

A proposed approach to discover nearest users on social media networks based on users' profiles and preferences

Mahmood Shakir Hammoodi¹, Ahmed Al-Azawei²

¹Computer Center, University of Babylon, Hillah, Iraq

²Department of Software, College of Information Technology, University of Babylon, Hillah, Iraq

Article Info

Article history:

Received Jul 18, 2022

Revised Sep 30, 2022

Accepted Nov 3, 2022

Keywords:

Data mining

Information systems

Knowledge-based systems

Social media networks

ABSTRACT

Social media sites (SMSs) become essential platforms used by people, customers, and companies for communication and marketing. Social media networks (SMNs) allow access to individuals who share their information and this, in turn, can lead to build users' profiles. Profiles can consist of a basic description of users' characteristics such as name, age, gender, education, marital status, email, phone number, and location. Preferences, on the other side, describe users' behavior on SMNs. Earlier literature defined a user's identity in different networks based on matching his/her name only. This research, however, proposes an integrated approach for discovering the nearest users by considering profiles and preferences. The proposed approach includes three key steps. First, users are grouped based on their preferences such as personal interests. Second, properties' values of a user profile and preferences are integrated to identify the nearest users. Finally, the nearest users to a certain user are identified by measuring the boundary distance. The findings show that the proposed approach can effectively identify the nearest users. Comparing the performance of the proposed approach with the two highly adopted approaches that rely on either users' profiles or preferences suggests that the proposed approach in this research is less prone to error.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Ahmed Al-Azawei

Department of Software, College of Information Technology, University of Babylon

Hillah, Babel, Iraq

Email: ahmedhabeeb@itnet.uobabylon.edu.iq

1. INTRODUCTION

The internet becomes an integrated platform in which one of its most intuitive aspects is the commercialization of social media networks (SMNs) [1], [2]. Many platforms have been established since the release of Web 2.0 such as media sharing (e.g., YouTube), blogs (e.g., Twitter), wikis (e.g., Wikipedia), social news (e.g., Reddit), and social networking (e.g., Facebook). Such platforms are used for content generation, information dissemination, and interactive communications [3]. SMNs are widely used in different fields of people's life such as health, education, marketing, and business. Such networks have become habits in many work areas [4]. Recently, business takes the advantage of SMNs to increase social interaction and drive growth [5]. People may share their ideas, perspectives, and views on several different topics with other users using these platforms. This may indicate that a huge amount of data is generated on social media sites (SMSs) where millions of users communicate with each other using these networks [6]. Networks refer to a set of relationships between a group of customers, suppliers, agents, and similar social connections [5]. SMNs allow access to individuals (customers, users, or companies) who share information about their social properties such as images, videos, and texts. This can increase their popularity and obtain

reactions and perspectives of other people [4], [7]. Users of SMNs also share their ideas, daily activities, and opinions [8].

The extended network includes a user's friends of friends, although there is no direct relationship between them [9], [10]. A user can create a profile on SMNs such as Facebook for free to communicate with different people [11]. A user's profile is considered a social identity. A profile consists of a basic description about the user such as name, age, gender, education, marital status, email, phone number, and location. Earlier literature exploits such kind of information to provide recommendations for users [12]–[14].

Information that may describe users' behavior on SMNs is called users' preferences such as accepting new friends, interested in sports news, sharing their posts, looking for new friendships, searching for a new job, and accepting requests from other people. Users may share similar preferences and interests in SMNs which could be analyzed to discover the nearest users [15]–[17]. However, searching for people, products, industries, businesses, or jobs is not a trivial task due to the huge amount of data generated on SMNs every day [7]. For example, if a user searches for a person nearby who shares a similar profile and/or preferences, the user needs to search posts and users' profiles. On the other hand, SMNs may suggest nearby people based on several different features of similarity.

Although the suggested results in earlier literature could be an easy way and with less effort to generate an individual network, a user may not have the willingness to interact with such results. Users are probably unfamiliar with the process of suggesting the nearest people through SMNs which may be suggested based on either profile or preferences. For example, it can be assumed that there is matching between a group of users in terms of age, gender, and education, but each user has his/her own preferences such as preferring a particular football team, accepting new friendships, and allowing to share his/her posts. Thus, it might be difficult to group users together based on these differences. To address this issue, users' profiles and preferences can be exploited together to search for the nearest users. This study, therefore, contributes to the body of research by proposing an approach for discovering the nearest users based on users' profiles and preferences. Proposing this hybrid approach can address the limitations in previous literature and provide more reliable results. It can also help recommend items that fit users' needs based on the preferences of other close users.

The rest of the paper is organized as follows. In section 2, previous approaches are reviewed and analyzed. In section 3, the proposed approach is presented. Section 4 presents the key findings of this study. Section 5 discusses the research findings, whereas section 6 concludes the key concepts of this research and the possible future directions.

2. RELATED WORKS

A review of previous approaches that have been used for discovering and matching users' profiles in SMNs is presented here. The discussed literature is grouped into two categories namely, approaches that relied on users' profiles and approaches that relied on users' preferences. About the first category, according to Raad *et al.* [18] a framework for matching users' profiles that refer to the same person on two SMNs was proposed. The suggested approach included two phases: i) profiles were converted into an extensible markup language (XML) file and ii) a set of profiles' attributes were compared using the similarity function. If the similarity score of two profiles is higher than a profile matching threshold, this means that both profiles are for the same user. In another research study [19], an approach for matching profiles was suggested to discover if different profiles belong to the same user. Screen names on Facebook were analyzed by removing nicknames, converting names into lowercase, splitting spaces, and matching the extracted names to the corresponding screen names on Twitter. According to Thelwall [20] an approach for analyzing users' profiles was used where it focuses on users' ages and gender more than other features such as location, status (i.e., married, single), and the number of registered friends. Although such frameworks were able to identify profiles on two SMNs that refer to the same user, users' preferences are ignored. Another research study exploited users' profiles on SMNs to address the cold start issue as users' demographic data was considered such as gender, age, and occupation [21]. In such context, [22] also used demographic information of users to compute the similarity weight of the users' behavior values and their rating value. In another research by Senefonte *et al.* [23], users' profiles on social media sites were considered to predict their mobility as tourists. In the proposed approach, users were grouped using fuzzy C-means and self-organizing maps.

About the second category, [24] investigated users' preferences based on their location to recommend, for example, the most appropriate restaurants. The study automatically identified user preferences based on their location history and opinions. Similarly, according to Bao *et al.* [25] a systematic review was provided on investigating the methods used in recommender systems based on SMNs. According to Jamil *et al.* [26] an approach was proposed for identifying the nearest users by measuring the similarity of preferences. Users with the same preferences were grouped and recommended to a particular user. According to Kim *et al.* [9] an approach to suggest similar users based on preferences was also suggested. This approach

consists of three tasks which are: i) clustering user's preferences; ii) measuring similarity between two users on a scale ratio; and iii) providing suggestions. Another research study proposed a model based on measuring the similarity between users' preferences [15]. Finally, according to Torrijos *et al.* [27] a method was suggested for discovering related users in terms of location-based social networks. The method assumes that users are more similar when they visit the same place at the same time and share similar preferences. However, the properties of users' profiles are ignored in such studies.

It can be seen that the reported literature did not take into consideration both users' profiles and users' preferences. In addition, some of the reported approaches are either designed for identifying the same user from different networks or for matching profiles in terms of users' names only. However, users' age, gender, education, and preferences are considered essential properties of a user which have to be taken into consideration when discovering the nearest users to a particular user [28], [29]. This research attempts to address this issue in previous work by proposing an approach for discovering the nearest users based on both profiles and preferences.

3. THE PROPOSED APPROACH

The proposed approach consists of three key stages. First, users are grouped based on their preferences. Second, the nearest users are identified by measuring the boundary distance of the grouped users using properties' values of profiles and preferences. Finally, the best nearest group of users are indicated through measuring similarity between them.

3.1. The research datasets

Two datasets were used which are the Book-Crossing dataset [30] and the personality dataset [31]. The Book-Crossing dataset consists of 278,858 users with 1,149,780 ratings, as well as 271,379 books. It was generated with three main tables namely, BX-Users, BXBooks, and BX-Book-Ratings.

The BX-Users table consists of userID and age to represent users' profiles. The BX-Books table includes ISBN, book-title, book-author, year-of-publication, and publisher in which ISBN represents a book identity. BX-Book-Ratings table encompasses users' preferences, book rating, userID, and ISBN. The dataset is available in two formats structured query language (SQL) dump [26.391 Kb] and comma separated value (CSV) dump [25.475 Kb]. In this study, the SQL dump was downloaded and converted into SQL server database where schema information and data insertion statements are found in SQL dump which can be run as a SQL script.

Ten users were selected randomly from the dataset in which the proposed approach was implemented. However, the results of only four users are reported in this research as an example of the performance of the proposed model. Table 1 shows a description of the selected users from the Book-Crossing dataset. The users' profiles include userID and age, whereas the number of preferred books and the average rating is considered as users' preferences.

Table 1. A description of the selected users

User profile features		User preferences features	
userID	Age	Number of preferred books	Average rating
10	26	2	3
2,276	46	498	3
7,400	22	5	7
13,877	27	4	6

The personality dataset consists of 1,834 users with 1,028,751 ratings, as well as 198,117 movies. It was generated with two main tables namely, personality data and ratings. Personality data consists of userID and the five personality types which are scored from 1 to 7. The five personality types are openness, agreeableness, emotional stability, conscientiousness, and extraversion. Openness evaluates users' attitudes to choose a new experience. Agreeableness assesses users' behavior toward others such as compassion, cooperation, suspicion, and antagonism. Emotional stability evaluates users' attitudes to psychological stress. Conscientiousness evaluates users' behavior such as organization, dependability, and show self-discipline. Extraversion assesses users' attitudes to be outgoing [32]. The rating table contains the ratings of users. Ten users were also selected randomly from the dataset, but the results are reported here for four users only to show the implementation of the proposed approach. Table 2 presents a description of the selected users from the personality dataset.

Table 2. Description of the selected users from the personality dataset

UserID	Openness	Agreeableness	Emotional stability	Conscientiousness	Extraversion	Average rating
8e7ceb9a234c064b75016249f2ac65e	5	2	3	2.5	6.5	3.48
8f45de0d6ded99d3c81f32100d08ab85	3.5	4	5.5	2.5	4.5	3.65
abb3365a8b5a8ed98497110974d21089	5	4	3	3	2	3.50
cdc60474a84852f28f42441686869c05	7	6	3.5	2	2.5	3.79

3.2. Users' grouping

In SMNs, users post their views or interact with others by sharing information based on their preferences. In this stage, they are grouped based on their preferences such as desires or interests. A user may prefer watching action movies, family movies, or follow documentary reports. Other users may share similar preferences. Accordingly, the similarity of preferences is used for grouping users. For example, it can be assumed that user X prefers action movies and follows documentary reports, whereas user Y prefers family movies and follows documentary reports. Thus, $X \cap Y = 1$ preference based on the documentary reports.

Hence, X and Y can be grouped with the number of movies that they liked or tagged, as shown in Algorithm 1. It is worth mentioning that the number of preferences (i.e., X prefers six documentary reports) is not used for grouping users here. This is used in the next stage for measuring the boundary distance between the grouped users. G refers to a group of users who share similar preferences, whereas P refers to a preference. Hence, $X \cap Y = \{P1, P2, P3\}$. $X \cap Z = \{P8\}$. $X \cap W = \{\}$. Therefore, $G = \{X, Y, Z\}$.

Algorithm 1: Grouping users

```

Input: User X, List of Users L           5:   Py ←= {list of preferences of Y}
Output: List of Grouped Users           6:   if Px ∩ Py ≠ ∅ then
1: Px ←= {list of preferences of X}     7:     G ←= Y
2: G ←= ∅ list of grouped users         8:   end if
3: for each user in L do                 9: end for
4:   Y ←= L

```

3.3. The boundary distance of the grouped users

In this stage, the properties' values of a user profile and preferences are used to identify the nearest users. Although a user may have more than one profile in different SMNs, the properties of each profile are designed and assigned based on the terms and conditions of each social media network which differ from one network to another. In this research, the properties of a profile are used to identify the nearest users, whereas different profiles of same users on different SMSs are not considered.

The grouped users (the output of Algorithm 1) are analyzed by measuring the boundary distance to user X (1). A boundary distance between user X and the grouped users is represented as a radius of the circle and indicated as a threshold to identify which users are assigned within the radius boundary. Hence, the accuracy of discovering the nearest users may be improved using boundary distance [4], [6].

$$\text{Boundary Distance } (X, Y) = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{x})^2}{n} - \left(\frac{\sum_{i=1}^n (y_i - \bar{x})}{n}\right)^2} \quad (1)$$

where n refers to the number of features of user X and Y . \bar{x} refers to μ of features' values (profile properties and preferences) of user X (2). Each feature's value of Y is subtracted from the μ such as $(y_i - \bar{x})$.

$$\text{Boundary Center } (\bar{x}) = \sum_{i=1}^n \frac{v}{n} \quad (2)$$

where v refers to the features' values of user X . Y is ignored when it is assigned outside of the boundary distance. Users who are assigned inside the boundary distance are considered the nearest users to user X , as shown in Algorithm 2.

Algorithm 2: Nearest users

```

Input: User X, List of Grouped Users L   5:   Bdy ←= Boundary Distance between X and Y
Output: List of Nearest Users to X       6:   if Bdy ≤ Bdx then
1: N ←= ∅ list of nearest users to X     7:     N ←= Y
2: Bdx ←= Boundary of X                 8:   end if
3: for each user in L do                 9: end for
4:   Y ←= L

```

3.4. Measuring similarity of the nearest users

In the previous stage, the nearest users to X are identified by measuring boundary distance. Although nearest users who share preferences and properties' values of profiles may be obtained by using the boundary distance as described previously, measuring the similarity between X and the nearest users is required. For example, it can be assumed that user X prefers 20 documentary reports, whereas the nearest user Y prefers 20 family movies and only 1 documentary report. User Z , on the other hand, prefers 15 documentary reports. Hence, Z is identified as the best nearest user to X according to the similarity function which is known as cosine similarity (3):

$$\text{Similarity}(X, Y) = \frac{\sum_{i=1}^n (x_i * y_i)}{\sqrt{\sum_{i=1}^n (x_i)^2} * \sqrt{\sum_{i=1}^n (y_i)^2}} \quad (3)$$

Algorithm 3 shows the main steps to measure similarity between X and the nearest users (i.e., within the boundary of X) using features' values (profile's properties and preferences) which represent users' desires and interests. In Algorithm 3, a threshold t is used to evaluate the similarity between users. Where t is a user-defined in which it is set between 0.5 and 0.9.

```

Algorithm 3: Measuring similarity
Input: User X, List of Nearest Users L,      5:   Vy ←= {list of featureso values of Y}
threshold t                                6:   CSxy ←= Similarity between Vx and Vy
Output: List of Similar Users to X          7:   if CSxy ≥ t then
1: S ←= {∅} list of similar users to X      8:     S ←= Y
2: Vx ←= {list of featureso values of X}    9:   end if
3: for each user in L do                    10: end for
4:   Y ←= L

```

3.5. Evaluation

The performance of the proposed approach is evaluated based on precision. It measures the number of users with a similarity value greater or equal to a threshold (between 0.5 and 0.9) divided by the number of users as shown in (4):

$$P = C / (C + I) \quad (4)$$

where C refers to the number of users with a similarity value greater or equal to the threshold. I refer to the number of users with a similarity value less than the threshold.

4. RESULTS AND EVALUATION

The proposed approach is implemented using SQL and server management studio 2017 to store users' profiles and preferences and run queries. To evaluate the ability of the proposed approach for discovering nearest uses, three experiments were conducted which are labeled Exp1, Exp2, and Exp3. Each experiment uses the same set of data. The experiments were conducted on ten users from each dataset, but the results are reported for four users only as other results agree with the reported findings. Therefore, repetition has been avoided.

Exp1 refers to the use of the proposed approach using users' profiles and preferences. Exp2 refers to the application of the proposed approach using users' preferences only, whereas Exp3 refers to the implementation of the proposed approach using users' profiles only. The findings of these three experiments are compared.

4.1. Results of the Book-Crossing dataset

Table 3 shows the number of nearest users discovered with Exp1 and Exp2. The number of the discovered users with Exp2 is greater or equal to that discovered with Exp1. However, there is a possibility to discover users incorrectly according to the similarity function. Figure 1 presents a comparison between the outcomes of Exp1 and Exp2 in terms of the number of nearest users discovered correctly and incorrectly for each user in Table 1.

Based on the findings presented in Figure 1, the performance of the proposed approach with Exp1 is better than Exp2. Although the number of the discovered users with Exp1 is less than Exp2 in some cases, the number of the nearest users discovered incorrectly in Exp1 is significantly less than the number reported in Exp2 in most cases, supporting the validity of the proposed approach. Figure 2 (see in Appendix) shows a comparison between Exp1 and Exp2 for each user in terms of precision using the reported number of the

discovered users correctly and incorrectly (see Figure 1). The similarity measure reported higher precision between 0.5 and 0.9 in most cases with Exp1. The reason behind this is that discovering the nearest users requires both the profiles' properties and preferences.

Table 3. The number of nearest users with Exp1 and Exp2

UserID	Number of nearest users with Exp1	Number of nearest users with Exp2
10	1	2
2,276	2	2
7,400	1	3
13,877	3	14

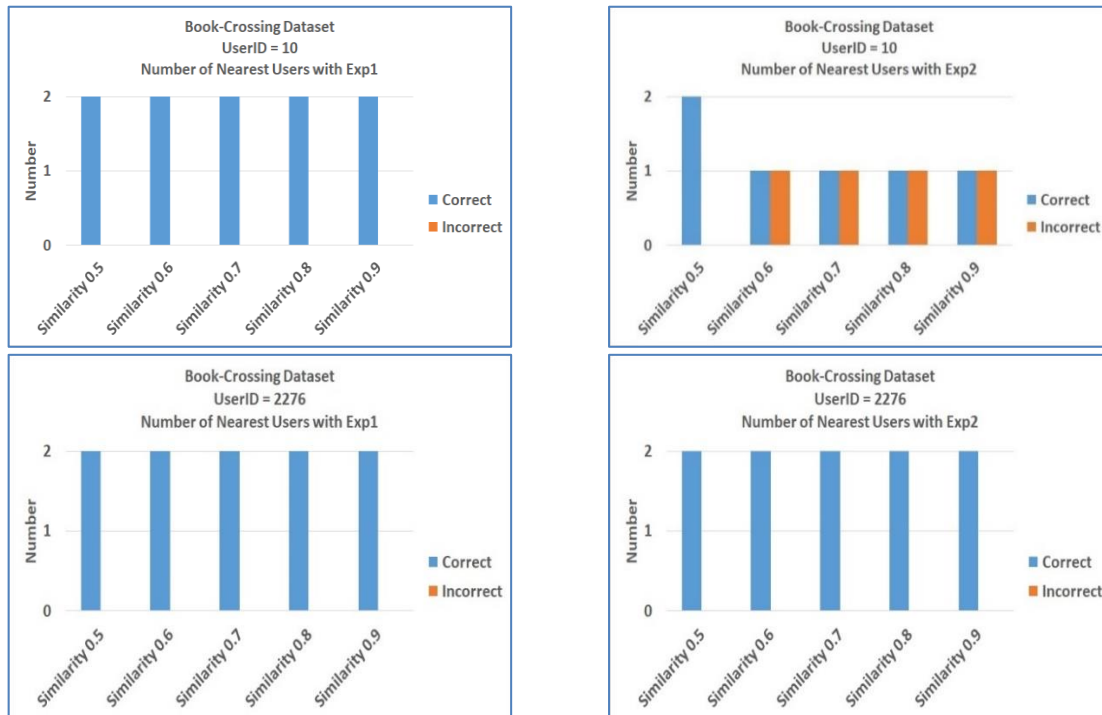


Figure 1. The number of nearest users of the first two users with Exp1 and Exp2

4.2. Results of the personality dataset

Table 4 shows the number of the nearest users discovered with Exp1 and Exp3. The number of the discovered users with Exp3 is greater or equal to that discovered with Exp1. However, there is a possibility to discover users incorrectly according to the similarity function, as shown in Figure 3 which presents a comparison between Exp1 and Exp3 in terms of the number of the nearest users discovered correctly and incorrectly for each user in Table 4.

In Figure 3, the performance of the proposed approach with Exp1 is better than Exp3. Although the number of the discovered users with Exp1 is smaller than Exp3 in some cases, the number of nearest users discovered incorrectly in Exp1 is less than the number reported in Exp3 in most cases. Figure 4 shows a comparison between Exp1 and Exp3 for each user in terms of precision using the reported number of the discovered users correctly and incorrectly (see Figure 3). It can be seen that the similarity measure reported higher precision between 0.5 and 0.9 in most cases with Exp1. This finding confirms that the proposed method is adequate in predicting the nearest users than relying on either users' profile features or users' preferences only.

Table 4. Number of nearest users with Exp1 and Exp3

userID	Number of nearest users with Exp1	Number of nearest users with Exp3
8e7cebf9a234c064b75016249f2ac65e	34	20
8f45de0d6ded99d3c81f32100d08ab85	75	94
abb3365a8b5a8ed98497110974d21089	16	38
cdc60474a84852f28f42441686869c05	81	89

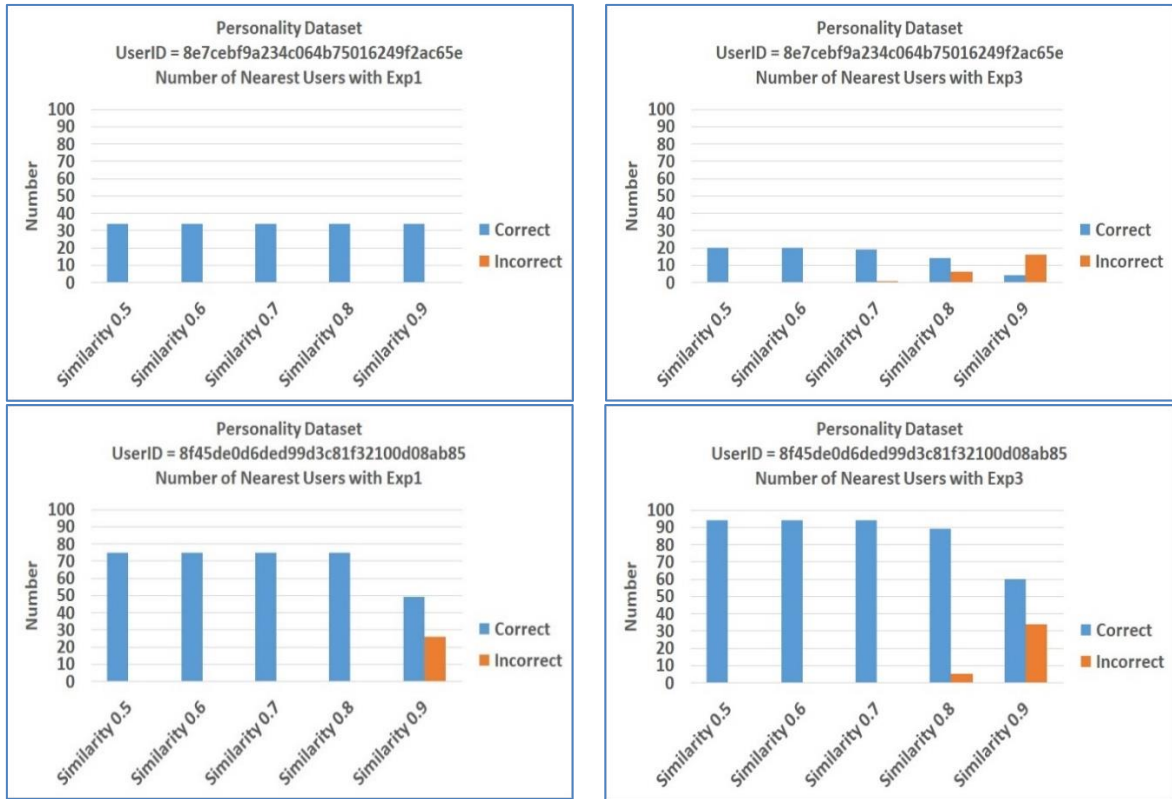


Figure 3. The number of nearest users of the first two users with Exp1 and Exp3

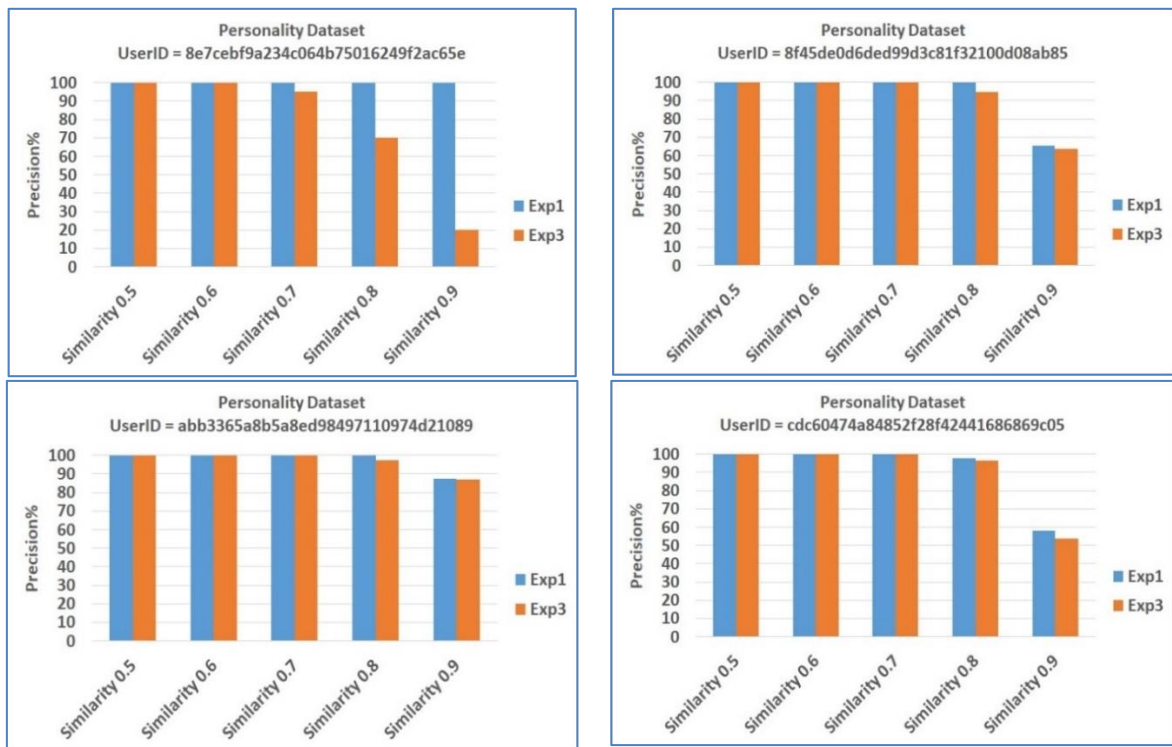


Figure 4. The results of the personality dataset for four users

5. DISCUSSION

Current literature develops several different algorithms to search for similar users based on either users' profiles or users' preferences. However, considering both is necessary as it provides more reliable results in identifying the nearest users. Some of the reported approaches are either designed for identifying the same user from different networks or matching profiles in terms of users' names only. This research, however, attempted at addressing this gap by proposing an approach that can discover the nearest users based on users' profiles and preferences.

The proposed approach relies on three key phases. First, users are grouped based on their preferences (i.e., desires or interests). Second, the properties' values of a user profile and preferences are used to identify the nearest users. Although some users may have more than one profile on different SMNs, the properties of each profile are designed and assigned based on the terms and conditions of each social media network which differ from one network to another. In this research study, therefore, properties of profiles are used to identify the nearest users, whereas the same or similar users are not considered. Third, the nearest users to a certain user are identified by measuring the boundary distance.

Although nearest users which share similar preferences and properties' values may be discovered using the boundary distance, measuring the similarity between the user and the nearest users are required. The performance of the proposed approach was evaluated concerning precision. It has been found that the proposed approach achieved higher accuracy in comparison with the approaches that rely on either users' profiles or users' preferences. In all cases, the number of incorrectly predicted nearest users in the proposed approach was less than the incorrectly predicted nearest users in Exp2 and Exp 3 which relied on users' profiles or preferences respectively.

The approach presented in [22] is compared with the proposed approach of this present research. According to Widiyaningtyas *et al.* [22] however, the approach identifies the nearest users by measuring the similarity of preferences only. Users with the same preferences were grouped and recommended to a specific user. The best precision result achieved in [22] was less than 60 where the similarity is greater or equal to 0.83. On the other hand, the proposed approach in this present research achieves better results in all aspects based on the integration of both profiles and preferences.

6. CONCLUSION

Based on the research outcomes, it is recommended to use both users' profile information and their preferences in SMNs or other personalized sites such as Amazon or customers' sites. This can provide better recommendations that may meet their actual needs effectively. This research is distinguished from previous works in that the proposed approach is capable for discovering the nearest users based on both profiles and preferences. Although the findings of the proposed approach are promising, future work in this field can be conducted in other directions. Further implementation of the proposed approach can be done by integrating Algorithm 1 with the k-means clustering algorithm to discover a group of the nearest users. Thus, the number of clusters can be identified dynamically. Another possible research direction is to evaluate the proposed approach using other datasets and implement it in several domains.

APPENDIX

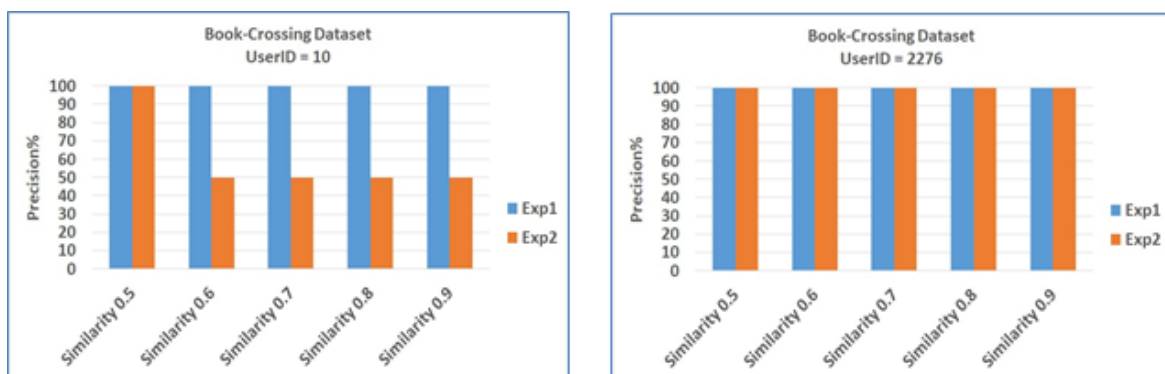


Figure 2. The results of the Book-Crossing dataset

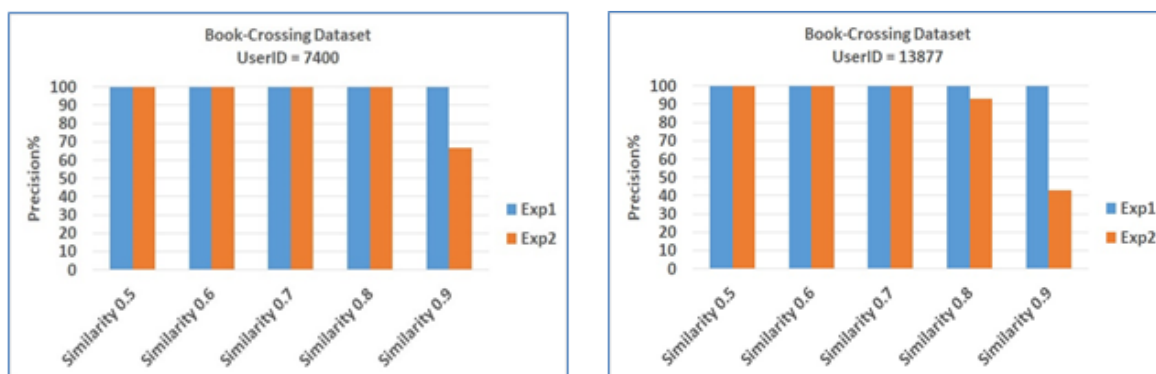


Figure 2. The results of the Book-Crossing dataset (continue)




REFERENCES

- [1] D. Zeng, H. Chen, R. Lusch, and S. Li, "Social Media Analytics and Intelligence," *IEEE Intell. Syst.*, vol. 25, no. 6, pp. 13–16, 2010.
- [2] B. J. Jansen, K. Sobel, and G. Cook, "Classifying Ecommerce Information Sharing Behaviour by Youths on Social Networking Sites," *J. Inf. Sci.*, vol. 37, no. 2, pp. 120–136, 2011.
- [3] C. S. Yang and L. C. Chen, "Personalized recommendation in social media: A profile expansion approach," *Proc. - Pacific Asia Conf. Inf. Syst. PACIS 2014*, 2014.
- [4] A. Tutgun-Ünal, "Social Media Addiction of New Media and Journalism Students," *J. Educ. Technol.*, vol. 19, no. 2, pp. 1–12, 2020.
- [5] S. Fraccastoro, M. Gabriellson, and S. Chetty, "Social Media Firm Specific Advantages as Enablers of Network Embeddedness of International Entrepreneurial Ventures," *J. World Bus.*, vol. 56, no. 3, p. 101164, 2021.
- [6] U. Can and B. Alatas, "A new direction in social network analysis: Online social network analysis problems and applications," *Phys. A Stat. Mech. its Appl.*, vol. 535, p. 122372, 2019.
- [7] V. G. U., S. K., P. Deepa Shenoy, and V. K. R., "An Overview on User Profiling in Online Social Networks," *Int. J. Appl. Inf. Syst.*, vol. 11, no. 8, pp. 25–42, 2017.
- [8] P. D. Shenoy, K. G. Srinivasa, K. R. Venugopal, and L. M. Patnaik, "Evolutionary Approach for Mining Association Rules on Dynamic Databases," *Adv. Knowl. Discov. Data Min.*, no. April, pp. 325–336, 2003.
- [9] Y. S. Kim, F. Natali, Y. Kim, F. Natali, F. Zhu, and E. Lim, "Investigating the influence of offline friendship on Twitter networking behaviors," in *49th IEEE Hawaii International Conference on System Sciences (HICSS)*, 2016, pp. 736–745.
- [10] E. D. Andrea, P. Ducange, B. Lazzarini, and F. Marcelloni, "Real-Time Detection of Traffic From Twitter Stream Analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 2269–2283, 2015.
- [11] C. Lorenzo-Romero, M.-C. Alarcon-del-Amo, and M.-A. Gomez-Borja, "Do you have social profile? Users and non-users of social networking sites in the Web 2.0," *Rev. Bus. Inf. Syst.*, vol. 15, no. 5, pp. 41–50, 2011.
- [12] S. Ahmadian, M. Afsharchi, and M. Meghdadi, "An effective social recommendation method based on user reputation model and rating profile enhancement," *J. Inf. Sci.*, pp. 1–36, 2018.
- [13] S. Ahmadian, M. Meghdadi, and M. Afsharchi, "A social recommendation method based on an adaptive neighbor selection mechanism," *Inf. Process. Manag.*, vol. 54, no. 4, pp. 707–725, Jul. 2018.
- [14] S. G. K. Patro *et al.*, "A Hybrid Action-Related K-Nearest Neighbour (HAR-KNN) Approach for Recommendation Systems," *IEEE Access*, vol. 8, pp. 90978–90991, 2020.
- [15] Y. Li, J. Liu, and J. Ren, "Social recommendation model based on user interaction in complex social networks," *PLoS One*, vol. 14, no. 7, pp. 1–17, 2019.
- [16] C. Wang *et al.*, "Behavior-interior-aware user preference analysis based on social networks," *Complexity*, vol. 2018, 2018.
- [17] A. Bahulkar, B. K. Szymanski, N. Chawla, O. Lizardo, and K. Chan, "Influence of personal preferences on link dynamics in social networks," *Complexity*, vol. 2017, 2017.
- [18] E. Raad, R. Chbeir, A. Dipanda, and E. Raad, "User profile matching in social networks," *Network-Based Inf. Syst.*, no. September, pp. 297–304, 2010.
- [19] H. Hazimeh, E. Mugellini, O. A. Khaled, and P. Cudré-Mauroux, "Socialmatching++: A novel approach for interlinking user profiles on social networks," *CEUR Workshop Proc.*, vol. 1927, 2017.
- [20] M. Thelwall, "Social networks, gender, and friending: An analysis of myspace member profiles," *J. Am. Soc. Inf. Sci. Technol.*, vol. 59, no. 8, pp. 1321–1330, 2008.
- [21] F. Tahmasebi, M. Meghdadi, S. Ahmadian, and K. Valiollahi, "A hybrid recommendation system based on profile expansion technique to alleviate cold start problem," *Multimed. Tools Appl.*, vol. 80, no. 2, pp. 2339–2354, 2021.
- [22] T. Widiyaningtyas, I. Hidayah, and T. B. Adji, "User profile correlation-based similarity (UPCSim) algorithm in movie recommendation system," *J. Big Data*, vol. 8, no. 1, pp. 1–21, Dec. 2021.
- [23] H. C. M. Senefonte, M. R. Delgado, R. Luders, and T. H. Silva, "PredicTour: Predicting Mobility Patterns of Tourists Based on Social Media User's Profiles," *IEEE Access*, vol. 10, pp. 9257–9270, 2022.
- [24] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data," in *In Proceedings of the 20th international conference on advances in geographic information systems*, 2012, pp. 199–208.
- [25] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: A survey," *Geoinformatica*, vol. 19, pp. 525–565, 2015.
- [26] N. J. Jamil, S. A. Mohd Noah, and M. Mohd, "Collaborative Item Recommendations Based on Friendship Strength in Social Network," *Int. J. Mach. Learn. Comput.*, vol. 10, no. 3, pp. 437–443, 2020.
- [27] S. Torrijos, A. Bellogín, and P. Sánchez, "Discovering Related Users in Location-based Social Networks," *UMAP 2020 - Proc.*




- 28th ACM Conf. User Model. Adapt. Pers., pp. 353–357, 2020.
- [28] D. Shwartz-Asher, S. Chun, N. R. Adam, and K. L. Snider, “Knowledge sharing behaviors in social media,” *Technol. Soc.*, vol. 63, p. 101426, 2020.
- [29] J. Vithayathil, M. Dadgar, and J. K. Osiri, “Social media use and consumer shopping preferences,” *Int. J. Inf. Manage.*, vol. 54, no. March, 2020.
- [30] C.-N. Ziegler, S. M. McNee, J. A. Konstan, and G. Lausen, “Improving recommendation lists through topic diversification,” in *Proceedings of the 14th International World Wide Web Conference (WWW '05)*, 2005, pp. 10–14.
- [31] T. T. Nguyen, F. Maxwell Harper, L. Terveen, and J. A. Konstan, “User Personality and User Satisfaction with Recommender Systems,” *Inf. Syst. Front.*, vol. 20, no. 6, pp. 1173–1189, 2018.
- [32] M. Komaraju, S. J. Karau, and R. R. Schmeck, “Role of the Big Five personality traits in predicting college students’ academic motivation and achievement,” *Learn. Individ. Differ.*, vol. 19, no. 1, pp. 47–52, 2009.

BIOGRAPHIES OF AUTHORS



Mahmood Shakir Hammoodi    received a B.Sc degree in computer science from the College of Science in 1998, and an M.Sc degree in computer science from University Putra Malaysia. He received a Ph.D degree in Computer Science from the University of Reading, UK. Currently, he is a Senior Lecturer at the University of Babylon. His research interests include data mining, data stream mining, big data analysis, and social media. He can be contacted at email: pre.mahmood.shakir@uobabylon.edu.iq.



Ahmed Al-Azawei    received a BSc degree in computer science from the College of Science in 1997, and an MSc degree in information technology from University Utara Malaysia. He received a Ph.D degree in Computer Science from the University of Reading, UK. Currently, he is an Assistant Professor in the Department of Software at the University of Babylon. His research interests include technology adoption, educational data mining, social media, and text mining. He can be contacted at email: ahmedhabeeb@itnet.uobabylon.edu.iq.