give users access to tourist information. There is a growing need for information on tourism and tourist activities to satisfy user queries in this domain. Social networks, such as Facebook, Twitter, and Foursquare, among others, store substantial volumes of check-in data, which are a valuable resource for recommending tourism attractions. However, using Facebook check-in data has rarely been considered in conventional recommendation systems (RSs). This presents not only a new research challenge for the computer science and information technology fields but also an interesting opportunity for the tourism industry: knowing what kind of attractions tourists are interested in and how to acquire their user preferences without adding tasks to users of an RS. We propose a tourism RS that is based on its recommendations on data dynamically aggregated and extrapolated from the Facebook check-in data. In addition, the so-called ''cold-start'' problem has been resolved by using users' Friends' check-in data to analyze ongoing Facebook activity and update user profiles in the system. Most Facebook users have a well-extended list of Friends. Consequently, the proposed system can dynamically learn user behavior and appropriately adapt recommendations. This paper demonstrate the usefulness of the data available on Facebook through the example studies involving attraction recommendations, resolving the cold-start problem, and adapting the user model to improve recommendation quality in the tourism domain.

INDEX TERMS Tourism, social network, cold-start problem, recommendation system, personalization.

I. INTRODUCTION

It is often essential for travelers to consult with experts, natives, or friends about which tourism attractions to visit at their desired destinations (e.g., where to go, where to stay, and how to get there, Customs and Immigration rules, warnings, and so on). Exchanging travel-related information between travelers using the Internet and online social networks has become more important and effective. People have access to tremendous amounts of travel information, but finding and/or identifying the most relevant information is difficult. As such, a recommendation system (RS) is a key technology for this problem. The high availability of information on online social networks has huge benefits for the tourism domain.

An RS is an information filtering system that analyzes user preferences and adapts its functions to individual users [1] and finding user interests or preferences is therefore a key process of an RS. RS-based techniques have been applied in diverse applications, including movies [2], music [3], news [4], books [5], search queries [6], restaurants [7], financial services [8], life insurance [9], personal activities such as online dating [10], and Twitter followers [11]. To the best of our knowledge, there are a few RSs in the tourism domain. In those RSs that have been studied, the researchers have paid special attention to the inclusion of social network functions that allow tourists to share some information [12]. However, almost all of these RSs have ignored the opportunity to exploit data from social network services for personalized recommendations.

The source of user data in conventional RSs comes from two approaches: 1) explicitly, by directly asking the user for individual information by allowing them to fill out forms or answer some questions on the Web site; and 2) implicitly, by gathering user information from various on-line sources without interrupting the user's on-line activities. The explicit approach can elicit more accurate data from users, but it also adds extra tasks to them. However, many users usually do not want to give their information to RSs because of their privacy concerns. Thus, the second method is preferred and deployed by some intelligent RSs. The main challenge of the implicit method is how to collect user information without interrupting them. Typically, the implicit approach collects user information by inference from the users' behaviors and actions in interacting with the RS. However, the accuracy of user preferences extrapolation is one of major issues of the implicit method.

The rapid growth of online social networks such as Facebook, Twitter, Foursquare, and Instagram has led to the availability of massive amounts of user data and provide much more detailed knowledge about users and attractions through the analysis of collected data [13]. Data from social networks, such as tags, social bookmarks, check-ins, likes, photos, videos, and comments, are useful information for automatically analyzing and finding user interests without adding extra tasks to users. For example, a music RS using social bookmarks has been introduced by Firan *et al.* [3] and personalized news RS using textual information from Twitter has been presented by Abel *et al.* [4]. RSs are also applied to tourism to recommend specific domains [16], hotels, restaurants, travel packages [17], itinerary plans [18], and navigational planning services to travelers [19], and for urban point-of-interest recommendation [14]. However, extracting useful information and exploiting these data remains a great challenge for researchers in several research fields, such as data mining [14], RSs [4], and forecasting [15].

Insufficient information creates the ''cold-start'' problems [20]. There are two types of cold-start problems [21]: 1) cold-start items and 2) cold-start users. The cold-start item problem, also known as the ''new item problem'', results because there are insufficient previously submitted ratings about items or products that are available to be recommended to users. The cold-start user problem, also known as the ''new user problem'', is caused by the entry of a new user who is therefore as yet unknown to the RS. Lack of previous data and a history of access means RSs are unable to make recommendations based on accessing experience to these new users [22].

An RS therefore needs to find user-related information from other sources, which would usually be other users somehow related to the new user. In this research, we hypothesize that a possible solution for the cold-start user problem is that information from the user's designated Friends on social networks could be used to represent the user's interests. The implicit relationships among users and their Friends can be considered through their shared interactions.

There are many user activities, e.g., comments, postings, Likes and so on that represent a rich source of knowledge about a user's implicit relationships with Friends, even when that is not explicitly declared by users, but can be derived from an analysis of their social network interactions. Thus, we can deploy information about a user's Facebook Friends to solve the cold-start user problem. As a result, social networking information is used to not only solve the explicit data-acquisition problem, but also the cold-start problem and, consequently, improve the prediction accuracy of the RS. There are, however, limits to the usefulness of much of the data available in this way.

Social media tags, also known as the hashtags (identified with the symbol ''#''), for example, which are widely used, are generally described as issues or topics of interest to particular users on social media sites. Hashtags make topics and conversations easier to find and to follow as these will be arranged by category. However, advertisers and marketers using social media sites are the main beneficiaries [23]. Research by Twitter [24] shows that the use of hashtags in a post can lead to higher user engagement from individuals and 50% higher engagement with referenced brands. While they may be useful for determining user interests, most hashtags are considered emotional expressions of user interests i.e. #love, #sad, #happy and are therefore usually unrelated to tourism attractions and are rarely used to analyze user interests in the tourism domain. Similarly, messages posted on Facebook, which are the major data types on Facebook, have little relevance to tourism, being, typically, more about users' personal opinions of someone or something unrelated to the tourism domain.

A further problem is that these data are always generated by individuals and almost inevitably contain diverse synonyms and polysemes used to describe users personal circumstances or are unlabeled photos (the ubiquitous 'selfie'), and certainly do not always refer to experiences, events or locations that might be helpful in the 'tourism' domain. Therefore, an intelligent natural language processing algorithm is required to extract the actual semantic information from the messages. Facebook users typically click the ''like'' button on the photos and comments by their Friends, and, as observed above, only some, very few, likes are relevant to tourism experiences. Check-in data is the most relevant type of data for the tourism industry because it indicates the places where users have visited or nominated as Favorite. In addition, as reported by Zocial's online analysis [25], check-in is the second most shared data between all Facebook users (Figure 1). Given these trends, it becomes more and more important to investigate the benefits of social networks such as Facebook in a tourism context.

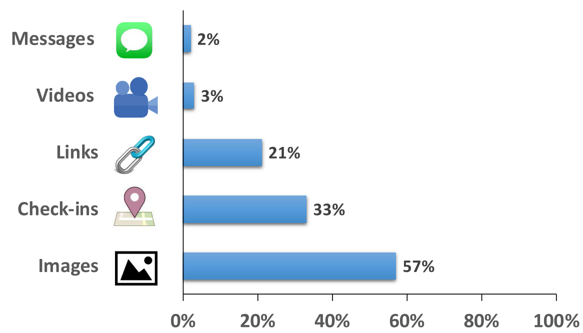

FIGURE 1. Most shared data between Facebook users in descending order.

Apart from the cold-start problems, most frameworks that we have identified ignore the problem of users changing their preferences over time. Those frameworks that we investigated provide a static mechanism to detect what the user is interested in at that moment of time and the quality of the recommendation is reduced with the passage of time bringing changes to the data. These frameworks also ignore the need to identify a suitable data sample size for user interest analysis: data overload can interfere with the RS's ability to match attractions to user interests, which affects the recommendation quality.

Hence, the proposed approach has to effectively overcome the problems of cold-start user, data overload, and adaptive user profile. In our study, we developed a personalized tourism information service (PTIS) framework to dynamically recommend attractions based on user interests. The framework has several specific features, including personalized attraction recommendations for tourists, solving the cold-start user problem using Facebook Friends' data, using appropriate user data quantity, and dynamically updating user profiles in the tourism domain.

Social networks offer opportunities to gather user data to aggregate and analyze for individual preferences to find their travel interests. Unlike other social networks, Facebook provides check-in data with direct relevance to the tourism domain. The vast amount of Facebook data, especially check-in data, makes Facebook a valuable source of personal information that is essential for the development of RSs for the tourism domain. This leads to our first hypothesis.

Hypothesis 1: Facebook check-in data can be used for personalized preference analysis in the tourism domain to make attraction recommendations.

Our second hypothesis considers the ability of an RS to overcome the cold-start problem by utilizing information from alternative external sources that can be substitute for, or complement, missing data to facilitate accurate recommendations, as discussed in [62].

Hypothesis 2: Check-in data from close Friends in social networks can overcome the cold-start problem.

Our third hypothesis is based on the situation of users changing their behaviors and their interests, and commenting on that in posts, and with Friends. User behavior in this context means users clicking on, searching for, and reading information differently than before, and more frequently on one topic than on others. These actions can be used as implicit feedback information to adjust the user model and enhance the recommendation efficiency.

Hypothesis 3: User feedback information can be effectively used by the system to adjust the user preference model and recommend attractions to users.

This paper introduces the PTIS framework that we developed. Ease of use, minimal required interaction and user responses were considered essential, and ways to achieve this were introduced in our model. As indicated in [12], users can be overwhelmed by complex and time-consuming interactions.

The main contributions of this research are, first, a novel technique that uses Facebook data to extrapolate user interests in tourism attractions. In contrast to conventional RSs, we also use Facebook check-in data to recommend attractions for users. This serves as an implicit user-data acquisition approach. Second, Facebook Friends' check-in data are also aggregated and exploited for personalized attraction recommendations, which makes the system more robust against the cold-start user problem using only a single type of data. Finally, we also proposed and evaluated a technique that can dynamically capture contemporary user-interest data, which is highly changeable, and flexibly update a user's profile by analyzing Facebook user actions. This allows the RS to learn user behaviors, extrapolate their interests, and adaptively recommend appropriate attractions.

The remainder of this paper is organized as follows: Section II presents the background and related work in the field. Section III describes our proposed technique to analyze user interests, user model construction, and attraction recommendations. Section IV discusses our experimental results. Finally, section V summarizes our key contributions, limitations, and further work.

II. BACKGROUND AND RELATED WORKS

In recent decades, many RSs have been presented to increase product-selling opportunities, facilitate product search, and assist in planning, including e-tourism [16], travel planning [18], personal travel assistance [26], traveling information [27], [28], and personalized search [6]. Several researchers used social network information to identify user interests in various ways. The main advantage of social networks is that they contain useful personal data about users. This information is used for personalized products and services in RSs.

A. SOCIAL NETWORK MESSAGES, TAGS, AND BOOKMARKS

This section surveys related state-of-the-art frameworks that use different data available from social network systems, such as messages, tags, and bookmarks. In recent years, users have begun to share more photos than text messages [29]. This visual information contains unstructured data, leading to the semantic gap problem [30], whereby a given platform cannot express the actual meaning of visual contents using low-level features. However, photos are usually accompanied by text message captions expressing more particular details about the visual information. These text messages are usually generated by humans; therefore, they are useful for detecting user interests in numerous ways.

Twitter is an online social network that allows users to share news or opinions using short messages called "tweets." Twitter users can post text messages, photos, or videos on their own or Friends' timelines. However, the messages posted on Twitter may not contain useful keywords related to visual contents. For example, a picture of the moon could have "good night" as a caption, which imperfectly represents the content in the picture. Several studies made an effort to resolve this problem. Poslad and Kesorn [31] proposed a method to separate tweets into two types, News and Chat, to assist users in assessing the reliability of tweets for further use. Abel *et al.* [4] identified topics and entities (e.g., persons, events, or products) mentioned in tweets for news recommendation. Castillo *et al.* [32] and Wasim *et al.*[33] used WordNet and Wikipedia to map entities in interest categories that will be used to define topic profiles. Lim and Datta [34] presented an approach for user-interest detection using tweets together with Wikipedia. The main weakness of this strategy is that an incorrect mapping problem could occur; e.g., mapping ''Apple'' (the fruit) to ''Apple'' (the technology company). Another disadvantage of this method is that it is time consuming to compute user interests, which is an impractical process in an on-line situation requiring immediate, or at least fast, response, leading users to potentially reject this system.

Many social network services create a tagging system for categorizing posted messages to allow users to more easily find the desired contents. The social bookmark is a service available to users who want to bookmark some Web pages and share them with their Friends or other users. The tag is attached to the bookmark. For example, an Internet surfer can bookmark tourism Web sites and add the ''hillside'' tag using the social bookmark service. Hence, the different types of tags used can represent different user preferences. Additionally, different tags on the same social bookmark services are likely to have semantic relationships. For example, Kumar *et al.* [6] used a social bookmark service to build a user profile (or interest profile) for personalized searches. An interest profile is constructed from the tags annotated by a community of users to Internet resources of interest. The interest profile is then used to assist search engines to generate a set of personalized search results. Michelson and Macskassy [35] proposed the add-a-tag approach to construct a user-interest model from social bookmark tags using Delicious.com services. Likewise, Firan *et al.* [3] proposed a tag-based interest model from tags on Last.fm for music recommendations. However, the drawback of these two approaches is they cannot support personal recommendations because the model generated used all user preferences with the so-called ''global tag.'' Michlmayr *et al.* [36] proposed an approach to use an individual tag for personalized recommendations. This solution explores the similarity between tags and users. Tags that are used more often can be considered as a user-interest topic. Durao and Dolog [37] deployed social tagging to provide suggestions about interesting communities that users may want to join. Kim and Saddik [38] also proposed a personalized Web site RS based on social bookmark tags. The RS can recommend the most relevant Web pages for individual users based on the users' profiles using Open Directory Project as an external knowledge data source. However, the proposed system relies on a simple string-matching technique, which provides low matching power to find the most relevant Web pages for user interests.

Several researchers have tried to exploit data such as comments, photos, likes, check-ins (of particular interest in our work) and relationships of Friends, from various social networks. Davoodi and Fatemi [39] introduced a framework that exploits check-in data from location-based social networks (LBSNs) for location recommendation. The algorithms presented in their research are based on four factors: 1) past user behavior, 2) the location of each venue, 3) the relationships among users in the social network, and 4) the similarity between users. The experiments showed promising results that can significantly improve recommendation performance.

In a similar way, Wang *et al.* [40] extracted location information from Foursquare (a New York mobile startup primarily for showing location information of other users). These authors proposed the use of location information from social networks such as Foursquare for identifying relationships among users. They termed this Friend-based collaborative filtering (FCF), and the location information and relationship information are then later integrated with collaborative ratings of places entered into the system by social Friends. In addition, geospatial characteristics of places previously observed and identified are also deployed to enhance the performance of FCF. Their experimental results demonstrated that their FCF approach demonstrates equivalent recommendation effectiveness when compared against state-of-the-art algorithms previously published.

Ye *et al.* [41] presented a novel framework to reduce the gap between tourist requirements and information extracted from dominant tourism resources using cross-region community matching techniques. Local interests of people from different geographical regions are captured and heterogeneous relationships among users, places, and times are taken into consideration. Information from different communities from across different regions is then correlated and venue recommendations generated and presented to the tourists via cross-region community matching. The experimental results of Zhao's work outperformed existing frameworks.

Check-in data from an online social network (Gowalla) is also used for personalized venue recommendations by Zhao *et al.* [42]. The collected data is used to analyze the possibilities and requirements for location spot recommendations by utilizing a collaborative filtering scheme. The recommendation performance of this work is superior to others when statistically measured by root mean square error (RMSE) and mean absolute error (MAE).

Berjani and Strufe [43] exploited multi-modality travel information such as comments, pictures and rating scores for personalized tourism purposes, whereas single-modality data had been used by other researchers previously for the same purpose. In addition, context information, such as the user's location, was used to refine the recommendation to better suit that user's preference. The experimental results demonstrated that the method achieved promising performance in terms of both effectiveness and efficiency.

Association rules, complemented by the classification method, were effectively applied to predict user preferences [45]. Their proposed method is mainly based on fuzzy logic and associative classification and is less sensitive to the data sparsity (cold start) problem. Their experimental results demonstrated that their method can overcome the limitations of existing recommendation systems and increases recommendation quality. Instead of a personalized recommendation system, Lucas *et al.* [45] recently proposed a social group recommendation approach in the tourism domain which aims to satisfy a group of users as a whole. In that research, a group profile rather than an individual model is constructed to store user preferences as well as the social relationships among members of a group. To generate recommendations, three techniques are combined to compensate each other: collaborative, content-based, and demographic filtering. The method can recommend tourist attractions to both individual and groups of users. The proposed method achieved more accurate recommendations than the classic approaches.

B. USER PROFILE REPRESENTATIONS

In social networks like Facebook, users provide their individual data (e.g., name, workplace, email, or affiliations) when setting up their profiles during the registration process. In the RS, a user model typically stores user preferences derived from individual data and the model can be represented using various data structures. User profiles allow the RS to make personalized recommendations, and they constitute a key component of the system. This research represents the user model in the form of a relational database (RDB). The main advantages of this are the stability, consistency, integrity, ease of maintenance (insert, update, or delete), and better security offered by an RDB. The user modelsare designed to store several types of data of users e.g., personal data, interest in attractios, feedback information, and interactions between users and PTIS).

C. RECOMMENDATION TECHNIQUES

The recommendation process aims at suggesting attractions to users. Various approaches have been deployed in an RS, including: 1) Content-based [47], 2) Collaborative [48], and 3) Hybrid [21] methods. The most popular approaches are Content-based and Collaborative. Content-based filtering is recommended based on the analysis of the user's previous actions whereas collaborative filtering refers to recommending items based on information from other users.

1) CONTENT-BASED FILTERING

Content-based filtering suggests activities, events, or services to a user by matching the users' interests with the information about these things. Travelocity.com [16], for example, deploys the content-based approach, in which users specify his or her requirements or interests to the system. The system will match the users' interests with the products available from the destinations. The matching technique used by both systems syntactically compares product and user profile attributes. However, this technique cannot represent user interests precisely. For example, a user may refer to products on behalf of others, not for themselves; perhaps it may be a gift that may suit a friend or family member; this interaction would provide incorrect information about the user to the system, which leads to poor recommendation performance. Nonetheless, this problem is minimized in RSs for the tourism domain because almost all users usually purchase services for themselves, not for somebody else, as would happen in other RSs such as e-retailers for books or movies [2]. Another problem with the content-based approach is the interaction with a new (cold-start) user where the system usually does not have adequate user information to deliver recommendations resulting in poor or imprecise suggestions being delivered to the new user. However, a newer collaborative approach can resolve such a problem by constructing a new user profile from information derived from the user's Friends on social media (collaborating).

2) COLLABORATIVE FILTERING

The new collaborative filtering approach is the process of suggesting products and services using a technique involving collaboration among multiple users existing in the system. Similar to content-based approaches, new users will obviously not have any historical preferences, meaning that this collaborative filtering approach is not able to generate any recommendations, except for generalized popularity-based approaches. TripleHop Technologies [16] uses a statistical computation of past queries to predict user similarities. The MAIS project [49] uses a collaborative filtering approach to identify similar users and recommend items based on the information extracted from those similar users. However, if the system does not have adequate user interactions, it is difficult to find user neighbors. This is because user interactions allow the system to learn about a user and identify the user's neighbors. The main drawback of this method is that the selected neighbors may be controversial, which may result in diverse recommendations and a lack of specialized suggestions.

Another disadvantage of this approach is that it recommends products based on user ratings, which creates a cold-start item problem that result in inaccuracy recommendations [50]. A new item may not be recommended to any users after it is added to the system because it does not yet have any user ratings. However, the cold-start item problem is out of the scope of this paper and will not be discussed further. Therefore, the term "cold-start problem" used in this paper always refers to the cold-start user problem.

3) HYBRID FILTERING

The hybrid filtering approach has been proposed to eliminate the limitations of content-based and collaborative filtering approaches such as the framework in Pazzani [51]. In this work, the content-based approach is used to construct user profiles whereas the collaborative technique is deployed to compute the similarities of user interests. Schiaffino and Amandi [17] proposed an expert software agent, Traveler, that combines content-based and collaborative approaches with demographic information to recommend tour packages to tourists. Burke [52] presented an RS that uses a knowledge-based method that can recommend restaurants in a different city from where the user lives. Maw *et al.* [53] used rule-based and collaborative filtering approaches to overcome the sparsity problems that can reduce the consideration set of items. Table 1 compares some of the major problems of these three different recommendation techniques.

TABLE 1. Disadvantages of the recommendation techniques.

Problems	Recommendation techniques		
	Content-based	Collaborative-based	Hybrid
1. Cold-start user.	×	×	✓
2. Data overload.	✓	×	×
3. Flexible user profile.	×	×	×

III. PTIS FRAMEWORK

To develop the personalized RS for tourists, we present the PTIS framework that recommends attractions to tourists based on using Facebook check-in data. Although using location or event check-in data to recommend an attraction to the tourist is not a new method, we consider that both the processes of the recommendation system, and the user interest extraction method developed in our research, are significantly different from existing state-of-the art frameworks. The PTIS framework was briefly introduced in [54] and different from this work which is mainly focused on the user preference analysis module and the user model construction based on Facebook data. The three major components of the PTIS are illustrated in Figure 2, and are explained below.

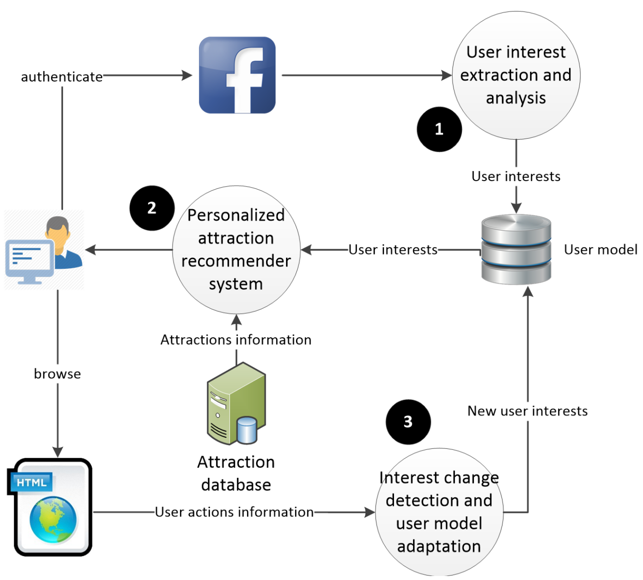

FIGURE 2. Overview of the PTIS framework, which has three main components: 1) user-interest data extraction and analysis; 2) personalized attraction recommendations; and 3) detection of user-interest changes and adaption of the user model.

A. EXTRACTING USER PREFERENCES

This section describes the method for computing user interests and the methodology for attraction recommendations. This process extracts Facebook user-interest information using the Facebook graph API [55] for the data extraction. First, to extract Facebook data, users must be authenticated in the system to retrieve their private data e.g. check-in and Friends' information. Second, the extracted data are processed to derive the user's interests. The system then generates a user-preference model represented in an RDB structure. For any new user of our system, a user profile will be constructed during this step which is skipped this step for existing users. For non-Facebook members, our system cannot perform personalized recommendations for them because it does not have their user information, but it is still able to recommend some attractions based on popularity scores measured by the number of Facebook check-ins. Table 2 demonstrates an example of Facebook check-in data extracted using the Facebook graph API. This information will be used for further user-interest analyses.

TABLE 2. Example of extracted Facebook check-in data.

Extracted data	Descriptions
{"place": {	
"id": "196230743801741",	Place identification number
"name": "Wat pra sri rattana mahathat",	Place name
"location": {	Location of place
"latitude": 17.429238993972,	Latitude
"longitude": 99.811871548508,	Longitude
"street": "Srisuchanalai" }	Address
}, "id": "750723874966970",	Message id/number
"created_time": "2013-04-13T09:21:03+0000"}	Check-in time

1) USER-INTEREST MATRIX

The user-interest matrix shows the level of interest in the attraction categories using Facebook check-in data. Our assumption was that the number of places visited (check-ins) in a particular category indicates the level of interest in attractions in that category. Therefore, we apply the item frequency (*if*) scheme [56] to measure user interest levels and represent them in a user model. A weight is assigned to each attraction in a category based on the number of check-ins at that attraction. The item frequency (*if*) technique is deployed because some other candidate techniques (e.g., *idf* and *tf-idf*) are not applicable to PTIS. For example, *tf-idf* will adjust for the fact that some attractions that are highly checked-in by a user are more generally applicable and do not impact the weighting given to that attraction. In other words, those techniques do not represent the real level of user interest in an attraction. Therefore, the *if*-weighting scheme is preferred.

There are two important concepts that require definition: the level of interest of a user and Friend interactions of a user.

Definition 1: The interest level (I) of user (u) shows their preference levels in each attraction category:

$$I_u(c) = \{(c, l)|c \in C \text{ and } l \in L, l \text{ is in the range } [0,1]\}$$

where l is the level of interest associated with category c for a given user u. l is the proportion of the number of attractions that a user has checked-in in category c out of the total number of attractions in the same category. l must be in the range 0 to 1. c are the members of attraction category C and can be grouped into the six main categories. For example, $I_u(c)$ = *{(historic, 0.44), (natural, 0.37), (cultural, 0.19)}* where historic, natural, and cultural are categories. $I_u(c)$ denotes the level of interest of users where the weight is normalized and the sum of all weights in $I_u(c)$ is equal to 1. A summed weight of 0 shows that a user is not interested in the category and a value of 1 indicates the highest level of interest in the category.

2) FRIEND INTERACTION MATRIX

This research adopts the Friend analysis algorithm, ay-fb-friend-rank [57], to separate close Friends from others. This algorithm is based on an EdgeRank technique [58] to identify close Friends. There are three main components of this algorithm: affinity score, edge weight, and time decay.

The affinity score is the interaction score. For example, John often writes on Mary's wall. John has a very high affinity score with Mary. The ay-fb-friend-rank calculates an affinity score based on: (1) explicit actions that users take (e.g., clicking, liking, commenting, tagging, sharing, and mutual Friends); (2) the proximity of the person who took the action in relation to a user; and (3) how long ago the action was taken.

Edge weight gives different weights to different actions. Every action that a user takes creates an edge, and each of those edges, except for clicks, creates a potential story. For example, comments have higher edge weight than likes. Time decay adjusts the score of the story. As a story gets older, it loses points because it is "old news." When a user logs into Facebook, their newsfeed is populated with edges that have the highest score at that very moment in time. Based on the idea of those algorithms, Friend interactions can be defined as:

Definition 2: Friend interactions (F) of user (u) that show the level of interaction between a user and their Friends. F is a set of Friend-interaction score pairs:

$$F(u) = \{(f, s)|f \in \mathbb{R} \text{ and } s \text{ is in the rang } [0, 1]\}$$

where f is a Friend in Facebook ($\Re$) and s is the interaction score, which varies in the range 0 to 1, where 0 means that users have no interaction between them and 1 shows that they have a high level of interaction.

3) USER-INTEREST ANALYSIS

There are two scenarios for extracting Facebook check-in data to detect user preferences: adequate and inadequate information for PTIS. Adequate information refers to users with check-in data greater than the threshold, which is chosen experimentally to find user interests while inadequate information is the number of check-ins of a user that are less than the threshold. In the latter case, the system needs to retrieve data from a user's Friends from Facebook instead of the user's own data. To analyze user interest based on personal data from Facebook, the computation scheme of user interests can be defined as (1), where $I_u(c)$ is an interest level in a category c of the user u, n_c is the number of check-ins for a category c and $I_u(c)$ is normalized in the range of 0 and 1.

$$I_u(c) = \frac{n_c}{\sum_{i=1}^{6} n_i}, \quad 1 \leq c \leq 6 \tag{1}$$

To analyze user interests based on information from Facebook Friends, the computational scheme from (1) can be modified as shown in (2). The extrapolated interest level of user u can be calculated from an aggregation of the user's Facebook Friends' check-in data. $F_i(u)$ is the level of interaction between users and Friends, $I_{fi}(c)$ is the interest level of each category of the user's Friends, n refers to numbers of Friends of the user on Facebook, and $n \in N$ where N is the set of close Friends on Facebook.

$$I_u(c) = \frac{\sum_{i=1}^{n} F_i(u) \times I_{fi}(c)}{\sum_{i=1}^{n} F_i(u)}, \quad i \leq n \tag{2}$$

A user-interest analysis algorithm is shown in Algorithm 1 (described below). User interests, once obtained, are then used to construct a user model for attraction recommendations.

Algorithm 1 User Interest Analysis

Input: All check-ins data of a user.
Output: Interest level $I_u(c)$ in each category.
1: Remove duplicated check-ins from the same day.
2: Remove any check-ins that are not attractions.
3: Classify check-ins into six categories.
4: **If** all classified check-ins of a user $\geq$ threshold
5: Compute level of interest $I_u(c)$ using Eq. (1)
6: **Else**
7: Identify close Friends.
8: Get all Facebook check-ins of close Friends
9: **If** numbers of check-ins of close Friend $\geq$ threshold
10: Compute level of interest $I_{fi}(c)$ using Eq. (2)
11: **Else**
12: Define level of interest $I_u(c)$ with average value from popularity.
13: **Return** $I_u(c)$

B. ATTRACTION RECOMMENDATIONS

The recommendation process was employed to suggest the attractions that should be most appealing to a user. Once user interests are known, the PTIS needs to match user interests with attractions based on mathematical computation. If user check-in data is inadequate, the system will extract information from close Friends of the user using the ay-fb-friend-rank [57] algorithm. The matched attractions have to be ranked by descending scores. Here, our proposed approach is applied to recommend attractions based on the analysis of user check-ins by matching the characteristics of the attractions with the user characteristics [16]. This process consists of four main steps: category proportion, context-aware recommendation, attraction weighting, and attraction ranking.

1) CATEGORY PROPORTION

Human beings are usually interested in many things at the same time. Likewise, a traveler is usually interested in various attractions in different categories. Therefore, not only will the attractions in the category with the highest interest level be recommended, but also other attractions in other categories will be suggested. Category proportion performs such a task to find the ratio of places in different categories to suggest to a user. Category proportion can be computed as shown in (3).

$$R(c) = \frac{I_u(c)}{\sum_{k=1}^{6} I(k)}, \quad 1 \leq c \leq 6 \tag{3}$$

where $R(c)$ is the proportion of places in each category suggested to a user based on their interest level or $I_u(c)$ in (*1*) divided by the interest level of all categories, $I(k)$.

2) CONTEXT AWARENESS RECOMMENDATIONS

Context is any information that can be used to characterize the situation of an entity [59]. Context in this work can be any information regarding the tourism situation in which a user experiences an attraction (e.g., location, time, or weather). Among these, time, information can be considered as important and useful data in RSs, particularly in the tourism domain [60]. Time-aware recommendations have received much attention and have been proven to increase recommendation performance and facilitate tracking the evolution of user models [61]. Here, we incorporate time information within the recommendation process to improve the RS accuracy by having information available about the recommended time for visiting the attraction. For example, recommending beautiful beaches in monsoon season to a traveler is questionable, although those beaches may have very high interest scores in the traveler's user profile. Therefore, our system may be considered as a time-aware RS and should not recommend any attractions that are out of context for users. This will provide greater user trust in the RS recommendations. Therefore, Definition 1 can be extended to Definition 3 by incorporating the time dimension into the interest level (I) as shown below.

Definition 3: The interest level (I) of user (u) can be modified by incorporating time dimension (t) into Definition 1 which can then be formulated as follows:

$$I_u(c) = \{(c, l, t)|c \in C, l \text{ is in the range [0,1], and t} \in \text{T}\}$$

where t is the target time information (discrete factor) stated as the month specified by the user. The use of the time context information here allows the system to provide differentiated recommendations depending on the target time of a user. This factor affects the ranking of the final results. Thus, the system can recommend different places to a user if his or her preferred visiting time is different.

3) ATTRACTION WEIGHTING

This process weights all attractions in each category for selecting the top $R(c)$ places to the user. The weighting scheme has three variables:

1) Popularity of place (P): Tourists usually want to visit the popular places or iconic places at the destinations. This information is obtained by extracting the number of Facebook check-ins. The higher number of Facebook check-ins a destination has, the more popular it is.
2) Places that Friends have visited (F): Tourists are also interested in places where their Friends have been. Therefore, this parameter cannot be ignored and may lead to significant improvements in recommendation accuracy.
3) Appropriate time to visit attractions (T): As discussed in the previous section, time information is also an important condition for making a decision to travel in the tourism domain.

Based on the reasons mentioned above, the attraction weighting uses a linear regression analysis model to consider all variables for the attraction recommendations as shown in (4).

$$W(p) = \alpha P(p) + \beta F(p) + \gamma T(p) \tag{4}$$

where $W(p)$ is the attraction weight, α is the popularity weight, β is the weight of the user's Facebook Friends' check-ins at locations, and γ is the weight of the appropriate time for visiting the attractions. Popularity of place, $P(p)$, is measured by the number of Facebook likes and check-ins for a place. $n_{ch,p}$ is the number of Facebook check-ins at a place and $n_{li,p}$ is the number of Facebook likes for a place. Max is the highest number of Facebook check-ins for each attraction. We normalize the range of check-in and like variables to ensure that the data are not overloaded by each other in terms of distance measures [62] as shown in (5).

$$P(p) = \frac{n_{ch,p}}{\max(n_{ch,p})} + \frac{n_{li,p}}{\max(n_{li,p})} \tag{5}$$

Tourists are often interested in places that their Friends have visited. Therefore, we also consider this parameter. Places visited by Friends, $F(p)$, is measured by the Facebook check-ins by close Friends as shown in (6). F_i is the level of interaction between users and close Friends i^{th}, $C_i = 1$ if Friend i^{th} has checked-in on Facebook at this place p and $C_i = 0$ otherwise.

$$F(p) = \frac{\sum_{i=1}^{n} F_i C_i}{\sum_{i=1}^{n} F_i} \tag{6}$$

Almost all attractions have an appropriate time to visit, which is known as high season or the popular season. Therefore, the time parameter will assist the system in recommending the right place at the right time. We defined rules for the time values as follows:

$T(p) = 1$ indicates that this is the best time to visit the attraction, as recommended by the Tourism Authority of Thailand.

$T(p) = 0.5$ indicates that this is an acceptable (reasonable, convenient) to visit the attraction.

$T(p) = 0.2$ indicates that this is an acceptable, but does carry caveats against travelling to this attraction.

$T(p) = -1$ indicates that this is not an appropriate or convenient time to visit the attraction, and tourists are recommended not to visit at this time.

Each rule is applied to (4) depending on the chosen time of a user.

4) ATTRACTION RANKING

This process measures the relevance of attractions to the user's interests that are identified in the user's profile. The recommended attractions are ranked by descending order and deliver a list of the top-N ranked attractions that the user may prefer. The attraction ranking is measured by (7). If $Rank\ (p_1)$ has a greater score than $Rank\ (p_2)$, this indicates that p_1 has more relevance to user interest than p_2.

$$Rank(p) = \frac{R(c) \times W(p)}{N} \tag{7}$$

where N is the number of attraction recommendations.

C. DETECTING CHANGES IN USER INTEREST

Users change their interests over time depending on several factors i.e. social environment, seasons, Friends, news or age. These factors have a great influence on user interests. An obvious example is where travelers' interests change with the seasons, being interested in mountains and snow or ski resorts in winter while preferring beaches, tropical islands and hiking trails in summer. They are also influenced by special, often annual, events at specific times.

The PTIS monitors this change by evaluating user actions (both implicit and explicit) during their interaction with the system. Every action has a score and this data will be used to detect changes in user interests and adapt or update the user model generated by the first component in Figure 2. The process to detect changes in user interests manipulates the user preference model to show the current interests of a user in real-time using relevant feedback techniques. We applied the Rocchio algorithm [63] to this task. We separated user feedback information into two groups, positive and negative feedback, using the different actions by users during their interaction with the system. Positive feedback refers to the feedback about relevant attractions and negative feedback is feedback about irrelevant attractions. The proposed technique is presented in the following section.

Information feedback is generated by user interactions with the system. The action score and action confidence values [64] were deployed to consider the importance of attractions based on individual actions using the Rocchio algorithm to classify feedback actions into positive or negative feedback.

TABLE 3. Action scores and action confidence values of users' feedback actions.

Feedback Actions	*Action Score (S)*	*Action Confidence (C)*
1. Browse the detail of recommended item (Implicit)	1.0	0.2
2. Establish recommended item as interested place (Explicit)	1.0	0.5
3. Remove recommended item and require new item (Explicit)	–1.0	0.5

The action confidence value is calculated based on implicit and explicit actions. Implicit actions are the actions indirectly required from users. PTIS automatically collects this information from the users' browsing click-stream. Explicit actions are the actions directly required from users such as clicking on items when the recommendation list appears. Here, the explicit action, which directly contributes to the relevant feedback method, is given a greater confidence score than for the implicit action. Table 3 shows action scores and action confidence values with feedback behavior. All user actions in their interaction with the system are used to determine their action scores and confidence values. The user feedback actions related to the same place are measured by their adaptive weight of attraction and defined as $A(p)$ in (8), S_i is the action score of feedback action ith, C_i is the action confidence of feedback actions ith, and n_p is the number of all actions with place p. $A(p)$ presents changes in user interests for a particular place. The positive $A(p)$ shows that the user is more interested in that place, and the negative $A(p)$ shows that the user is less interested in that place.

$$A(p) = \frac{\sum_{i=1}^{n_p} S_i C_i}{n_p} \tag{8}$$

The Rocchio technique was also applied to the automatic adaption of user models to renew the initial recommendations by adapting to the changes in interest. The adaptive weight of place $A(p)$ is used to adjust the weight of place $W(p)$ and the interest level $I_u(c)$ in each attraction category by adding to the old one. The new weight of place ($W'(p)$) is defined as in (9), the new interest level in each attraction category ($I'_u(c)$) is defined as in (10), and α_c is defined as (11).

$$W'(p) = W(p) + A(p) \tag{9}$$

$$I'_u(c) = I_u(c) + \alpha_c A(p) \tag{10}$$

$$\alpha_c = \left[\log \frac{1}{I(c)}\right] / 10 \tag{11}$$

The algorithm for detecting changes in user interests is illustrated in Algorithm 2.

Algorithm 2 Detecting Changes in User Interest

Input: Action Score (S), Action Confidence (C), n_p
Output: $Rank(p)$
1 : Measure adaptive weight $A(p)$ using (8).
2 : Calculate the new weight $W'(p)$ using (9).
3 : Compute the new interest level $I'_u(c)$ using (10).
4 : Measure the new category proportion $R(c)$ using (3) with $I'_u(c)$.
5 : Measure the new attraction ranking $Rank(p)$ using (7) with $W'(p)$and new $R(c)$.
6 : Recommend the new attraction with the list of new $Rank(p)$

D. HANDLING CHECK-IN DATA INTEGRATION UNCERTAINTY

As previously discussed, it is not an easy task to group check-in places into the six categories that we created, given the hundreds of informal categories created by Facebook users. An additional difficulty arises from the freely created naming by users. Therefore errors arise from misspellings, inappropriate synonyms and heterogeneous terms that are used by different users. To resolve the misspelling problem those categories are automatically grouped based on the similarity of their names and keywords tagged by users using the Fuzzy string matching technique [56], which is a technique for finding strings that approximately match a pattern. Fuzzy string matching (FSM) is based on the measurement of distance between two strings. Distance means the number of characters separating the strings. In addition, Facebook freely allows users to create and name attraction categories, which has the inevitable outcome of duplication of categories with a variety of synonyms for the same attraction, and polysemes for different categories. To overcome these problems, PTIS exploits WordNet [65] as a dictionary which can be consulted to disambiguate category names by obtaining the top-level categories that can be deduced from their correspondence with the lexicographer file structure.

IV. PERFORMANCE EVALUATIONS AND DISCUSSION

To evaluate the PTIS, several experiments have been performed using a dataset extracted from participants with active Facebook accounts. We recruited participants by invitation. However, the ages and number of check-ins in specific age groups of participants in this experiment were not normally distributed as shown in Figure 3. Participants came from Naresuan University (NU), Thailand and the majority were aged between 18–32 years, whereas only 10% of all participants were older than 35 years. This is because elderly people are less likely to have Facebook accounts and lack information technology experience. A total of 520 volunteers were recruited and participated. All experiments were undertaken with the same group of participants. Figure 3(a-d) shows the distribution of the demographic features of the participants. As illustrated in Figure 3 (a), the highest number of participants was in the single marital status. The largest age group was between 18-22 years, as shown in Figure 3 (b). The majority of participants were studying Bachelor or Masters

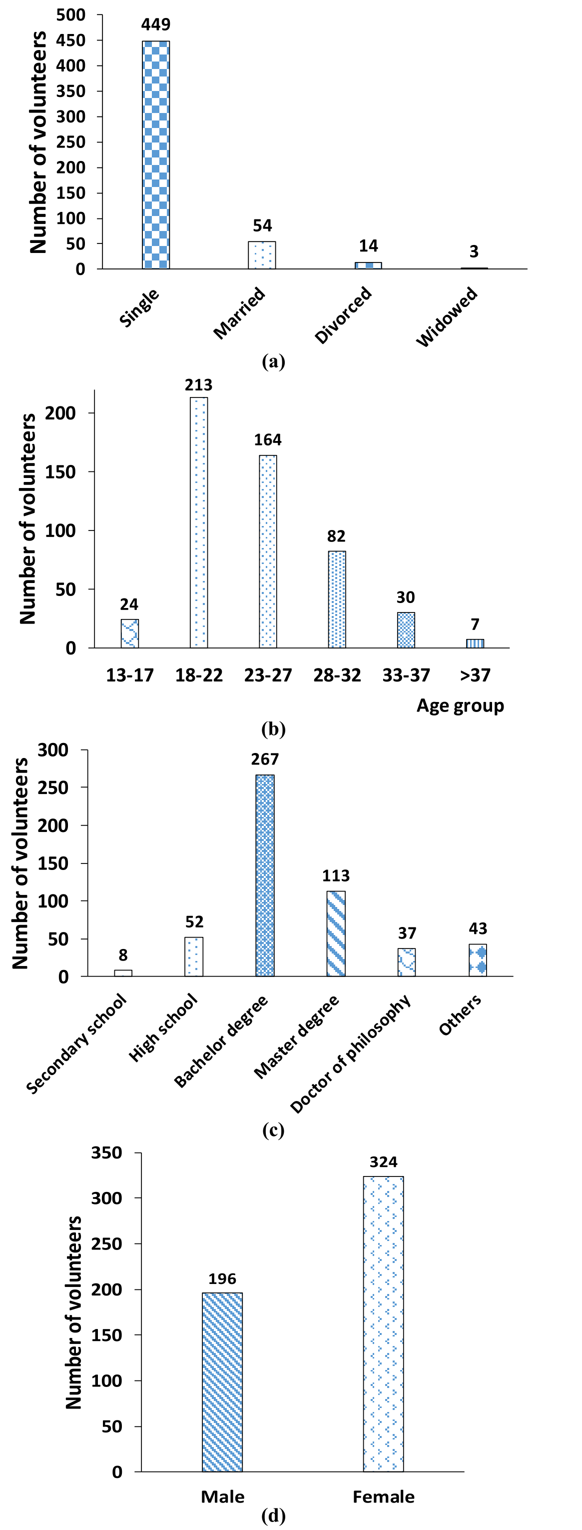

FIGURE 3. The distribution of selected demographic features with respect to the number of participants: (a) marital status, (b) age group, (c) education, and (d) gender.

degree (Figure 3 (c)) (thus the predominance of 18-22 year olds) and some participants were lecturers and researchers working in NU, thus the 37 PhDs. Figure 3 (d) shows 324 female participants and 196 male participants.

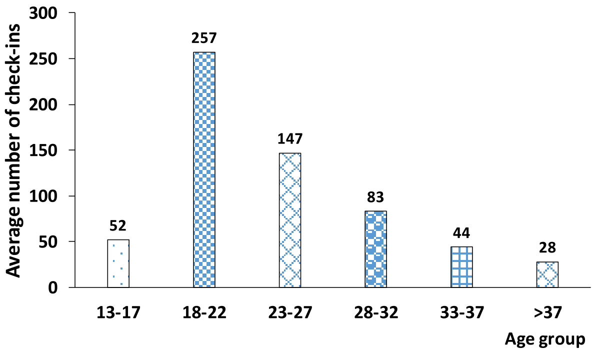

FIGURE 4. The histogram of the number of check-ins of users in specific age groups over seven years.

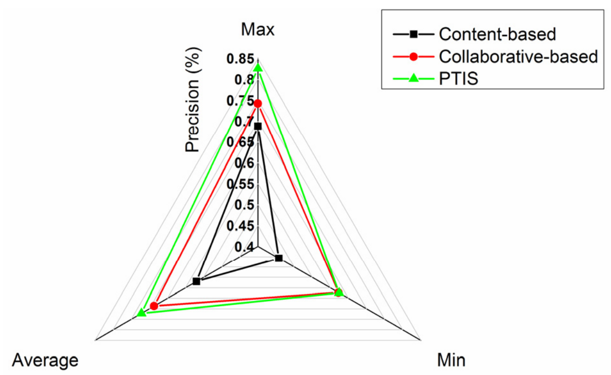

FIGURE 5. Average precision (%) comparison of the PTIS, content-based, and collaborative techniques.

Figure 4 depicts the histogram of the number of check-ins for various age range of users. The majority of participants had about 100-200 check-ins over the previous seven years. This is not a large number in the circumstances, probably due to these participants studying either in secondary school or in a Bachelor degree program and had not travelled extensively in that 7-year period. The Masters and Doctoral degree holders had travelled more and had therefore greater check-in numbers; more than 200 times in the seven years. Figure 5 does illustrate a fair and reasonable distribution of check-in behavior over age groups, where it is understood that over-37's for example, use Facebook significantly less than 18-32 year olds. However, it must be stated that the demographics of the participants are discussed here for interest only. The primary focus of our research was the effectiveness of the model that we developed and tested. It was the extent of data that we were likely to have available from the participants that was important. Even given that the sample of the participants were drawn mainly from students, and many of them were Bachelor degree students without a large number of check-ins recorded, we feel that the data was nonetheless extensive enough and variable enough to be valid.

It is also important to note that we did not compare the evaluations with alternative tourism RSs because these experiments were conducted on different groups of participants, with correspondingly different demographic features, cultures, and data sets. Therefore, it is difficult to compare our results with those from existing RSs. However, we theoretically compared the performance of the PTIS with those systems through the discussion of our experimental results. The experiments were conducted to measure the recommendation quality of the personalized recommendation method of our paper in comparison with conventional content-based and collaborative recommendations published elsewhere. The standard measures, average precision, and rank score are deployed to evaluate the recommendation efficiency of PTIS.

A. RECOMMENDATION PERFORMANCE EVALUATION

To evaluate the PTIS, we comp are the recommendation efficiency of the proposed system with some conventional techniques. The purpose of this evaluation was to study the effectiveness of using Facebook data for attraction recommendations compared with conventional methods. As only the top-n recommended attractions have to be chosen from all attractions, this unknown distribution influences the recommendation accuracy and user satisfaction in practice. The precision provides a direct evaluation of recommender accuracy [66], [67]. To compute the precision for each user (u), the attractions are ranked using the computation scheme in (7). The attraction is defined as relevant to a user if they find it appealing. Thus, the precision can be defined as the fraction of relevant attractions that are in the top-n of the ranking list, denoted by $N(n,u)$, from all recommended attractions, $N(u)$. In this study, $N(u)$ is 10 because the system will recommend only the top-10 relevant places to users. The precision can be aggregated from all users to obtain the average precision for the test dataset. Hence, the average precision is given by (12).

$$H(n,u) = \frac{\sum_u N(n,u)}{\sum_u N(u)} \tag{12}$$

It is vital to note that we do not evaluate the PTIS performance using recall because it is calculated from the fraction of $N(n,u)$ from all relevant attractions in the database, denoted by $R(n,u)$. This means that the recall value is always small because $R(n,u)$ is normally much larger than $N(n,u)$, which is fixed at 10 (top-10 highest score recommendation) for our system.

The top 10 recommendation performances using those approaches are shown in Figure 5. which shows that PTIS is superior to other methods, with 72.24% average precision. By contrast the content-based and collaborative approaches, yielded only 57.03% and 68.76% average precision. This may be explained by the fact that PTIS exploits individual Facebook user check-in data, which effectively represents their preferences and, consequently, PTIS has a maximum

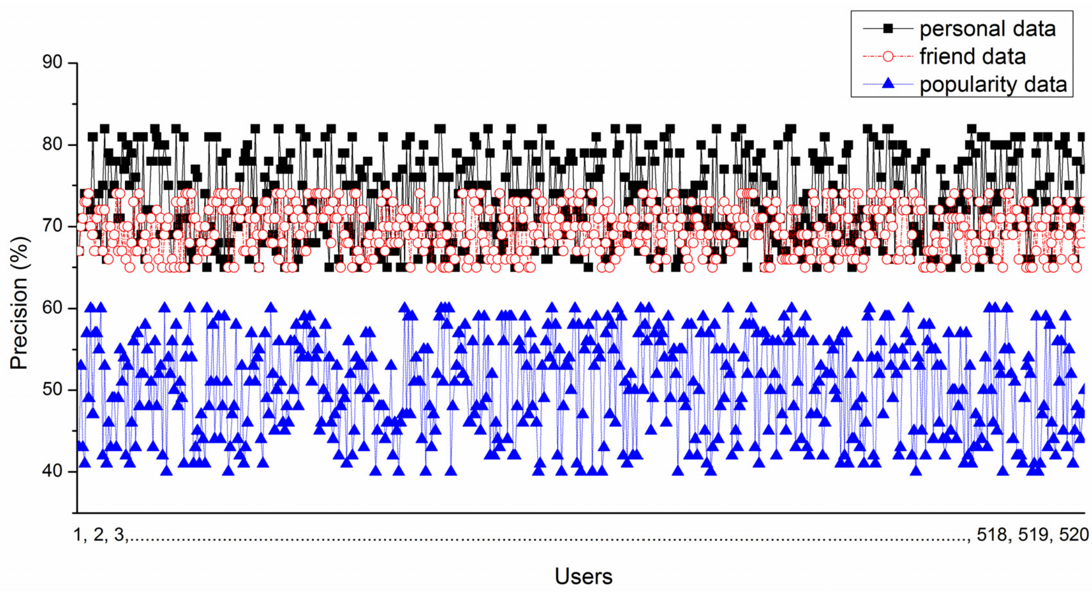

FIGURE 6. User interest prediction accuracy of 520 participants using three different data sources: personal, Friend, and popularity data.

of 82.51% precision. This enables PTIS to overcome the cold-start problem effectively, which is the principal reason for the poor performance of the content-based technique. As such, it is no great surprise that the content-based approach has a maximum precision of only 68.80%. The collaborative approach also deploys information from Facebook Friends (all Friends) of a user and this facilitates the system to find relevant attractions for users meaning that the cold-start problem is resolved. From our perspective, using Friends' information may not allow the collaborative technique to recommend attractions as accurately as PTIS because the aggregated data of all Friends could become noise which confuses the RS in matching attractions to user interests. Consequently, the prediction of interests by the collaborative approach is not as precise as the PTIS. As a result, the collaborative approach has a maximum precision of 74.23%, which is still lower than for PTIS.

TABLE 4. Statistical comparison between three approaches.

Methods	n	$\overline{x}$	*S.D.*	F	*p*-Value
1.Content-based	10	0.57^a	0.08		
2.Collaborative-based	10	0.69^b	0.04	14.8	< 0.001
3.PTIS	10	0.72^c	0.06		

Note: a, b, c indicate statistically significant difference.

In order to investigate statistical difference between those methods, the average precisions of all three approaches were compared using one-way ANOVA, followed by Bonferroni method for multiple comparisons (Table 4). The result illustrates that they are significantly different (F = 14.80, p-value < 0.001).

This study confirms our first Hypothesis that Facebook check-in data can be useful and used for personalized preference analyses and recommendations.

B. EVALUATION OF CHECK-IN DATA

This evaluation investigated the effectiveness of using check-in data to tackle the cold-start problem, that is Hypothesis 2, that users' Facebook check-in information can represent their preferences and overcome the cold-start problem. To evaluate this hypothesis we divided our analyses into three cases using personal, Friend, and popularity data. The so-called *personal data case* uses personal user information to represent their own interests. To test the cold-start problem, the check-in data of some participants was removed which forced the system to use data from others by aggregating data from Facebook Friends. This scenario is the second case in our experiment, the so-called *Friend data case*. The third case exploits the popularity data of attractions based on the number of Facebook check-ins. A higher number of check-ins indicates the greater popularity of the attraction. This is called the *popularity case* in our experiment. This number is gathered from the number of check-ins, likes, and shares of each attraction. Participants in the study manually ranked the top ten attractions in each category in which they were interested; this is called *true ranking*.

We compared the recommendation results generated by the system to those in the true ranking. As shown in Figure 6, the average ranking accuracy was up to 80.88% and 72.83% for the *personal data case* and *Friend data case*, respectively. These encouraging results show that when the amount of data is sparse, recommendation accuracy relying solely on Facebook Friend data is slightly lower than the results obtained from using personal data. This confirms that individual data has higher quality than Friend data to represent user interests. Therefore, both cases can significantly increase the precision of results compared with using popularity data, which achieved only 51.24% accuracy.

Besides using check-in data, some researchers proposed the use of social bookmarks [6], [39] and tags [3], [36]. It has been reported [68] that the precision of tag-based RSs usually is in the range of 10%–30%, which is significantly low precision compared with the results yielded by the PTIS. This low precision indicates that check-in data is valuable for increasing the RS performance more than social bookmarks or tags because the majority of bookmark and tag information is not directly related to tourism and does not effectively represent user interests for attractions in the tourism domain. In addition, personalized recommendations by those methods have not been achieved because of the failure to acquire personal user preferences; instead, they recommend items to individuals based on the preferences of all users. Based on the processes described in the literature, the recommendation performance could be similar to the popularity case of the PTIS.

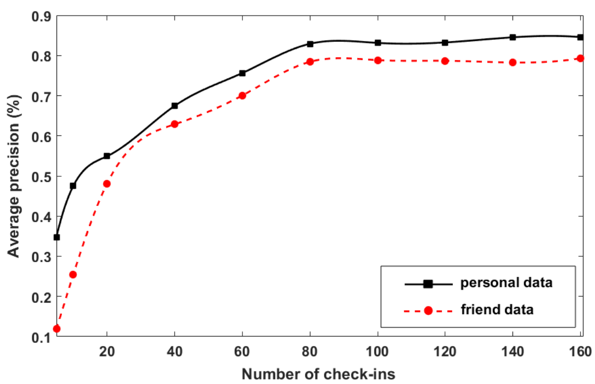

FIGURE 7. The average accuracy for category rankings with different numbers of check-ins using personal and friend data.

Because extracting check-in data is expensive and time consuming, more experiments were conducted to examine the optimal number of check-ins for precision compared with time spent in data extraction. As depicted in Figure 7, ranking accuracy tends to be higher when the number of check-ins and amount of Friend data increase. In the *personal data case*, the average recommendation accuracy is up to 86.91% when the system exploits at least the last 80 check-ins, whereas there is 72.83% average prediction accuracy using Friend data. Therefore, these trends show that the more check-in data are used, the higher the recommendation performance. However, this is not always the case because the number of check-ins and ranking performance is not always directly variant. Based on the results shown in Figure 7, ranking performance trends to remain steady when the number of check-ins exceeds 80.

In the *Friend data case*, the average precision is even slightly decreased when there are more than 120 check-ins. We analyzed this result and found that data from too many Friends can introduce noise in the user-interest computation model, which can lower the accuracy of the RS. We also found that using all check-in data of a user can cause imprecise recommendations because they do not represent current user interests. Therefore, the last 80 check-ins is the optimum number for the RS in the tourism domain. Figure 8 illustrates the time related (yearly) distributions of check-ins for seven years which are close to symmetric. This might be because the majority of participants were in the same age range and, thus, their check-ins behaviors are similar. The check-in medians are in the range 80-90 which is consistent with the optimal value of check-in data used for the recommendations demonstrated in Figure 7.

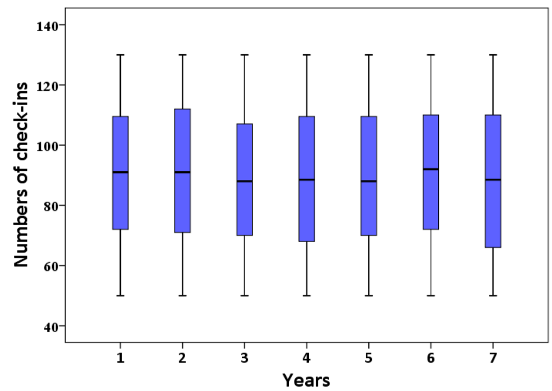

FIGURE 8. Check-in distributions of 520 volunteers for 7 years.

We also found that the system requires exponentially longer time to acquire Facebook check-in data when the number of check-ins increases. Two variables are examined in Figure 9: execution time and the number of years of check-ins. Figure 9 (a) and (b) demonstrate the user preference prediction precision using personal data and Facebook Friend data. The results show that the ranking accuracies in both line charts trend higher when the number of years increases. This result indicates that more check-in data produces higher prediction accuracy. Both figures illustrate the results for determining the optimal value of years of check-ins and execution time. The preference prediction precisions of both cases are increased exponentially from 1 to 5 years and remained steady thereafter. This is because five years of data are sufficient to analyze user preference. Although more information is added into the analysis model, they have little effect to the precision value and this makes the system stable at five years of check-ins. Figure 9 (a) shows that the highest preference prediction precision of 82.50% is achieved after 15 seconds of executing the acquisition of personal data. Figure 9 (b) indicates that the highest preference prediction precision of 72.83% took 66 seconds when using Friend data due to it not being replicated in the local database. The system will acquire data from Friends only when the user data is inadequate and the execution time would be up to one minute higher i.e. more than double the time required. These results are consistent with those of the the previous experiment; that

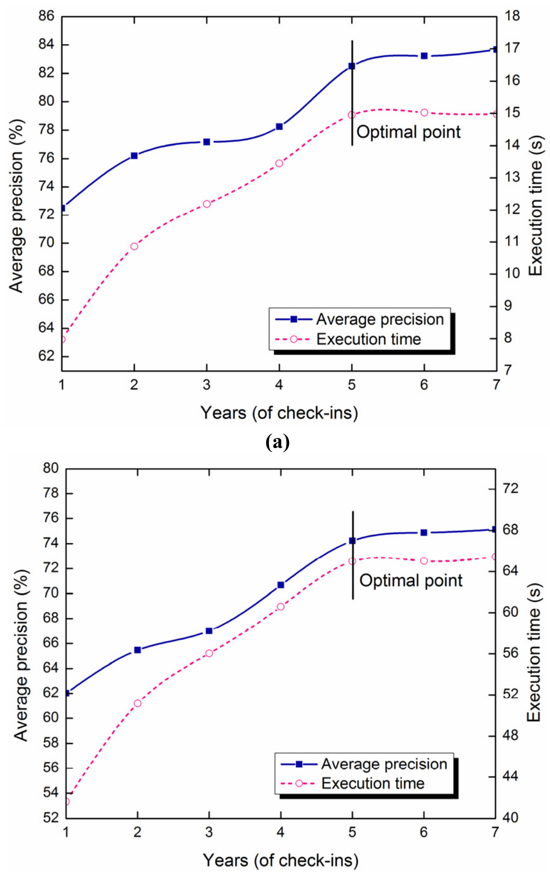

FIGURE 9. **The average precision and execution time for different numbers of years of (a) personal data and (b) Facebook friends' data.**

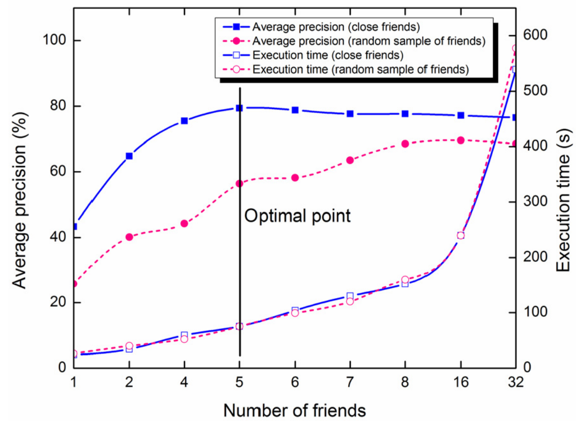

FIGURE 10. **The average accuracy and execution time using various numbers of close friends.**

is, using personal data is better than using Friend data. Based on these results, we conclude that 5 years of check-in data is the optimal value for user preference analysis in both cases.

C. OPTIMAL NUMBER OF FRIENDS

A user can have hundreds to thousands of Friends but few of them are close Friends. Using information of all Friends on Facebook could confuse the user interest analysis algorithm and also dramatically increase the time for data extraction. We investigated how much data from close Friends is necessary for making recommendations. To conduct this experiment, we separated our data into two datasets: 70% for a training set and 30% for the test set. Close Friends are identified by the algorithm ay-fb-friend-rank [57] described in section III. We further evaluated the preference prediction accuracy by varying the number of close Friends that were used to aggregate the check-in data. The average precision and execution times were compared to the results attained from using the data of a random sample of Friends.

Based upon the results in Figure 10, we found that five close Friends can establish the highest average accuracy of all category ranking and consume the lowest execution time. A greater number of close Friends seems to provide a higher accuracy of recommendation, but having too many Friends can exponentially increase the execution time. What was also interesting was that the average precision obtained from using a random sample of Friend data is lower than when close Friend data is used. This implies that a random sample of Friend data does not effectively represent user interests and can become noise for the analysis model.

We can conclude that a selection of five closest Friends is the optimal number, allowing the system to produce high-performance recommendations of relevant attractions for tourists and does not overload the RS. This demonstrates the potential of the exploitation of Facebook Friend data for the improvement of personalized recommendation services when no personal data are available, such as in the cold-start situation. This finding is particularly important because it can be used as a guideline for other RSs for the appropriate amount of data necessary for use in user-interest analysis while ensuring good recommendation performance and minimal processing time.

D. RECOMMENDATION FACTOR ANALYSIS

Several factors affect the quality of recommendations. In this experiment, we studied the effectiveness of those factors, such as popularity, Friend data, and appropriate times or seasons to visit an attraction. Precision and rank score (13) metrics were used for the evaluation. The rank score [69] is used for evaluating recommendation performance which will increase if the attractions at the top positions in the recommended list are chosen by users.

$$rankscore = \frac{\sum_{i \in \text{interest} \cap \text{recommended}} \frac{1}{2^{\frac{rank(i)-1}{\alpha}}}}{\sum_{i \in \text{interest}} \frac{1}{2^{\frac{k(i)-1}{\alpha}}}} \tag{13}$$

where α is the ranking half-life (chosen experimentally) (i.e., an exponential reduction factor), and k is the number of interesting places. We evaluated the recommendation performance using five different parameters as shown below.

1) Popularity (P): The popularity of place is solely considered for (4) and then we define β and $\gamma = 0$.

2) Friend data (F): The places Facebook Friends have visited is solely considered for (4) and then we define α and $\gamma = 0$.
3) Popularity and Friend data (PF): The popularity of places and places visited by Facebook Friends are considered for (4) and then we define $\gamma = 0$.
4) Popularity and recommended time (PT): The popularity of places and appropriate times to visit places are considered for (4) and then we define $\beta = 0$.
5) Popularity, Friend data, and suitable time (PFT): All variables are considered for (4).

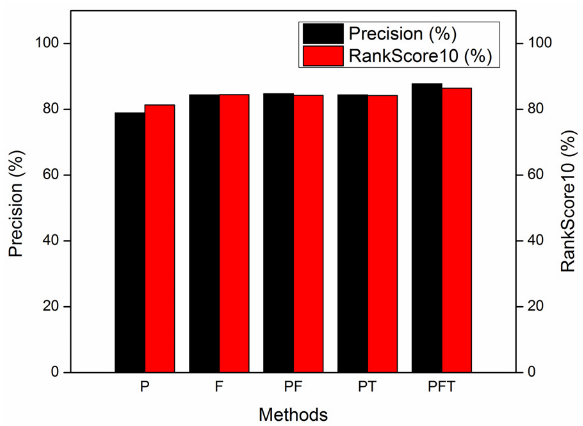

FIGURE 11. The average precision and rank score for attraction recommendations.

Figure 11 illustrates that PFT achieves superior average precision (87.75%) compared with other methods. These results are similar to the rank score evaluation of 86.47% on average. This result indicates that those three factors, popularity, Friend data, and suitable time for visiting are significant and affect the recommendation power. This is explained by the fact that travelers usually select their destinations based on those factors (e.g., where and when they should go, and if any of their Friends have ever visited; in order to ask them for some information). Tourists usually make their decisions based on these criteria. Therefore, considering those variables can improve the accuracy of RSs. Only PT can obtain similar average precision and rank scores to PF. When we investigated the results further, we found that travelers were slightly more interested in where their Friends have been (F), with 84.40% of precision, than in the popular places aggregated from other Facebook users (P), with only 78.94% of precision. In other words, travelers are more interested in the attractions where Friends have checked-in than in popular places visited by other Facebook users. Nonetheless, incorporating time-aware (T) analysis with Friend data (F) and popularity (P) in the recommendation can significantly improve the precision of recommendations as shown by PFT. Rank scores also illustrate that not only are the recommendations precise, but also the users' places of interest are shown in the top positions in their recommendation results. As such, this result is consistent with the previous experiments that Facebook Friends' check-in data can be used to represent user interests and enhances the recommendation performance of the RS. Thus, Hypothesis 2 is validated successfully.

Our findings show that Facebook Friend data can increase recommendation performance and enable the RS to identify the attractions in which a user is likely to be interested. Facebook Friend data are easy to collect without additional user effort or special devices and are valuable sources of information for RSs, especially in the tourism context, which has never previously been studied or deployed, according to [6], [13], and [26]. In addition, temporal context information is also exploited by PTIS for recommendations because it has been proven very useful for improving the recommendation power, as Campos *et al.* [60] found.

Based on our survey of personalized RS, as discussed previously, it is apparent that those methods have limitations and shortcoming related to the recommender environment. Although some published work [39], [40], [42] used similar data to our research, the methodology we developed to exploit the data is very different. We consider that those other researchers ignored the case where users change their interests over time. Some [43]–[45] did not consider the time factor at all. In other words, similar weight was given to the places visited a year ago as was given to any venues checked into only six months prior. This lack of consideration of the time-aware factor potentially lowers the quality of the recommendations. Ye *et al.* [41] introduced the time dimension into the recommendation model, but their time splitting was over the course of a day, as in morning, afternoon or evening, which is more suitable for restaurants, coffee houses, offices or hotels where time of day clearly influences decision making. This is too narrow a timeframe, where seasons are more appropriate to travel planning. Our proposed context awareness is different inasmuch as within these time frames we refine the recommendations that may influence user's option at a particular moment by considering seasonal factors among other longer timeframes that are more suitable for selecting attractions and events to visit and the weightings given to any particular place or event is reduced as time passes. In addition, our framework attempts to detect changes in a user's interest using a relevance feedback technique which makes our user model more dynamic, and provides more accurate and reliable recommendations.

E. OPTIMAL VALUES OF α, β, AND γ FOR ATTRACTION WEIGHT

There are three parameters, α, β, and γ, for attraction weight, $W(p)$, computation as shown in (4). Here, the values of those parameters are defined as $0 \leq \alpha + \beta + \gamma \leq 1$, and we vary from 0 to 1 step by 0.1 in order to study their optimal values using four different techniques, P, F, PT, PF, and PFT. There is no structural way to select the optimal values of α, β, and γ. This method is called a grid search and involves numerous trials and errors [70] by varying one parameter at a time. Various pairs of α, β, and γ values were tried, and the one with the best average precision was selected. Table 5 shows

TABLE 5. The optimal α, β, and γ values of all recommendation methods and average precisions obtained from those variables setting.

	α	β	γ	Avg. precision (%)
Popularity (P)	0.8	-	-	78.94
Friend data (F)	-	0.4	-	84.40
Popularity + Friend data (PF)	0.4	0.6	-	85.73
Popularity + Time (PT)	0.8	-	0.2	85.01
Popularity + Friend data + Time (PFT)	0.2	0.5	0.3	87.75

that places visited by Facebook Friends (β) significantly affected the recommendation performance. The highest value of average precision was obtained when β equals 0.5.

This result confirms that places visited by Facebook Friends affect travelers' decision making to choose their desired destinations more than do both popularity (α) and appropriate time to visit (γ). The highest average precision value of 87.75% was obtained when the value of α was 0.3, and γ equals 0.2. The optimal values of the other recommendation methods are shown in Table 5.

F. ADAPTIVE USER MODEL EVALUATION

The purpose of the following experiment was to evaluate the flexibility of the user model of PTIS compared with the state-of-the-art techniques and content-based and collaborative-based methods. Similar to the other experiments, this experiment was conducted based on the same group of participants. The user models of 10 people were removed to force the system to face the cold-start problem. We used a PFT-weighting scheme for this investigation. Figure 12 shows the values of average precisions obtained from four methods.

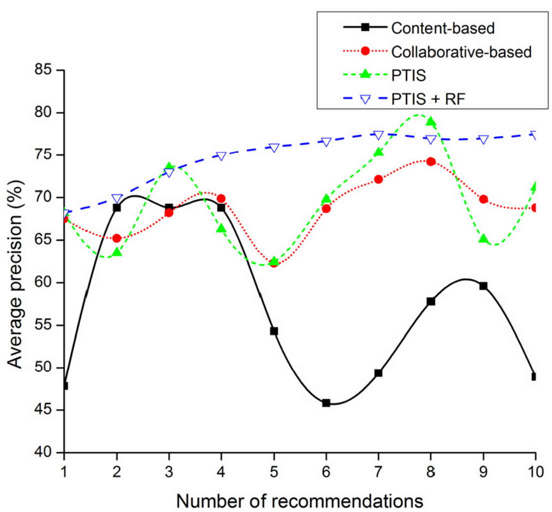

FIGURE 12. The average precision of four different techniques in various recommendation iterations.

The content-based technique yielded significantly low precision at 47.80% in the first iteration of evaluation. Its performance fluctuated in the rest of the experiment because sometimes people are not interested in the places recommended by the system. As such, the content-based technique has the lowest average precision. The collaborative-based method exploited Facebook data; therefore, the average precision value is higher than for the content-based method. However, this technique lacks a mechanism to find close Friends. It deploys all Facebook Friends' check-in data, which is extensively time consuming and some of this data confuses the system's analysis of user interests, which affects the recommendation quality. In contrast, the mechanism for finding close Friends on Facebook makes PTIS outperform those two methods in the average precision of ten rounds of iterations. However, the recommendation performance of PTIS fluctuated because it does not consider changes in user interests; therefore, the system always recommends the same kinds of attractions to users. The average precision described by several existing works [2], [16], [17], [21], [47], [48] may suffer from changes in the travelers' interests over time, being unable to learn from changing user behaviors, meaning that these user models are static and cannot adapt to time-related user behavior changes or new, important and perhaps one-off events that modify the range of local destinations that tourists should not miss. Therefore, the recommendation performance is decreased when users change their interests over time. In addition, this research incorporates time-awareness into recommendations which are different to, or are in addition to, existing works ([43]–[46], [71]). Consequently, the performance of those frameworks could be lower than for the PTIS because they could recommend to the venues that may be not recommended to visit at specific time. Our views as expressed here have not been experimentally confirmed, however, but we are confident that, theoretically at least, we are correct.

In addition, this experiment reveals the effectiveness of the method to capture the users' change of interests by deploying the relevance feedback (RF) technique [63]. Figure 12 shows that the average precision tends to be greater when the RF is incoporated with our system (PTIS+RF). When the number of recommendation iterations increases, greater average precision is obtained. This is because user actions are implicitly collected and this data causes the adjustment of parameters in the user model, which can improve the recommendation accuracy. Therefore, this indicates that the RF technique is able to effectively detect user interest changes, make user models more adaptive, and improve the recommend mechanism of the system significantly. Therefore, Hypothesis 3 is evaluated successfully for the tourism domain.

V. LIMITATIONS

However, there are some other limitations for the PTIS. First, our approach relies solely on user check-in activities which only represent visits to specific places without indicating like or dislike preferences. Some people might check-in for criticizing the unsatisfied service they received [50]. One possible way to do this is we need an additional Natural Language Processing tool to process comments or captions of users and exploit them as a factor for the recommendations. However, the results of this work are at least a good approximation for point-of-interest recommendations and our findings can be used by other researchers to benefit tourism industries.

Second, social networks such as TripAdvisor or Foursquare provide opportunities to aggregate user preference data to address the cold-start problem and enhance the recommendation accuracy for tourists. Those social networks provide some missing attraction information that Facebook rarely offers; e.g., reviewing, rating, or attraction information. This information is useful to make the user model more complete and the system more reliable by enhancing its recommendation power. Instead of relying on data from a single source of information, aggregating data across social networks is possibly another direction for our future work.

Third, while we have addressed demographic information in our research, we acknowledge that we include relatively simple demographic data. Other researchers such as Reinecke and Bernstein [72] have proposed the novel idea to allow the system to generate an interface that corresponds to users' cultural preferences thereby extending the type of demographic data included.Different cultural backgrounds influence people to have different preferences, as found by Callahan [73] and Burgmann *et al.* [74], and these background differences can inhibit the successful use of information technology [75] and user acceptance [76]. Several companies have realized this and now offer localized versios of their Web site, such as Google and Microsoft. Our future work will be to makea PTIS that also addressesdifferent user cultural backgrounds to improve recommendations.

Finally, the Web service policies and methodologies of social networks like Facebook are always dynamic. They can change their service policies at any time and these changes can create problems for the PTIS to acquire data from the social network. Consequently, the recommendation quality and robustness of the system could be affected and the system may not work properly. This risk may arise at any time in the future. RS developers should be seriously concerned by this issue and prepare a plan for modification of the system to respond to any future changes in social network services.

VI. CONCLUSIONS

Our motivation for developing the PTIS was to contribute to the sustainable growth and development of tourism industries. Our contribution is novel in that we proposed an approach to analyze user interests and perform personalized attraction recommendations using check-in information extracted from Facebook services. This information is directly useful for user attraction preference analysis and significantly benefits tourism industries. Our approach differs from existing approaches presented in the literature by overcoming the cold-start problems by collecting information from individual users and Friends available in Facebook. Here, close Friends are identified based on three factor: affinity score, edge weight, and time decay. The system uss close Friends' information to detect attractionsin which the target user may be interested. By this means theRS overcomes the cold-startproblems which are the weakness of the current state-of-the-art approaches. We also ound that aggregating and exploiting onl the information of a certain number of close Friends can overcome the data overload problem and effectively reduces the use-interest extracton and processing time while providing the same recommenation accuracy.

We also make the user profiles more flexible by employing the relevantuser feedback derivd from the analysis of user actions. The RF technique calculates the relevantattraction scors and continuously updates the user's profile with the new scores which enables adaptiveuser profiles and therefore enhances therecommendation performance.

An experimental evaluation of the presented approach has demonstrated that it is capable of predicting user preferences with high accuracy. The recommendations using information from popularity, close Friends, and suitable visiting time data gained higher precision compared with other techniques. With popularity, Friend, and time awareness recommendations, our system can improve its recommendation accuracy.

ACKNOWLEDGMENTS

The sponsors had no role in the design and conduct of the research; data collection, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript. We also acknowledge the contribution of Mr. Roy I. Morien of the Division of Research Administration at NU for his editing and checking of English grammar and expression in this paper.

AUTHOR CONTRIBUTIONS

Collected data for the experiment: A. Salaiwarakul, and W. Juraphanthong. Conceived, designed the experiments, and wrote the paper: K. Kesorn and A. Salaiwarakul. Involved in the discussions and analysis plans for the paper from its inception, including the idea of the data analysis: K. Kesorn. All authors read and approved the final manuscript.

REFERENCES

[1] F. Abel, E. Herder, G.-J. Houben, N. Henze, and D. Krause, "Cross-system user modeling and personalization on the social Web," *User Model. User-Adapt. Interact.*, vol. 23, nos. 2–3, pp. 169–209, 2013.

[2] M. Montaner, B. López, and J. L. de la Rosa, "A taxonomy of recommender agents on the Internet," *Artif. Intell. Rev.*, vol. 19, no. 4, pp. 285–330, 2003.

[3] C. S. Firan, W. Nejdl, and R. Paiu, "The benefit of using tag-based profiles," in *Proc. Web Conf.*, 2007, pp. 32–41.

[4] F. Abel, Q. Gao, G.-J. Houben, and K. Tao, "Analyzing user modeling on twitter for personalized news recommendations," in *Proc. 19th Int. Conf. User Modeling Adapt., Personalization*, 2011, pp. 1–12.

[5] B. Cui and X. Chen, "An online book recommendation system based on Web service," in *Proc. 6th Int. Conf. Fuzzy Syst. Knowl. Discovery*, vol. 7. 2009, pp. 520–524.

[6] H. Kumar, S. Lee, and H.-G. Kim, "Exploiting social bookmarking services to build clustered user interest profile for personalized search," *Inf. Sci.*, vol. 281, pp. 399–417, Oct. 2014.

[7] A. Gupta and K. Singh, "Location based personalized restaurant recommendation system for mobile environments," in *Proc. Int. Conf. Adv. Comput., Commun. Informat.*, 2013, pp. 507–511.

[8] A. Felfernig, K. Isak, K. Szabo, and P. Zachar, "The VITA financial services sales support environment," in *Proc. 19th Nat. Conf. Innov. Appl. Artif. Intell.*, vol. 2. 2007, pp. 1692–1699.

[9] A. Gupta and J. Anwiti, ''Life insurance recommender system based on association rule mining and dual clustering method for solving cold-start problem,'' *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 3, no. 10, pp. 951–954, 2013.

[10] K. Tu *et al.*, ''Online dating recommendations: Matching markets and learning preferences,'' in *Proc. 23rd Int. Conf. World Wide Web*, 2014, pp. 787–792.

[11] P. Gupta, A. Goel, J. Lin, A. Sharma, D. Wang, and R. Zadeh, ''WTF: The who to follow service at Twitter,'' in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 505–514.

[12] J. Borrís, A. Moreno, and A. Valls, ''Intelligent tourism recommender systems: A survey,'' *Expert Syst. Appl.*, vol. 41, no. 16, pp. 7370–7389, 2014.

[13] P. De Meo, A. Nocera, G. Terracina, and D. Ursino, ''Recommendation of similar users, resources and social networks in a social internetworking scenario,'' *Inf. Sci.*, vol. 181, no. 7, pp. 1285–1305, 2011.

[14] J. J.-C. Ying, W.-N. Kuo, V. S. Tseng, and E. H.-C. Lu, ''Mining user check-in behavior with a random walk for urban point-of-interest recommendations,'' *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, pp. 40-1–40-26, 2014.

[15] M. Arias, A. Arratia, and R. Xuriguera, ''Forecasting with Twitter data,'' *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 1, pp. 8-1–8-24, 2014.

[16] K. Kabassi, ''Personalizing recommendations for tourists,'' *Telematics Informat.*, vol. 27, no. 1, pp. 51–66, 2010.

[17] S. Schiaffino and A. Amandi, ''Building an expert travel agent as a software agent,'' *Expert Syst. Appl.*, vol. 36, no. 2, pp. 1291–1299, 2009.

[18] U. Rabanser and F. Ricci, ''Recommender systems: Do they have a viable business model in e-Tourism?'' in *Information and Communication Technologies in Tourism*. Vienna, Austria: Springer, 2005, pp. 160–171.

[19] Y. Zheng, Y. Liu, J. Yuan, and X. Xie, ''Urban computing with taxicabs,'' in *Proc. 13th Int. Conf. Ubiquitous Comput.*, New York, NY, USA, 2011, pp. 89–98.

[20] B. Lika, K. Kolomvatsos, and S. Hadjiefthymiades, ''Facing the cold start problem in recommender systems,'' *Expert Syst. Appl.*, vol. 41, no. 4, pp. 2065–2073, Mar. 2014.

[21] G. Adomavicius and 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.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.

[22] V. A. Rohani, Z. M. Kasirun, S. Kumar, and S. Shamshirband, ''An effective recommender algorithm for cold-start problem in academic social networks,'' *Math. Problems Eng.*, vol. 2014, p. e123726, Mar. 2014.

[23] M. Jackson. Social Media Today. (2014). *Use Hashtags Effectively for Greater Branding Benefits*. [Online]. Available: http://www.socialmediatoday.com/content/use-hashtags-effectively-greater-branding-benefits

[24] M. Sluckie. (2012). *Best Practices for Journalists*. [Online]. Available: https://blog.twitter.com/official/en_us/a/2012/best-practices-for-journalists.html

[25] Zocial Inc. (2014). *Thailand Zocial Award 2014 Insightful Information of Thai People on Social Media*. Accessed: Jun. 14, 2014. [Online]. Available: http://zocialinc.com/zocialawards2014/ThailandZocialAwards2014.pdf

[26] B. Shapira, L. Rokach, and S. Freilikhman, ''Facebook single and cross domain data for recommendation systems,'' *User Model. User-Adapt. Interact.*, vol. 23, nos. 2–3, pp. 211–247, 2012.

[27] L. Coyle and P. Cunningham, ''Exploiting re-ranking information in a case-based personal travel assistant,'' in *Proc. 5th Int. Conf. Case-Based Reasoning*, 2003, pp. 11–20.

[28] F. Carmagnola *et al.*, ''Tag-based user modeling for social multi-device adaptive guides,'' *User Model. User-Adapt. Interact.*, vol. 18, no. 5, pp. 497–538, 2008.

[29] C. Gena, F. Cena, F. Vernero, and P. Grillo, ''The evaluation of a social adaptive website for cultural events,'' *User Model. User-Adapt. Interact.*, vol. 23, no. 3, pp. 89–137, 2013.

[30] A. Ali. Social Media Today. (May 8, 2015). *What People Like to Share More on Social Media [INFOGRAPHIC]*. Accessed: Sep. 12, 2015. [Online]. Available: http://www.socialmediatoday.com/social-networks/2015-05-08/what-people-share-more-social-media-infographic

[31] S. Poslad and K. Kesorn, ''A multi-modal incompleteness ontology model (MMIO) to enhance information fusion for image retrieval,'' *Inf. Fusion*, vol. 20, pp. 225–241, Nov. 2014.

[32] C. Castillo, M. Mendoza, and B. Poblete, ''Information credibility on Twitter,'' in *Proc. 20th Int. Conf. World Wide Web*, 2011, pp. 675–684.

[33] M. Wasim, I. Shahzadi, Q. Ahmad, and W. Mahmood, ''Extracting and modeling user interests based on social media,'' in *Proc. IEEE 14th Int. Multitopic Conf.*, Dec. 2011, pp. 284–289.

[34] K. H. Lim and A. Datta, ''Interest classification of Twitter users using wikipedia,'' in *Proc. 9th Int. Symp. Open Collaboration*, 2013, pp. 22-1–22-2.

[35] M. Michelson and S. A. Macskassy, ''Discovering users' topics of interest on Twitter: A first look,'' in *Proc. 4th Workshop Anal. Noisy Unstruct. Text Data*, 2010, pp. 73–80.

[36] E. Michlmayr, S. Cayzer, and P. Shabajee, ''Add-a-tag: Learning adaptive user profiles from bookmark collections,'' in *Proc. Collections Int. Conf. Weblogs Social Media*, 2007, pp. 1–2.

[37] F. Durao and P. Dolog, ''A personalized tag-based recommendation in social Web systems,'' *Adapt. Pers. Web 20*, vol. 485, pp. 40–49, Mar. 2009.

[38] H.-N. Kim and A. El Saddik, ''Exploring social tagging for personalized community recommendations,'' *User Model. User-Adapt. Interact.*, vol. 23, nos. 2–3, pp. 249–285, 2012.

[39] F. G. Davoodi and O. Fatemi, ''Tag based recommender system for social bookmarking sites,'' in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, Aug. 2012, pp. 934–940.

[40] H. Wang, M. Terrovitis, and N. Mamoulis, ''Location recommendation in location-based social networks using user check-in data,'' in *Proc. 21st ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, 2013, pp. 374–383.

[41] M. Ye, P. Yin, and W.-C. Lee, ''Location recommendation for location-based social networks,'' in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, 2010, pp. 458–461.

[42] Y.-L. Zhao, L. Nie, X. Wang, and T.-S. Chua, ''Personalized recommendations of locally interesting venues to tourists via cross-region community matching,'' *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, pp. 50-1–50-26, 2014.

[43] B. Berjani and T. Strufe, ''A recommendation system for spots in location-based online social networks,'' in *Proc. 4th Workshop Social Netw. Syst.*, New York, NY, USA, 2011, pp. 1–6.

[44] J. Shen, C. Deng, and X. Gao, ''Attraction recommendation: Towards personalized tourism via collective intelligence,'' *Neurocomputing*, vol. 173, pp. 789–798, Jan. 2016.

[45] J. P. Lucas, N. Luz, M. N. Moreno, R. Anacleto, A. A. Figueiredo, and C. Martins, ''A hybrid recommendation approach for a tourism system,'' *Expert Syst. Appl.*, vol. 40, no. 9, pp. 3532–3550, 2013.

[46] I. Christensen, S. Schiaffino, and M. Armentano, ''Social group recommendation in the tourism domain,'' *J. Intell. Inf. Syst.*, vol. 47, no. 2, pp. 209–231, 2016.

[47] P. Lops, M. de Gemmis, and G. Semeraro, ''Content-based recommender systems: State of the art and trends,'' *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, P. B. Kantor, Eds. New York, NY, USA: Springer, 2011, pp. 73–105.

[48] L. Terveen and W. Hill, ''Beyond recommender systems: Helping people help each other,'' in *HCI in the New Millennium*, J. Carroll, Ed. Boston, MA, USA: Addison-Wesley, 2001, pp. 487–509.

[49] A. Corallo, G. Lorenzo, and G. Solazzo, ''A semantic recommender engine enabling an eTourism scenario,'' in *Knowledge-Based Intelligent Information and Engineering Systems*. Berlin, Germany: Springer, 2006, pp. 1092–1101.

[50] G. Xu, B. Fu, and Y. Gu, ''Point-of-interest recommendations via a supervised random walk algorithm,'' *IEEE Intell. Syst.*, vol. 31, no. 1, pp. 15–23, Jan./Feb. 2016.

[51] M. J. Pazzani, ''A framework for collaborative, content-based and demographic filtering,'' *Artif. Intell. Rev.*, vol. 13, nos. 5–6, pp. 393–408, 1999.

[52] R. Burke, ''Knowledge-based recommender systems,'' *Encycl. Libr. Inf. Syst.*, vol. 69, no. 32, pp. 175–186, 2000.

[53] S. Y. Maw, M. M. Naing, and N. L. Thein, ''RPCF algorithm for multi-agent tourism system,'' in *Proc. Int. Symp. Micro-NanoMechatronics Hum. Sci.*, 2006, pp. 1–6.

[54] D. Asavasuthirakul, A. Harfield, and K. Kesorn, ''A framework of personalized travelling information services for Thailand,'' *Adv. Mater. Res.*, vols. 931–932, pp. 1382–1386, May 2014.

[55] Facebook. Facebook Developers. (2016). *Graph API*. Accessed: Mar. 15, 2016. [Online]. Available: https://developers.facebook.com

[56] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. London, U.K.: Cambridge Univ. Press, 2008.

[57] K. Kuizinas. GitHub. (2012). *Facebook-Friend-Rank*. [Online]. Available: https://github.com/gajus/facebook-friend-rank

[58] J. Widman, (2014). *EdgeRank*. [Online]. Available: http://edgerank.net/

[59] A. K. Dey, "Understanding and using context," *Pers. Ubiquitous Comput.*, vol. 5, no. 1, pp. 4–7, 2001.
[60] P. G. Campos, F. Díez, and I. Cantador, "Time-aware recommender systems: A comprehensive survey and analysis of existing evaluation protocols," *User Model. User-Adapt. Interact.*, vol. 24, no. 1, pp. 67–119, 2014.
[61] L. Xiang *et al.*, "Temporal recommendation on graphs via long- and short-term preference fusion," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2010, pp. 723–732.
[62] Y. Yusof and Z. Mustaffa, "Dengue outbreak prediction: A least squares support vector machines approach," *Int. J. Comput. Theory Eng.*, vol. 3, no. 4, pp. 489–493, 2011.
[63] G. Salton, *The SMART Retrieval System-Experiments in Automatic Document Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1971.
[64] A. Moreno, A. Valls, D. Isern, L. Marin, and J. Borràs, "SigTur/E-destination: ontology-based personalized recommendation of tourism and leisure activities," *Eng. Appl. Artif. Intell.*, vol. 26, no. 1, pp. 633–651, Jan. 2013.
[65] G. A. Miller, "WordNet: A lexical database for English," *Commun. ACM*, vol. 38, no. 11, pp. 39–41, 1995.
[66] H. Steck, "Training and testing of recommender systems on data missing not at random," in *Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2010, pp. 713–722.
[67] X. Yang, H. Steck, Y. Guo, and Y. Liu, "On top-k recommendation using social networks," in *Proc. 6th ACM Conf. Recommender Syst.*, 2012, pp. 67–74.
[68] T. Bogers and A. van den Bosch, "Recommending scientific articles using citeulike," in *Proc. ACM Conf. Recommender Syst.*, 2008, pp. 287–290.
[69] A. Gunawardana and G. Shani, "A survey of accuracy evaluation metrics of recommendation tasks," *J. Mach. Learn. Res.*, vol. 10, pp. 2935–2962, Dec. 2009.
[70] L. J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1506–1518, Nov. 2003.
[71] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, "Exploiting geographical influence for collaborative point-of-interest recommendation," in *Proc. 34th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, 2011, pp. 325–334.
[72] K. Reinecke and A. Bernstein, "Knowing what a user likes: A design science approach to interfaces that automatically adapt to culture," *MIS Quart.*, vol. 37, no. 2, pp. 427–453, 2013.
[73] E. Callahan, "Cultural similarities and differences in the design of University Web sites," *J. Comput.-Mediated Commun.*, vol. 11, no. 1, pp. 239–273, 2005.
[74] I. Burgmann, P. J. Kitchen, and R. Williams, "Does culture matter on the Web?" *Marketing Intell. Planning*, vol. 24, no. 1, pp. 62–76, 2006.
[75] D. E. Leidner and T. Kayworth, "Review: A review of culture in information systems research: toward a theory of information technology culture conflict," *MIS Quart*, vol. 30, no. 2, pp. 357–399, 2006.
[76] A. Kappos and S. Rivard, "A three-perspective model of culture, information systems, and their development and use," *MIS Quart*, vol. 32, no. 3, pp. 601–634, 2008.

K. KESORN received the Ph.D. degree in electronic engineering from Queen Mary College, University of London, U.K. He is currently an Associate Professor with the Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Thailand. His current research interests include semantic multimedia retrieval, knowledge-based modeling for multimodal information retrieval, semantic data processing and data mining, and also social business intelligence. He has participated in several research projects and has been a reviewer for several world-class journals, including the IEEE transactions and the Elsevier journals. He is currently receiving research grants as a new scholar from the Thailand Research Fund from 2016 to 2017.

W. JURAPHANTHONG received the master's degree in information technology from the Department of Computer Science and Information Technology, Faculty of Science, Naresuan University, Thailand, where she is currently pursuing the Ph.D. degree in computer engineering with the Faculty of Engineering. She is involved in the research topic of security in embedded systems.

A. SALAIWARAKUL received the Ph.D. degree in computer science from the University of Birmingham in 2009. She is currently a Lecturer with the Department of Computer Science and Information Technology, Naresuan University, Phtisanulok, Thailand. Her research interests include computer security, semantic webs, and ontologies.