Journal of AI and Data Mining
Vol 8, No 3, 2020, 379-389.
DOI: 10.22044/JADM.2020.7326.1871

Improving Accuracy of Recommender Systems using Social Network Information and Longitudinal Data

B. Hassanpour[1], N. Abdolvand[2*] and S. Rajaee Harandi[2]

1. Department of Electrical, Computer and IT Engineering, Qazvin Islamic Azad University, Qazvin, Iran.
2. Department of Management, Faculty of Social Sciences and Economics, Alzahra University, Tehran, Iran.

Received 01 August 2018; Revised 21 July 2019; Accepted 31 March 2020
**Corresponding author: n.abdolvand@alzahra.ac.ir (N. Abdolvand).*

Abstract
The rapid development of technology, the Internet, and the development of electronic commerce have led to the emergence of the recommender systems. These systems assist the users in finding and selecting their desired items. The accuracy of the advice in recommender systems is one of the main challenges of these systems. Regarding the fuzzy system capabilities in determining the borders of user interests, it seems reasonable to combine it with social network information and the factor of time. In this work, for the first time, we try to assess the efficiency of the recommender systems by combining fuzzy logic, longitudinal data, and social network information such as tags, friendship, and membership in groups. Also the impact of the proposed algorithm for improving the accuracy of the recommender systems is studied by specifying the neighborhood and the border between the users' preferences over time. The results obtained reveal that using longitudinal data in social network information in memory-based recommender systems improves the accuracy of these systems.

Keywords: *Recommender System, Social Network, Longitudinal Data, Fuzzy Logic, Tags, Membership.*

1. Introduction
In the recent years, with the growth of Internet users, many online networks have been developed, and thousands of people talk online together [1]. Every day, the number of articles, music files, movies, books, and web pages on the Internet is increasing. In such an environment, people do not know what to do with the enormous amount of information. They are often unaware of the opportunities because of the high volume of data and, in some cases, irrespective of the decision in this regard. In the past, the opinions of friends, classmates or colleagues were used to overcome this type of problem but today no one can offer various suggestions according to all the available information. This data growth and massive data sources have created new fields of data mining and pattern discovery [1]. The recommender systems were developed in response to the changing needs of people and the number of choices in every field that help users to find and select items [2]. Today, many websites including Amazon, CDNOW, Barnes & Noble, and IMDb use the recommender systems in order to provide a good service for their customers [3]. Of course, not only the consumers and individuals use the recommender systems but the suppliers and vendors will also benefit from these tools. Because of a large number of customers and goods, the recommender systems try to give users a better suggestion. Regarding the information requirements of these systems, the mechanisms that enable these systems to collect the required information are necessary. Another issue that is considered in some networks, and is also considered in this work, is the matter of time. A person may change or forget her/his interest over time and reminding them may change one's preferences and desires [4]. The time of rating goods in social networks indicates the users' behavior over time. Therefore, using the time dimension will improve advice estimation. Failing to pay attention to the time people spend on networks and neglecting to update the considerations of network requirements can lead to a potentially significant loss of information [5].

Moreover, the information explosion may not necessarily improve the quality of people's lives, and finding facts and knowledge in the wealth of information available can be time-consuming and even frustrating. Therefore, having an intelligent system capable of automatically learning the users' interests, filtering unrelated interests, offering the relevant information in a limited time, and helping users with product-selection decisions is essential. The recommender systems counter the problem of information overload and assist with the problem solving by recommending items to users [6]. Although several memory-based recommender algorithms have been proposed, an algorithm that uses the social network data and addresses the time factor has not been introduced [7]. Furthermore, many of the current approaches in the recommender systems focus on offering users the most relevant recommendations and overlook the background information such as time, place, and people participating in a given action. In other words, the traditional recommender systems use only the two entities of users and elements in the process of providing advice. Thus the time series analysis in combination with other features of social networks can provide a good measure of suitable recommender systems [8]. Given the different preferences of users at different times, the recommendations users receive should be based on these preferences and should not be independent from the recommendations received at other times [9]. This raises an interesting question: given the importance of time and the influence of friends and people in decision-making, how can a time component be added to social network information to improve the efficiency of the recommender systems? Because of the acceptable results of the Jin *et al*. [4] method, in comparison with other methods, especially techniques based on social networks, it is used as a reference method in this work. Hence, in this work, we aimed to design memory-based recommender systems through the integration of longitudinal data and social network information. Therefore, a recommender system is designed by combining fuzzy logic, longitudinal data, social network information such as tags, friendship, and membership in groups, and by using the impact of long-standing interest, although with variable coefficients so that the proposed method can improve the efficiency of previous recommender methods. Overall, this system has a huge impact on improving the efficiency of the proposed method.

This work begins with a review of the literature, followed by a description of the research method used and proposed algorithm. Finally, the implications and conclusions will be explained.

2. Literature review

The recommender systems have become an essential research area. Their main purpose is to identify the users' neighbors based on profile similarity and then suggesting something that the neighbors had already liked [10]. The recommender systems are intelligent systems that provide appropriate recommendations to everyone and help the users to find and select their required items by finding and then analyzing their data [2, 11]. Usually, these systems are not able to offer suggestions without an accurate information about the users and their desired items (e.g. movies, music, books) [12]. According to Liang *et al*. [11], the recommender system is a subset of the decision support systems and is defined as the information system with the ability to analyze the past behavior and make recommendations for current issues. Moreover, the recommender systems are algorithms that provide the best and most accurate recommendations through the exploration of users' associated information from the relevant databases. Such systems find patterns in the users' data by examining their past choices and displaying appropriate recommendations based on them [11]. Therefore, one of the essential goals of such systems is to collect data related to the users' interests and items in the system such as videos [12]. In a general classification, the recommender systems are divided into two categories: traditional recommender systems and computational intelligence methods [13]. The recommender systems based on computational intelligent methods use artificial intelligence tools such as Bayesian techniques, neural networks, clustering techniques, genetic algorithms, and fuzzy theory to build the proposed model. The traditional recommender systems are divided into three general categories: collaborative filtering, content-based recommendation, and hybrid approaches [2,13]. The content-based recommendation method chooses the characteristics of the items with the highest popularity among celebrities like directors and actors to offer them to others. Hybrid approaches make recommendations by combining the collaborative filtering and content-based recommendation approaches.

Collaborative filtering is the most widely used technology in the recommender systems and makes suggestions based on the ranking of active users compared to their neighbors [13,14]. Collaborative filtering is divided into two categories: memory-based and model-driven approaches. Memory-based algorithms operate on the entire ranking matrix and make recommendations by identifying the target user's neighborhood. Such recommendations are based on past ratings. Model-driven techniques use the rated data for training the model, and thereafter, the model is used to derive recommendations [5,8,15]. With the advent of the Internet and the spread of social networks, extracting information from social web sites has grown, and integrating this information with the recommender systems to increase the efficiency of recommender systems has attracted a lot of attention [16,17]. The social network is a social structure between the users that can be individuals or organizations. This structure expresses the way in which different people are connected to each other [18]. In this regard, the hybrid approach of collaborative filtering and content-based approach has been proposed by Guy *et al*. [19], which indicates more interest in the tag-based recommender system than the user-based recommender system. Moreover, their results have indicated a better performance of the hybrid system. In addition, an algorithm that collect the user ratings and social network connections has been proposed by Liu and Lee [20], in which neighbors receive the same recommendations. Their results have indicated more accurate prediction algorithms by incorporating the social network information into CF. This algorithm only considers the friendship as a factor of social networks, while there are many other factors in social networks that can increase the accuracy of recommendations. In addition, the use of tags, its content, and longitudinal data, and its impact on people's interests are not considered in this work. Membership in a group is usually due to the interest in the similar topics so people of the same groups may have the same interest. Since people may have accidentally or out of curiosity join the group, it is not always true. However, in social networks, this can be considered as the criteria for measuring the individuals' similar interests and ideas. Therefore, a good use of this resource can help the recommender systems [7]. Lee and Brusilovsky [21] have used the community membership information to improve the personalized recommendations. They indicated their proposed system provided recommendations that were as accurate as those produced through a CF approach but with a better efficiency.

The ability to assign a tag to an item by the user is one of the features of most social recommender systems. Therefore, the users can allocate tags to each product based on their preferences. The social tagging systems have been through significant changes in the past five years and have been used in offers in the past two years [22]. Each tag is a keyword that is added to a digital object such as an audio, image, video or Web database to describe it; furthermore, it is one of the advanced features that represent the users' interests. Tags are selected words that, despite their simplicity in search, classification, and description, are very powerful sources. Most studies have been conducted in the field of creating the user models and predicting tags as well as exploring social networks. For example, a new method based on a combination of collaborative filtering and tag-based social relationships has been used by Naseri *et al*. [7] in order to improve the accuracy of recommendations based on the nearest neighborhood. In the proposed method, the new similarity of metric (proposed mathematical formulas) based on tags and social activities was used to calculate the nearest neighborhood. Then in the proposal part, a subset of users similar to the target user was selected based on the similarity matrix and was recommended based on their total points. In this algorithm, only very effective friends and active memberships were studied rather than all the friends. The limitations of this method are the lack of attention to behavioral changes (excluding the meaning and content of labels), and considering the same impact for all items in comparing similar users because a person may alter his/her interests over time, and items that are popular among all people cannot be a good measure of similarities. In order to improve collaborative filtering, Adeli and Moradi [23] proposed a web service recommendation method called Popular-Dependent Collaborative Filtering (PDCF), which predicts the quality of web services as well as users' dependency on a particular web service. They also proposed a Location-aware PDCF that considers the location of web services in the PDCF recommendation process. Halpin *et al*. [24] have proposed a general model based on group tagging, and indicated that the use of tags in collaborative filtering significantly increases the efficiency of the method. There are noises in using tags to the lack of restrictions on using them. This may be because of different interpretations. The method based on social tags for a book selling system, and used tags on different books for making recommendations were proposed by Heymann *et al*. [25], who have indicated that tags can improve sales. However, in this work, the data and label distribution were inadequate. Moreover, the methods for rating tags have examined the efficiency of using tags in the recommender systems. The date of using tags and the user clicks on them have been used for general label settings, and a common tagging system has been used for all users [26]. However, this method requires much more information about different users' conditions. Furthermore, tracking clicks from multiple users is costly. Moreover, an algorithm that uses the tag for identification and management of the required web sites has been developed by Durao and Dolog [27] to compute the tags' similarity. In this algorithm, the cosine similarity combined with other factors such as the popularity of tags, tag resolution, and user-tag information were used to rank. Given the tags' limitations, especially in finding the similarities, using tags alone is not a perfect measure for making recommendations. In order to resolve this problem, a standard classification system has been proposed by Liang *et al*. [11] to check the labels' integration in the recommender systems. Since in this method of classification, tags are collected by experts from different fields, the efficiency is high and there is a reasonable relationship between the tags and the goods. As the users' information and interests are tied with time, the primary methods of time-cognitive-based recommender systems are primarily concentrated on the dynamic and changing models [10,14,28]. Moreover, in an early research work on time-based algorithms, the researchers have tried to identify active users or items. Although the work has been done based on the profile that uses the user profile information over time, both methods are faced with information overload [29]. In order to resolve this problem, Baltrunas and Amatriain [28] have divided the user profiles into several sub-profiles such as text and music that have different contents; therefore, they could recognize people's different interests through different profiles. However, it also requires calculations, and information has to be gathered together in general decisions. As a result, much of the today's research works consider a set of pre-defined distributions of information over time to prevent information overload and abundance [1]. For example, an exponentially decreasing function of time has been used by Ding *et al*. [30] to compute time weights based on each user and each cluster of items. They have assigned a weight to each rating that has been defined with the function of time in the preference prediction phase of item-based collaborative filtering. Their experimental results indicated that the new algorithm could significantly increase the accuracy of prediction in the item-based collaborative filtering. Moreover, an algorithm based on item similarity and time has been proposed by Jin *et al*. [4], in which the similarities between items were used to measure the similarity between the ancient and recent behaviors because the lasting interest of the user is like his/her recent interest. Since the online visitors' behaviors are only shown in the database, there is a need for very real and updated data. In addition, an algorithm based on three strategies of weight-tag, weight-time, and a combination of weight-tag and weight-time has been used by Braunhofer *et al*. [31] to integrate the time and tag information. They indicated that the integration of time and tag information greatly improved the efficiency of the proposed method.

According to the research works in this field, different recommender systems have been proposed and each has its strengths and weaknesses. Therefore, the use of the strengths of these methods and elimination of their weaknesses is essential. It should be noted that a person may forget his/her interests over time, and people's interest and preferences may change by reminding them [4]. Moreover, the people are under the influence of their friends, and will be treated like their friends and other group members over time. Therefore, it is necessary to study the people's behavioural changes. Although the behavioural changes have been considered in the study of Jin *et al*. [4], the algorithms that involve information such as tags, friend, and groups have been neglected. Given the several memory-based recommender algorithms that have been proposed in different studies, the algorithm uses the social network information along with the time factor that has not been introduced [7]. Therefore, the integration of these factors seems to improve the performance of the memory-based recommender systems. Hence, this work aimed to assess the efficiency of the recommender systems by combining fuzzy logic, longitudinal data, and social network information such as tags, friendship, and membership in groups.

3. Research method

The objective of this work was to design a recommender system by combining the fuzzy logic, longitudinal data, and social network information such as tags, friendship, and membership in groups. Therefore, a multi-phase methodology was used to provide the recommender system. The proposed method consists of five phases of data collection, algorithm selection, implementation of the proposed system, making recommendations, and system evaluation. The first phase involves collecting, describing, and evaluating data. The primary data was collected, reviewed, and prepared by studying the existing recommender systems. During this phase, the previous studies that examined the social networking features or longitudinal data in the recommender systems and their strengths and weaknesses were reviewed to improve the recommender algorithm. In the next phase (implementation), which was the main phase in designing the recommender system, a memory-based collaborative filtering method was used using the social network information along with the longitudinal data. Then friendship similarity, tag-based similarity, similarities based on group membership were calculated, and recommendations were offered to the individuals based on the similarity between them. The third phase involved evaluating the proposed method. The proposed algorithm was tested using a simulation developed in the C# programming language and evaluated using the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and successful decision-making capacity (Top_N). It should be noted that the methods studied here were applied to a database using the MyMediaLite recommender system library. This is a free library for the recommender systems implemented in the C# language.

3.1. Model description

In this phase, depending on the type of collected data and based on need, the algorithm and model were selected. The synthetic algorithm that used the information for social networks was used in this work. Moreover, given the importance of time as the main characteristic of the information used in this system, the memory-based collaborative filtering method that used the social network information along with longitudinal data was used as the foundation of the proposed algorithm to increase the accuracy of the recommender system. Then the social network information collected from the users in social networks including four steps of similarity calculation based on friendship, tag, group membership, and overall similarity was used. In calculating the similarity based on the friendship, the similarity of people was determined based on the favorite items by considering the time. In addition, with respect to the specified intervals, the similar interests of people were involved in determining the amount of friendship. Since the interests are not forgotten, the previous friendship are involved in determining the overall similarity based on the friendships with lower coefficients. There are different criteria for calculating the similarity between the members. In this work, for the first time, the fuzzy method was used to determine the friendship between users that was computed as Equation 1:

$$\left\{\begin{array}{lll} HF(u,n) & cm_i > i-1 & \\ LF(u,v) & 0 < cm_i \le i-1 & i \neq 1 \\ NF(u,v) & \mathit{Otherwise} & \end{array}\right\} \tag{1}$$

where, i is the number of time intervals $i = <1,2,3,4,5>$, where 1 is the interval between 5 to 10 years, 2 is the interval between 10 to 15 years, and so it was proposed in the questionnaire. Moreover, cm_i is the number of common films between the two users u and v in a period of i; HF and LF. respectively, indicate the strong and poor friendships, and NF indicates no friendship. Therefore, the first two groups of friendship with high HF and low LF coefficients were used in the subsequent decisions (film selection and recommendation). The total friendship was computed using Equation 2, where μ and γ are the coefficients greater than zero:

$$Fsim_{u,v} = \mu * HF(u,v) + \gamma * LF(u,v) \tag{2}$$

The use of this system involves the same films in friendship in a fuzzy way. Thus the more the same number of films between the ages is, the chance of developing friendship between people is greater and is used in the subsequent decision with more coefficients. In order to obtain the highest efficiency in finding the similarities based on friendship, different values of μ and γ in the range of 0-1 were examined. In order to find the most appropriate values for μ and γ to find the best values, starting from zero for μ and γ and increasing to 0.1 per quantity, $FSim_{u,v}$ was used to compute d using HF and LF. Then the coefficients of other factors (tag and groups membership) were considered zero, and the accuracy of recommendations was computed using the Pearson formula. The best values for μ and γ that have the most relevant results in the process of advice selection were, respectively, 0.7 and 0.3.
Given that different tags can be used in various fields including film recommendation, it is necessary to select and use the most appropriate characteristics for tags. Therefore, in this part, tag information was used as the other indicator of calculating the similarity. For this purpose, the standard method used in the study of Naseri et al.

[7] for music recommendation was used. Regarding the variety of tags, the users were requested to score their favorite films with tags of excellent (10), very well (8), well (6), moderate (4), and weak (2), which were the conventional tags used in the films. However, the best tag was to allow the users to express their interest about a product. Due to the large number of different tags, the simplest method was used to provide tags [7]. The similarity between the two users u and v was computed according to Equation 3:

$$TSim_{u,v} = \frac{\sum_{i\in(I_u \cap I_v)}((|T_{u,v}|)^2 / (|T_{ui}| * |T_{vi}|))}{Max(|I_u|, |I_v|)} \quad (3)$$

where, $T_{u,v}$ is a set of shared tags between the users u and v for film I, $T_{u,i}$ is a set of tags allocated to item i by the user u, I_u is a set of items (films) selected by the user u, and $Max(|I_u|, |I_v|)$ indicates the total number of items (films) selected by the users u and v. Another criterion for calculating the similarity is the group membership. In this step, the similarity is computed based on the group membership using the fuzzy method to minimize the effect of random choices. In this work, the films are used to identify the group membership. Since selecting the film, especially on behalf of the lower age person, is not a good measure to put someone in a group, a fuzzy method that shows the interest of a person to the specific set of films is used to group the people. In general, there are several groups of films including action, science fiction, crime, comedy, romance, etc. Therefore, for subscribing any person in any of these groups, $km(i,j)$ is considered as the number of films that the user u observes from the group of $j =< Action, Fantasy, Crime, Romance, Comedy, ... >$ in the period of $i =< 1, 2, 3, 4, 5 >$. Then the membership in the mentioned group is obtained using Equation 4:

$$u \in G(j) \geq km(i, j) \geq i\text{-}2 \quad i \geq 3 \quad (4)$$

where, $G(j)$ represents the membership in the group j. In this way, the people aged 20-15, 25-20, and 30-25 from a specific group that have, respectively, viewed 1, 2, and 3 films are considered as a member of that group. Using the amount of $MSim_{u,v}$, the membership of the users u and v was computed based on Equation (4) in a way, the value of 1 is considered for being the membership and the value of 0 is considered otherwise. It should be noted that the tags can be used to improve the decision results about friendship and group membership. Finally, using the information obtained by calculating the similarity based on friendship, group membership, and tag, and according to the consistent methodology used by Naseri et al. [9], the overall similarity between the individuals is computed. The ultimate goal of this work is to use the social network information such as friendship, tags, and group memberships in collaborative filtering to enhance the accuracy of the recommender systems. Equation 5 is used to compute the overall similarity of the two users u and v.

$$Sim_{u,v} = a * TSim_{u,v} + (1-a)\left(b * FSim_{u,v} + (1-b) * MSim_{u,v}\right) \quad (5)$$

The values for α and β are used to create a significant relationship between the different criteria of friendships, group memberships, and tags. In this formula, α is used to make a relationship between the tags on one hand, and between friendship and membership on the other hand, and β is used to make a relationship between friendship and membership. The best values for α and β are obtained thorough an experimental method and the Pearson results. In order to obtain the best values, α and β start with zero and 0.1 is added to them at each stage (see Table 1). As indicated in Table 1, the best values for α and β are 0.7 and 0.7, respectively.

Table 1. Pearson values for α and β

β	α								
	0.9	**0.8**	**0.7**	**0.6**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
0.1	21.21	19.86	19.94	20.75	22.34	21.43	19.87	20.67	21.32
0.2	19.67	18.74	19.78	21.23	22.11	21.11	21.97	25.87	21.78
0.3	19.87	21.34	21.56	21.65	20.98	20.56	23.11	21.67	20.45
0.4	19.11	20.12	23.49	21.78	20.11	22.64	22.67	19.87	19.34
0.5	18.73	19.78	21.23	19.78	22.53	19.98	18.67	17.93	19.89
0.6	19.90	20.15	18.76	19.89	20.34	21.45	22.23	21.78	19.08
0.7	19.81	18.77	21.90	21.78	20.11	22.64	24.45	21.63	18.90
0.8	18.11	18.96	20.11	19.01	19.89	21.12	21.93	20.46	20.23
0.9	19.00	19.90	19.67	19.45	19.87	20.11	21.12	20.19	21.21

Prediction of the users' future behavior is one of the most important steps in the recommender systems. At this phase, the recommendations are given to the users based on the similarity between them so that a subset of users similar to the target user based on the similarity between them and their total weight of rates is used to recommend to the target user [8], and to predict the future behavior of the user. The proposed method recommends the items that the user u will tend to choose in the future. Calculation of the interest of user u to the specific items by his/her neighbor, user v has already been selected depending on the similarity between the two users and the similarity between item *i* that his/her neighbor, user v, has already been selected and items that have already been tagged by user u. The similarity between items is computed using the weighted Jaccard similarity metric [31] (Equation 6):

$$SimItem(i,j) = \frac{\sum_{t\in\left(v_i\cap v_j\right)} \mathrm{Min}\left(v_i(t).F_q, v_j(t).F_q\right)}{\sum_{t\in\left(v_i\cap v_j\right)} \mathrm{Max}\left(v_i(t).F_q, v_j(t).F_q\right) + \sum_{t\in\left(v_i\cup v_j - v_i\cap v_j\right)} \mathrm{Min}\left(v_i(ta).F_q, v_j(ta).F_q\right)} \quad (6)$$

According to this formula, the u_i vector including the name of tag and the frequency of its usage is considered for each item. In this formula, F_q represents the frequency of using tag t for item i. Finally, based on the similarity between users and items, the items that the target user may select in the future are estimated. This is done according to the following algorithm:

Algorithm: Generate Top-N Recommendation List

```
1- Avg. = null;
2- Top-N;
3- InterestItem = null;
4- FOREACH v as neighbor of u
     FOREACH item i in v's item list
     FOREACH item j in u's item list
     InterestItem [i].ADD (j, SimItemi, j *Simu,v);
     Avg. ADD (i, InterestItem [i].AVG ());
5- FOREACH gi as Avg.Groupby(i)
     Top-N.ADD (gi.Key, gi.Max ());
6- Top-N.Sort();
     RETURN Top-N.Select (Item);
```

According to the algorithm, the item similarity is used to find and offer the most appropriate recommendations.

3.2. Evaluation metrics

The aim of this study is to improve the accuracy of the recommender systems, and to provide the recommendations tailored to the needs and interests of the target user. Therefore, the necessary parameters for evaluating the recommender system are determined in order to measure the accuracy of the system. In order to evaluate the quality of a recommender system, its productive results can be compared with the user comments on various items, and its accuracy is determined based on the error between the predicted values and the real views of the users. In this study, the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Top_N criteria are used to measure the proposed system. The most common method for evaluating a recommender system is the MAE criterion. MAE is *"the average vertical distance between each point and the identity line"* [32, p. 268], in which the difference between the rate predicted by the system and the actual rating is computed using Equation 7.

$$\mathrm{MAE} = \frac{\sum_{(u,i)}\left|\hat{r}_{u,i} - r_{u,i}\right|}{N} \quad (7)$$

Where $\hat{r}_{u,i}$ is the predicted rate by the system for the user u and item i, $r_{u,i}$ is the actual rating of user u to item i, and N is the number of rows in the evaluated set of evaluation.

RMSE is a widely used method for the performance evaluation of the recommendation system models. It is a standard prediction score that uses the square root [32], and is calculated using Equation 8, in which $r_{u,i}$ is the binary variable that shows the available data, whether item i has been rated by user u or not.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{(u,i)|R_{u,i}}\left(\hat{r}_{u,i} - r_{u,i}\right)^2}{N}} \quad (8)$$

The lower RMSE and MAE values of the algorithm are the higher accuracy of recommendations.

3.3. Experimental results

This phase starts with the initial data collection and includes the activities to identify the data quality problems, discover the nature of the data, find the interesting subsets that form the hidden information hypotheses, and prepare the data for processing. Then the recommender system is designed.

- ***Datasets***

In this phase, the required data is extracted and used as the input of the proposed algorithms. Datasets in which the relationship between the user data and time is clearly stated are not available in time–based recommender systems. Therefore, in this study, the social network data is used. The data used in the proposed system is the favorite movies at specified time intervals and

interest in videos. The required data is gathered in two steps, collecting the user profile information on social networks and transferring the social networking data to a questionnaire.
The users of the proposed system were required to use their Facebook social network information to fill the questionnaire. Since the users usually had a limited ability to see their profile information, it was necessary for the users who had an unrestricted access to their profile information to complete this step. This approach was employed to provide an accurate information. Therefore, using a questionnaire was essential. A questionnaire was distributed online using the snowball method among the members of social networks. In this approach, the researchers ask certain members of a social network to complete a questionnaire and send it to their friends to complete and submit to the researcher.
The questionnaire was completed by thousands of active Facebook users who met the required conditions, and the gathered information was used as the input database for the proposed algorithm. The collected dataset included a total of 1,218 records. After deleting the incomplete records, a dataset of 1,000 permissible user opinions about 400 items was used. Then according to the data, people who had seen the same movies were placed in a virtual recommended friendship group, and based on this information, new movies that the users of the same friendship group had not already seen were recommended. Moreover, the people of the same age were classified in a friendship group based on their favorite movies. For example, the 20-year old users with the same interests of the 30-year old users were classified in the same friendship group in the age range of 15-20 years (high number of the same films had been seen in this group) and movies that the 30-year old users had observed at the age of 20-25 years (action, science fiction, crime, etc.) were recommended to the 20-year old users. Then a fuzzy general procedure was considered to obtain the friendship and give the recommendations.
The group membership is obtained from the type of movie. For this purpose, the people in different age categories with more interest on certain categories of movies will be in a group through a fuzzy procedure. The age interval within groups was fixed in five years. The participants aged 5–10 were provided movie suggestions, especially cartoons, as well as the opportunity to see their father's favorite cartoons. Friendship is computed based on the type of film and the similarity of films. For example, selecting the same three films during a specified period indicated a high similarity; using fuzzy logic, this seemed to be logical. Another criterion is the tag collected by the distributed forms and stored in the database. Data preparation included all activities that contributed to the construction of the final dataset. The data preparation tasks are usually performed for several times. The tasks for this study included selecting the records and features as well as data transformation and cleaning using data modeling tools. Data preparation includes the process of giving a unique ID number to each user (User_Id) and each item (Item_Id), which in this research work is a film. The purpose of data preparation is to prepare, clean, and extract the information required for testing. A review table is created based on a main table for maintaining user comments and included User_Id, Date, Item_Id, Itm_Group, and Itm_Ratings. Similar rows are removed from the review table. Additional changes include standardizing date display format, replacing the name of each item with an equivalent code in the item field, removing invalid ratings, replacing the name of each user with an equivalent code in the user field, and creating two tables of items and users to store user information and items extracted from the review table. Further data preparation involves creating a ratings table containing the information required for testing and inserting the refined information into the review table; summarizing all movies viewed by a user along with scores and history; sorting the table based on the user name and item; creating a friends table from the ratings table to capture the friendships between individuals; and finally, creating a groups table from the ratings table to specify the age groups with which people were associated.

- ***Performance evaluation***

Since the study [4] did not calculate RMSE and MAE for its proposed method, it was not possible to compare the proposed method with the method of Jin *et al*. [4] in terms of RMSE and MAE. However, the results of the proposed approach compared with the basic algorithms are presented in Table 2. According to this table, the proposed method produces better results in comparison to the basic algorithms, which reflect the positive impact of the time factor combined with fuzzy logic in the proposed algorithm and the higher accuracy of the proposed method in recommendations to the users. This indicates that the fuzzy algorithm combined with information from social networks such as friends, membership, and tags improves the accuracy of the recommender systems.

Table 2. The Proposed approach compared with the basic algorithms.

Method	RMSE	MAE
Proposed method	**0.9114**	**0.7130**
BipolarSlopeOne	0.9432	0.7324
FactorWiseMatrixFactorization	0.9263	0.7321
BiasedMatrixFactorization	0.9137	0.7220
SVDPlusPlus	0.9156	0.7198
ItemKNNPearson	0.9141	0.7167

• ***Top_N results***

The average accuracy rate (Top_N) of the proposed method compared to the method described by Jin *et al*. [4] has been shown over multiple experiments. Since the user preferences change over time, this research work used a model of groups of users with five-year intervals of items-time, which can provide detailed recommendations in a mobile business environment. The quantitative criteria (RCL) (Formula 9) and accuracy (PRC) (Formula 10) are adjusted to evaluate the Top_N criteria, as in the field of data recovery.

$$\mathrm{RCL}_u = \frac{\mathrm{L}_u^+ \cap \mathrm{I}_u^R}{\left|\mathrm{I}_u^+\right|} \tag{9}$$

$$\mathrm{PRC}_u = \frac{\mathrm{L}_u^+ \cap \mathrm{I}_u^R}{\left|\mathrm{I}_u^+\right|} \tag{10}$$

The variable I_u^+ represents the items used by the user u in the simulated test set, and I_u^R is the item recommended to the user u.

For doing experiments, on average, 10 recommendations were considered for the users. However, the given experiments were also done for 20, 40, and 60 recommendations. The results of comparing the proposed method with those used in the study of Jin *et al*. [4] for 10, 20, 40, and 60 recommendations is presented in Table 4. As indicated in this table, at least 10 repeat in the proposed algorithm on database indicates the superiority of the proposed method in most recommendations.

Table 3. Comparison of the proposed method with the Jin *et al*. [4] method.

Recommendations / Method	10	20	30	40	60
Proposed method	0.261	0.254	0.265	0.262	0.262
Jin *et al*. [4]	0.267	0.250	0.268	0.268	0.280

Moreover, the error rate of the proposed algorithm compared with the algorithm of Jin *et al*. [4] is indicated in Figure 1. According to this figure, the error rate of the proposed method is lower than that of [4], which indicates the good accuracy of the proposed method. As a whole, the results indicate the usefulness of fuzzy systems in the process of friendship and group memberships.

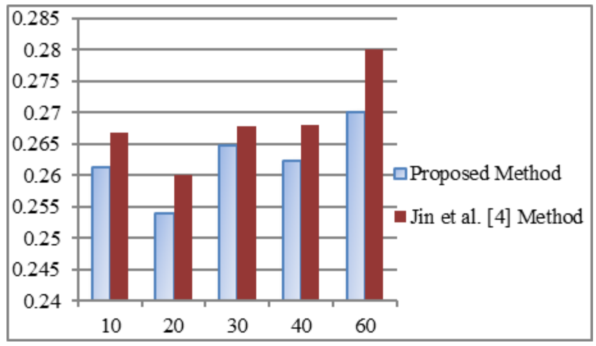

Figure 1. Error rate of the proposed algorithm compared with algorithm of Jin *et al*. [4].

4. Conclusion and recommendations

Regarding the social networks' development, and given the importance of people's ideas on the decisions of others with regard to using or not using a product or service, business owners use the recommender systems to increase the customer satisfaction and profits from providing goods or services. Social networks and trust affect the opinions of other users of a commercial/service site. Attention to the psychological aspects of people's tendency to rely on collective wisdom has opened a new horizon in the production and development of the recommender systems in the form of "social-network-based recommender systems." Therefore, using data from social networks, along with the time factor, which represents the old and forgotten interests of the users through the fuzzy system, this study tried to improve the evaluation criteria of the recommender systems in order to provide better results. Using fuzzy logic increased the accuracy of decision-making under uncertainty conditions. This research work, by studying a number of models in the field of social-network-based recommender systems, tried to take an effective step in the development of one of them. Therefore, because of the acceptable results of the method of Jin *et al*. [4] in comparison with the others, especially techniques based on social networks, it was used as the reference method in this study: a recommender system was designed by combining fuzzy logic, longitudinal data, and social network information such as tags, friendship, and membership in groups. The MAE, RMSE, and Top_N criteria were used to measure the proposed algorithm. In order to evaluate the efficiency of the proposed method, the results obtained were compared with those of basic algorithms. According to the results, the proposed method had better results than most of these algorithms, which reflected the higher accuracy of the proposed method in making recommendations to the users. Moreover, the results indicate that due to the appropriate separation in the process of computing the similarity based on friendship, group membership, and tags, the use of fuzzy systems has increased the efficiency of this method. Since the users' preferences will change over time, a weighting function based on the user-item and time was used to show the importance of their recent behaviors. According to the results, it is clear that the hybrid recommender systems based on the social networks' information and memory-based fuzzy systems have a better performance in predicting the score of a particular item than other methods. Better results on all users confirm the fact that the more information about the history, tastes, opinions, and previous ratings are available, the better performance the methods and techniques will have. In addition, the use of fuzzy systems in different choices can improve the options.

The purpose of the proposed method was to improve the limitations of the previous algorithms; however, this work has some limitations that can be the basis for future works. In this work, the experimental method was used to calculate the parameters of the algorithm. Therefore, it is suggested that future research works use formulas and do further calculations to determine the exact values of α and β, the optimum parameters and factors affecting the values of α and β, and to review the results of change in different datasets. In addition, it is recommended that future research works study the more demographic variables such as gender, nationality, and cultural characteristics in reviewing the selection of favorite films.

Regarding the fact that the users' behaviors change over time, it is suggested that future research works examine these changes and their effects on the users' future choices. In addition, there are other methods such as neural networks and other algorithms like learning automata, for which future research works can compute effectiveness to improve learning, provide more complete recommendations, and compare results with the proposed method of this work.

In order to promote and develop research and improve the model proposed in this work, it is suggested that future research works control the performance of the proposed model using the other datasets provided by commercial websites and compare the results to determine a future strategic plan.

In this work, friendship was computed using the type of film and the similarity of the films, for example, the same three films during the specified period using the fuzzy system. Therefore, other social network information can be used to specify the friendship with less weight and better results. It is also recommended that future research works use SVD, combined with social network information through fuzzy logic, to enhance the accuracy of selecting friends based on the users' favorite items and increase the efficiency of the proposed approach.

References

[1] Phulari, S. S., Khamitkar, S. D., Deshmukh, N. K., Bhalchandra, P. U., Lokhande, S. N., & Shinde, A. R. (2010). Understanding formulation of social capital in online social network sites (SNS). International Journal of Computer Science Issues, vol. 7, no. 1, pp. 92-96.

[2] Liao, C. L., & Lee, S. J. (2016). A clustering based approach to improving the efficiency of collaborative filtering recommendation. Electronic Commerce Research and Applications, vol. 18, no. 16, pp.1-9.

[3] Fasli, M. (2007). Agent technology for E-commerce. London: Wiley.

[4] Jin, X., Zheng, Q., & Sun, L. (2015, March). An Optimization of Collaborative Filtering Personalized Recommendation Algorithm Based on Time Context Information. In International Conference on Informatics and Semiotics in Organisations, Springer, Cham, Reading, United Kingdom. pp. 146-155.

[5] Lathia, N., Hailes, S., & Capra, L. (2009, July). Temporal collaborative filtering with adaptive neighbourhoods. In Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, Boston MA, USA. pp. 796-797.

[6] Bao, J., Zheng, Y., Wilkie, D., & Mokbel, M. (2015). Recommendations in location-based social networks: a survey. GeoInformatica, vol. 19, no. 3, pp. 525-565.

[7] Naseri, S., Bahrehmand, A., & Ding, C. (2015). An Improved Collaborative Recommendation System by Integration of Social Tagging Data, In: Ulusoy, O., Tansel A.U. & Arkun, E. (Eds.), Recommendation and Search in Social Networks. Springer International Publishing. pp. 119-138.

[8] Herlocker, J. L., Konstan, J. A., Borchers, A., & Riedl, J. (1999, August). An algorithmic framework for performing collaborative filtering. In Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, Berkeley California, USA. pp. 230-237.

[9] Esmaeili, L., Nasiri, M., & Minaei-Bidgoli, B. (2011, September). Personalizing group recommendation to social network users, In: Gong, Z., Luo, X., Chen, J., Lei, J. & Wang, L.F (Eds.), Web Information Systems and Mining. Springer Berlin, Heidelberg, 2011, pp. 124-133.

[10] Chen, T., Han, W. L., Wang, H. D., Zhou, Y. X., Xu, B., & Zang, B. Y. (2007, August). Content recommendation system based on private dynamic user profile. International Conference on Machine Learning and Cybernetics, Honk Kong, 2007, pp. 2112-2118.

[11] Liang, T. P., Hu, P. J., Kuo, Y. R., & Chen, D. N. (2007). A web-based recommendation system for mobile phone selection. The Pacific-Asia Conference on Information Systems. Auckland, Newzealand, 2007.

[12] Choi, K., Yoo, D., Kim, G., & Suh, Y. (2012). A hybrid online-product recommendation system: Combining implicit rating-based collaborative filtering and sequential pattern analysis. Electronic Commerce Research and Applications, vol. 11, no. 4, pp. 309-317.

[13] Chu, W., & Park, S. T. (2009, April). Personalized recommendation on dynamic content using predictive bilinear models. In Proceedings of the 18th international conference on World Wide Web, Madrid, Spain, pp. 691-700.

[14] Alhoori, H., & Furuta, R. (2017). Recommendation of scholarly venues based on dynamic user interests. Journal of Informetrics, vol. 11, no. 2, pp. 553-563.

[15] Koohi, H., & Kiani, K. (2017). A new method to find neighbor users that improves the performance of collaborative filtering. Expert Systems with Applications, vol. 83, no. C, pp. 30-39.

[16] Kazienko, P., Musial, K., & Kajdanowicz, T. (2011). Multidimensional social network in the social recommender system. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, vol. 41, no. 4, pp. 746-759.

[17] Chiu, Y. S., Lin, K. H., & Chen, J. S. (2011, December). A social network-based serendipity recommender system. In 2011 International Symposium on Intelligent Signal Processing and Communications Systems (ISPACS), Chaing Mai, 2011, pp. 1-5.

[18] Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: a survey. Decision Support Systems, vol. 74, no. 2015, pp.12-32.

[19] Guy, I., Zwerdling, N., Ronen, I., Carmel, D., & Uziel, E. (2010, July). Social media recommendation based on people and tags. In Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval, Geneva Switzerland, pp. 194-201.

[20] Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. Expert systems with applications, vol. 37, no. 7, pp. 4772-4778.

[21] Lee, D. H., & Brusilovsky, P. (2017). Improving personalized recommendations using community membership information. Information Processing & Management, vol.53, no. 5, pp. 1201-1214.

[22] Zheng, N., & Li, Q. (2011). A recommender system based on tag and time information for social tagging systems. Expert Systems with Applications, vol. 38, no. 4, pp. 4575-4587.

[23] Adeli, S., & Moradi, P. (2020). QoS-based Web Service Recommendation using Popular-dependent Collaborative Filtering. Journal of AI and Data Mining, vol. 8, no. 1, pp. 83-93.

[24] Halpin, H., Robu, V., & Shepherd, H. (2007, May). The complex dynamics of collaborative tagging. In Proceedings of the 16th international conference on World Wide Web, Banff Alberta, Canada, pp. 211-220.

[25] Heymann, P., Koutrika, G., & Garcia-Molina, H. (2008, February). Can social bookmarking improve web search? In Proceedings of the 2008 International Conference on Web Search and Data Mining, New York, United States, pp. 195-206.

[26] Sen, S., Vig, J., & Riedl, J. (2009, April). Tagommenders: connecting users to items through tags. In Proceedings of the 18th international conference on World Wide Web, Madrid, Spain, pp. 671-680.

[27] Durao, F., & Dolog, P. (2012). A personalized tag-based recommendation in social web systems. Journal CEUR Workshop Proceedings, vol. 485, pp. 40-49, arXiv preprint arXiv:1203.0332.

[28] Baltrunas, L., & Amatriain, X. (2009, October). Towards time-dependant recommendation based on implicit feedback. In Workshop on context-aware recommender systems (CARS 2009) in ACM Recsys. New York, USA, 2009.

[29] Wang, J., & Zhang, Y. (2013, July). Opportunity model for e-commerce recommendation: right product; right time. In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, Queensland, Australia, pp. 303-312.

[30] Ding, Y., Li, X., & Orlowska, M. E. (2006, January). Recency-based collaborative filtering. 17th Australasian Database Conference, Tasmania, Australia, 2006, pp. 99-107.

[31] Braunhofer, M., Kaminskas, M., & Ricci, F. (2011, October). Recommending music for places of interest in a mobile travel guide. In Proceedings of the fifth ACM conference on Recommender systems, Illinois, USA, pp. 253-256.

[32] Davagdorj, K., Park, K. H., & Ryu, K. H. (2020). A Collaborative Filtering Recommendation System for Rating Prediction, In Pan, J.S., Li, J., Tsai, P.W., & Lakhmi, C.J. (Eds.), Advances in Intelligent Information Hiding and Multimedia Signal Processing. Springer, Singapore, pp. 265-271.

نشریه هوش مصنوعی و داده کاوی

استفاده از اطلاعات شبکه‌های اجتماعی و زمان برای بهبود دقت سیستم‌های پیشنهاددهنده

بیژن حسن پور[۱]، ندا عبدالوند*[۲]، سعیده رجائی هرندی[۳]

[۱] دانشکده مهندسی برق، رایانه و فناوری اطلاعات، دانشگاه آزاد اسلامی قزوین، قزوین، ایران.

[۲] گروه مدیریت، دانشکده علوم اجتماعی و اقتصادی، دانشگاه الزهرا، تهران، ایران.

ارسال ۲۰۱۸/۰۸/۰۱؛ بازنگری ۲۰۱۹/۰۷/۲۱؛ پذیرش ۲۰۲۰/۰۳/۳۱

چکیده:

با پیشرفت سریع فناوری، اینترنت و افزایش تجارت الکترونیک، سیستم‌های پیشنهاددهنده به وجود آمدند که در یافتن و انتخاب آیتم‌های موردنظر، به کاربران کمک می‌کنند. دقت پیشنهاد در سیستم‌های پیشنهاد دهنده، یکی از چالش‌های اساسی این سیستم‌ها می‌باشد. با توجه به قابلیت سیستم‌های فازی در مشخص کردن مرز علایق کاربران، ترکیب آن‌ها با اطلاعات شبکه‌های اجتماعی و فاکتور زمان منطقی به نظر می‌رسد. لذا در این پژوهش، برای اولین بار با استفاده از منطق فازی و ترکیب اطلاعات زمان و اطلاعات شبکه‌های اجتماعی از جمله برچسب، دوستی و عضویت در گروه‌ها، کارآیی سیستم‌های پیشنهاددهنده مورد ارزیابی قرار گرفته است. سپس، با مشخص کردن همسایگی و مرز بین علائق کاربران در طول زمان، تأثیر الگوریتم پیشنهادی در بهبود دقت سیستم‌های پیشنهاددهنده مورد بررسی قرار گرفت. نتایج نشان دادند که استفاده از اطلاعات زمان و اطلاعات شبکه‌های اجتماعی در سیستم‌های پیشنهاددهنده مبتنی بر حافظه، باعث بهبود دقت این سیستم‌ها می‌شود.

کلمات کلیدی: سیستم‌های پیشنهاددهنده، شبکه‌های اجتماعی، اطلاعات زمان، منطق فازی، برچسب، هم‌گروهی.