PersoNet: Friend Recommendation System Based on Big-Five Personality Traits and Hybrid Filtering

Huansheng Ning, *Senior Member, IEEE*, Sahraoui Dhelim, and Nyothiri Aung

***Abstract*— Friend recommendation system (FRS) is an essential part of any social network system. With the popularity of social network sites, many FRSs have been proposed in the past few years. However, most of them are homophily based systems, homophily is the propensity to associate and bond with similar others. In other words, these systems will recommend people that you share common features with them as friends. Homophily based FRS is accurate when the common feature is a physical or social feature, such as age, race, location, job, or lifestyle. However, it is not the case with personality types. Having a given personality type does not necessarily mean that you are compatible with people that have the same personality type. Therefore, in this paper, we present and evaluate an FRS based on the big-five personality traits model and hybrid filtering, in which the friend recommended process is based on personality traits and users' harmony rating. To validate the proposed system's accuracy, a personality-based social network site that uses the proposed FRS named PersoNet is implemented. Users' rating results show that PersoNet performs better than collaborative filtering (CF)-based FRS in terms of precision and recall.**

***Index Terms*— Big five, five-factor model (FFM), friend recommendation system (FRS), hybrid filtering, personality computing, social computing, social networks.**

I. Introduction

WITH more than 3 billion active users around the globe [1], social networking sites (SNSs) have become the main method of making new friends. It had been proven that friendship in SNS can better describe self-report friendship compared to friendship created by frequent physical encounters [2]. Each one of these social networks relies on a friend recommendation system (FRS) that is used to detect common features between two people and, consequently, connect them to each other. Many FRS have been proposed in the past few years, but most of them are based on homophily (the propensity of people to associate and bond with similar others). In other words, these systems will recommend people that have a common feature with you as friends. Homophily based FRS is adequate when the common feature is a physical or social feature, such as age, race, location, job, or lifestyle. However, when it comes to personality types, things are different. Personality-based FRS brings up a very old psychological debate about the personality similarities between friends. While most of the mainstream researchers argue that there is no similarity in personality between friends [3], [4]. Recent researchers have suggested that friends and couples indeed are similar in their personality [5]. In addition to that, a major challenge for FRS is known as the cold-start problem, where the recommendation system does not have enough information about the new user, and the missing information is crucial in the recommendation process. In this case, personality information can help to alleviate the cold-start problem.

For the above-mentioned reasons, in this paper, we present and evaluate an FRS based on the big-five personality traits model and hybrid filtering, in which the friend recommendation process is based on personality traits and users' harmony rating. To validate the proposed system's accuracy, a personality-based social network site that uses the proposed FRS named PersoNet is implemented. The proposed system not only enhances the prediction accuracy of recommendation systems but also alleviates the cold-start problem of the legacy collaborative filtering (CF) systems. To compare PersoNet with the legacy FRSs, we have implemented three recommendation systems and compared them based on their precision and recall values: 1) FRS based only on personality matching; 2) FRS based only on CF; and 3) the proposed system PersoNet, which is based on personality traits and hybrid filtering. Users' rating results show that PersoNet performs better than the other two FRS in terms of precision and recall. Our contributions can be summarized as follows.

1) Propose a personality-based FRS based on big-five personality traits and hybrid filtering.
2) Implement the proposed system in a social network site.
3) Conduct an online experiment using the implemented site to validate the robustness of PersoNet.

The rest of this paper is organized as follows. In Section II, we present preliminaries about the topic's background. Section III reviews the recent related works. In Section IV, we present the system modeling details. While in Section V, experiment details of PersoNet are presented. In Section VI, we discuss and evaluate the performance of PersoNet. Finally, in Section VII, we conclude this paper and give some future research directions.

Manuscript received May 28, 2018; revised October 6, 2018 and December 29, 2018; accepted March 4, 2019. This work was supported by the National Natural Science Foundation of China under Grant 61872038. *(Huansheng Ning and Sahraoui Dhelim contributed equally to this work.) (Corresponding author: Huansheng Ning.)*

The authors are with the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: ninghuansheng@ustb.edu.cn).

Digital Object Identifier 10.1109/TCSS.2019.2903857

TABLE I

BIG-FIVE TRAITS AND ASSOCIATED CHARACTERS

Personality Trait	Related Characters
Openness to Experience	Artistic, Curious, Imaginative, Insightful, Original, Wide interests
Agreeableness	Trusting, Generous, Appreciative, Kind, Sympathetic, Forgiving
Conscientiousness	Efficient, Organized, Planful, Reliable, Responsible, Thorough
Extraversion	Energetic, Outgoing, Active, Assertive, Talkative
Neuroticism	Anxious, Unstable , Tense, Touchy, Worrying, Self-pitying

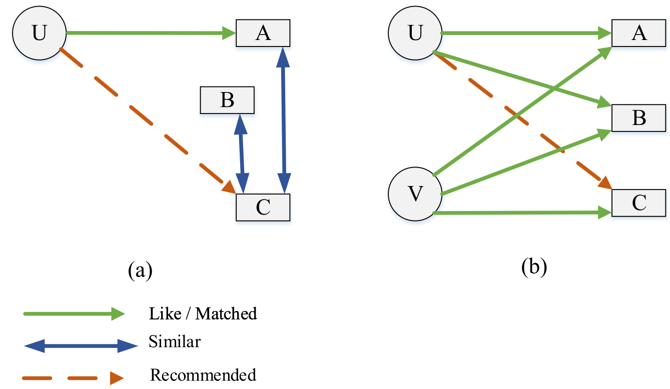

Fig. 1. (a) Content filtering. (b) CF.

II. BACKGROUND

In this section, we present some preliminaries about personality trait theory, as well as a recommendation system.

A. Human Personality

There is no general theory that defines the human personality. Nevertheless, many theories have elaborated the concept of human personality from different perspectives, including the cognitive perspective, biological perspective, learning perspective, humanistic perspective, psychodynamic perspective, and trait perspective [6].

Trait theory (also known as dispositional theory) is the most adapted personality theory. The trait theory suggests that human personality can be identified by the measurement of personality traits. Trait theorists define personality traits as habitual patterns of behaviors, thoughts, and emotions [7]. Personality traits are relatively stable over time, differ across individuals, relatively consistent over situations, and they influence human behaviors. There are two major methods used in trait theory to measure personality traits, Eysenck Personality Questionnaire (EPQ) also known as the three-factor model and big-five personality traits also known as the five-factor model (FFM).

The big-five traits are based on common language description of personality, which make trait theory an ideal model for computing technologies, such as natural language processing, machine learning, and semantic technologies. FFM is widely used for various purposes, such as job recruitment or mental disorders diagnosis. The model defines the five factors as openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism, often represented by the acronyms OCEAN or CANOE, in Table I, the five factors along some of the associated characters are presented.

There are many methods to measure one's personality; questionnaires where subjects answer with Likert scale questions about how they describe themselves are the most common personality measurement method. There are many well-known personality questionnaires with different lengths (items count). NEO Personality Inventory-Revised (NEO-PI-R, 240 items) [8] is one of the most adopted long personality questionnaires. For medium-size questionnaires, the NEO five-factor inventory (NEO-FFI, 60 items) [8], and the big-five inventory (BFI, 44 items) [9] is used frequently. Some other short questionnaires are much faster to fill (5–10 items), such as BFI-10 [10], short questionnaires retain only the most correlated items with each personality traits.

B. Recommendation Systems

A recommendation system is an information filtering system that is used to match a subject (e.g., user) with the best items (e.g., product and friend) that is suitable for its "needs" and/or "preferences." An FRS is a special case of recommendation system where the items are a set of users (potential friends). There are three main recommendation approaches as follows.

1) Content filtering [Fig. 1(a)] recommends items that are similar to those that a user (liked/bought/viewed) in the past (or is examining in the present). Particularly, various candidate items are compared with items previously (liked/bought/viewed) by the user and the best-matching items are recommended.
2) CF [11] [Fig. 1(b)] is based on the hypothesis that people who agreed in the past will agree in the future, and that they will (like/buy/view) be the same (or similar) items as they have (liked/bought/viewed) the same (or similar) items in the past.
3) Hybrid filtering is a combination of content filtering and CF.

III. RELATED WORKS

Many recent works have discussed the applications of human personality in computing systems. In this section, we review the related works about FRS in general and personality-based FRS specifically.

A. Friend Recommendation Systems

In the literature of social networks, many FRSs have been proposed, Wang *et al.* [12] proposed the Friendbook, an FRS that is based on semantic technologies, Friendbook recommends friends to users based on their lifestyles rather than social graphs. Friendbook identify users' lifestyles from user smartphone sensor data, after detecting their lifestyles, it recommends friends that have similar lifestyles. On the other hand, Yu *et al.* [13] presented GeoFriends, an FRS that recommends geographically related friends by social network structures analysis through combining GPS information. On the other hand, Silva *et al.* [14] developed an

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

algorithm that analyses the subgraph formed by a user and all the others connected users separately by three degrees of separation. Nevertheless, only users separated by two degrees of separation are candidates to be suggested as a friend. Hamid *et al.* [15] proposed a friend-recommendation system based on cohesion. They analyzed the cohesive subgroup on an augmented network formed by the physical connection network with the information of common interests and interactions. Bian and Holtzman [16] and Bian *et al.* [17] designed and implemented Matchmaker, a CF system that recommends friends to users on Facebook by matching and comparing user's online profile with the profiles of TV characters. For example, if Facebook user X is similar to TV character 1, and Facebook user Y is similar to TV character 2, and character 1 and character 2 are friends in the same TV show, then the Matchmaker system recommends user X to become friends with user Y.

However, none of the aforementioned works have incorporated personality traits in the friend recommendation process.

B. Personality and Social Networks

Many researchers have proved that the user's personality can be revealed by analyzing his social networks profile. Vinciarelli and Mohammadi [18] have surveyed the field of personality computing and its impact on the social computing system. While Kaushal and Patwardhan [19] surveyed the literature on automatic personality recognition from SNS data. Different works have used different methods for personality measurement using SNS data. The works in [20]–[22] have analyzed the language usage patterns and preferences to measure the user's personality, this was done by extracting linguistic features such as Linguistic Inquiry and Word Count, (Medical Research Council) Psycholinguistics Database, and parts-of-speech (POS). While the works in [23] and [24] discussed the relationship between social networks photographs and the user's personality. On the other hand, other works have found a positive relationship between the user's personality and other activities, such as online gaming [25]–[27] and online behaviors [28]. Ning *et al.* [29] discussed the relationship between excessive use of social networks and personality disorders.

C. Personality and Recommendation Systems

Personality traits have been used to enhance the recommendation systems in many domains, Tkalcic *et al.* [30] proposed a new approach for measuring the user similarity for CF recommender systems that are based on the big-five personality model in the context of product recommendation. Similarly, Hu and Pu [31] addressed the cold-start problem by incorporating human personality into the CF framework, they have tested the proposed system with movies and music public data sets. Ferwerda *et al.* [32], Ferwerda and Schedl [33], and Kleć [34] discussed the impact of personality traits on the accuracy of music recommendation systems. While Golbeck and Norris [35] have proved the positive correlation between personality and users' movie preferences. Using surveys and analysis of system data for 73 Netflix users, they have shown the correlations between personality and preferences for specific movie genres. Dhelim *et al.* [36] proposed a smart home architecture that incorporates the residents' personality traits into the smart home service recommendations.

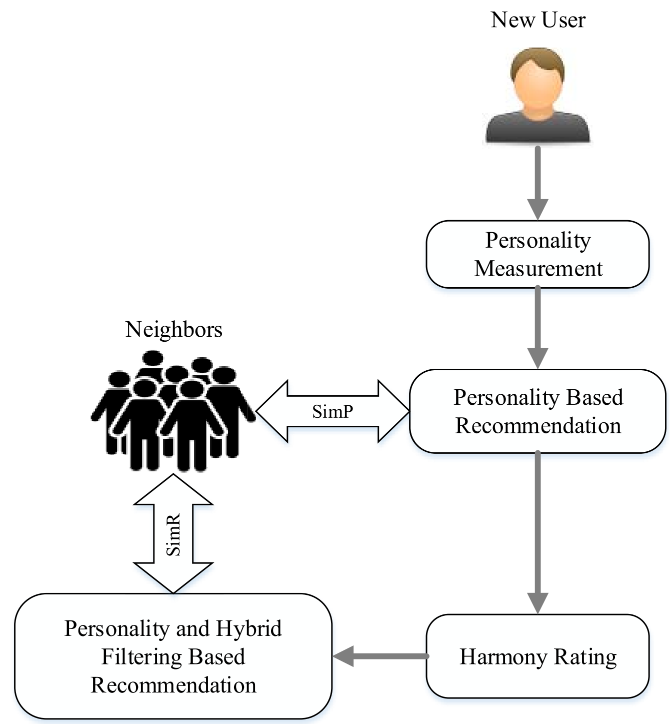

Fig. 2. PersoNet's system design.

However, in the context of FRS, as far as we know, we are the first that propose and implement a personality-based FRS.

IV. SYSTEM MODEL

The system design of the proposed system is presented in Fig. 2. After joining the network, the user must answer a personality measurement questionnaire. As the user has no preferences at this moment (cold start), to overcome this situation, the initial recommendation is based on personality similarity between the user and his neighbors (users with similar personality traits). In other words, the system recommends users that have identified as harmonic friends with neighbors of the new user. When the user passes the cold-start period, the recommendation will be gradually enhanced by incorporating the user's harmony rating preferences. As shown in Fig. 3, at the second stage, the recommendation is based on personality similarity and hybrid filtering approach (CF in terms of rating similarity with neighbors, and content filtering in terms of personality trait similarity between the previously rate friends and the potential friends).

A. Notations

For the sake of readability, the list of notations used in this paper is explained in Table II.

B. Similarity Measure

Similarity measure is the main component of any recommendation system and is used to measure the similarity between two entities (e.g., users and items) based on

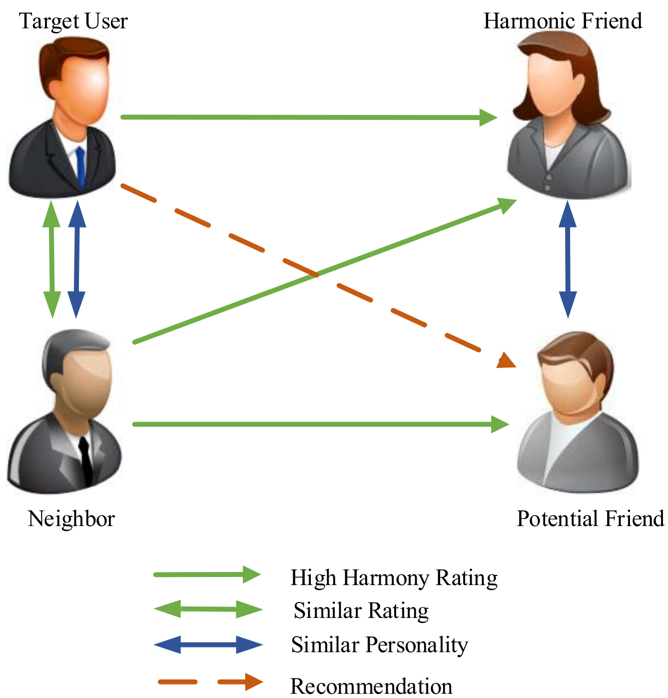

Fig. 3. Personality traits and hybrid filtering-based recommendations.

TABLE II
NOTATIONS AND SYMBOLS

Symbol	Meaning
$U = \{u_1, u_2, \ldots, u_n\}$	The set of all the users
$F_x = \{f_1, f_2, \ldots, f_i\}$	User u_x's friends
$\overrightarrow{P_x} = \left\{p_x^1, p_x^2, , \ldots, , p_x^j\right\}$	User u_x's personality traits vector
$SimP\,(u_x,\ u_y)$	Similarity between u_x and u_y based on their personality traits
$r_{x,y}$	Harmonic rating given by u_x to u_y
$R_x = \{r_{x,1}, r_{x,2}, \ldots, r_{x,n}\}$	The set of all harmonic rating given by u_x
$SimR\,(u_x,\ u_y)$	Similarity between u_x and u_y based on their given harmonic rating
$M_{x,i}$	The average of daily exchanged messages between u_x with u_i
$\overline{M_x}$	The mean of average of daily exchanged messages between u_x with his friends F_x
$SimM\,(u_x,\ u_y)$	Similarity between u_x and u_y based on their average daily exchanged messages
$N_{x,u}$	The set of users (neighbors) that their harmonic ratings are similar to u_x and have previously rated their harmony level with the potential friend u_u
$\widetilde{F_x}$	The set of recommended friends for target user u_x
α	The weight parameter that control the contribution of personality-based similarity in the overall similarity measure for neighbor formation procedure
β	The minimum rating similarity threshold
γ	The minimum personality similarity threshold
δ	The minimum neighbor formation threshold
ε	Neighbor formation threshold

a similarity factor (e.g., product ratings, browsing history, product category, and so on), in the context of recommendation systems, a precise similarity measurement enables the system to predict the future behaviors of the targeted entity based on the behaviors of its similar entities (neighbors). Many similarity measures have been proposed in the literature [37], [38]. In this paper, we have used the Pearson correlation coefficient, as it is one of the most commonly used similarity measures [39]. In this paper, we are interested in two kinds of similarity between the users as follows.

1) Personality Traits Similarity (SimP)*:* In this, we will measure the similarity between two users based on their personality trait (similarity factor). SimP is computed using the Pearson correlation coefficient as shown in (1)

$$\mathrm{SimP}(u_x,\ u_y) = \frac{\sum_i (p_x^i - \overline{p_x})(p_y^i - \overline{p_y})}{\sqrt{\sum_i (p_x^i - \overline{p_x})^2 (p_y^i - \overline{p_y})^2}} \tag{1}$$

where $\overline{p_x}$ and $\overline{p_y}$ are the average value of the personality traits vector for user u_x and u_y, respectively, and p_x^i is the ith trait in the personality traits vector.

2) Harmony Rating Similarity (SimR)*:* In this, we will measure the similarity between two users based on their harmony rating to other users. SimR is computed using Pearson correlation coefficient as shown in the following equation:

$$\mathrm{SimR}(u_x,\ u_y) = \frac{\sum_{i \in R_x \cap R_y} (r_{x,i} - \overline{r_x})(r_{y,i} - \overline{r_y})}{\sqrt{\sum_{i \in R_x \cap R_y} (r_{x,i} - \overline{r_x})^2 (r_{y,i} - \overline{r_y})^2}} \tag{2}$$

where $r_{x,i}$ and $r_{y,i}$ are the harmony rating given by u_x and u_y to u_i, respectively, and $\overline{r_x}$ and $\overline{r_y}$ are their mean harmony rating.

C. Recommendation System

There are two major steps in the recommendation procedure, the first step is neighborhood formation, in which the neighbors (most similar) of the targeted user are determined [see (3)], and the second step is friend prediction

$$N_{x,u} = \{u_y \in U\ \ : \mathrm{SimR}(u_x,\ u_y)\ >\ \beta \wedge u_u \in R_y\} \tag{3}$$

where u_x is the targeted user, u_u is the potential friend, and β is the minimum rating similarity threshold. $N_{x,u}$ is the list of users (neighbors) that are similar to u_x and have previously rated their harmony level with the potential friend, see Fig. 3.

To investigate the influence of personality traits on FRSs, we have proposed three recommendation systems with different recommendation algorithms as follows.

1) Birds-of-Feather: We have named this recommendation system based on the famous idiom "birds of a feather flock together." Birds-of-feather (BOF) is the simplest personality-based FRS, BOF will recommend people that have similar personality traits as friends, as shown in the following equation:

$$\widetilde{F_x} = \{u_u \in U\ \ :\ \ u_u \notin F_x \wedge \mathrm{SimP}(u_x,\ u_u) > \gamma\} \tag{4}$$

where u_u is the potential friend and γ is the minimum personality similarity threshold.

2) Collaborative Filtering: In this system, the recommendations are based only on the harmony rating that users provide after the data collection phase and do not consider the personality traits of both sides, as shown in the

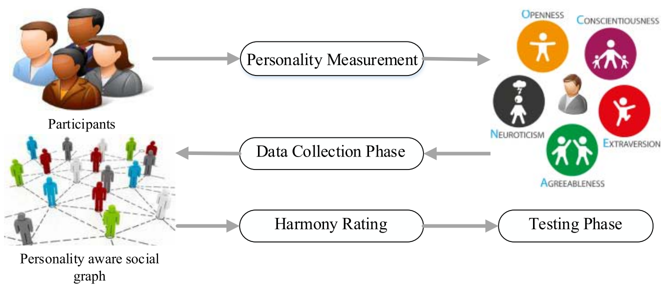

Fig. 4. Evaluation phases.

following equation:

$$\widetilde{F_x} = \{u_u \in U \ : \ u_u \notin F_x \wedge |N_{x,u}| > \delta\} \quad (5)$$

where $N_{x,u}$ is the set of targeted user's neighbors, u_u is the potential friend, and δ is the minimum neighbors threshold.

3) PersoNet: : In PersoNet, the recommendations are computed based on many factors.

1) The personality traits similarity between the user and his neighbors.
2) The personality traits similarity between the potential friend and the previously rated users (content filtering).
3) The rating similarity between the user and his neighbors (CF).

PersoNet's similarity measure is computed using the following equation:

$$\mathrm{Sim}(u_x,\ u_y) = \begin{cases} \mathrm{SimP}(u_x,\ u_y), & |F_x \cap F_y| < \varepsilon \\ \alpha \times \mathrm{SimP}(u_x,\ u_y) + (1-\alpha) \times \mathrm{SimR}(u_x, u_y), & |F_x \cap F_y| \geq \varepsilon \end{cases} \quad (6)$$

where α is the weight parameter that control the contribution of personality-based similarity in the overall similarity measure ($1 \geq \alpha \geq 0.5$). To solve the cold-start problem where the user has not giving its rating yet, the similarity will depend completely on personality similarity ($\alpha = 1$). Progressively, the more neighbors with common previous harmony rating, the less personality-based similarity contribute in the overall similarity. The group of recommended friends $\widetilde{F_x}$ for target user u_x is computed using the following equation:

$$\widetilde{F_x} = \{u_u \in U : u_u \notin F_x \wedge |N_{x,u}| > \delta \wedge \mathrm{Sim}P(\overline{p_{i \in F_x(r_{x,i} > \beta)}}, u_u) > \gamma\} \quad (7)$$

where $\overline{p_{i \in F_x(r_{x,i} > \beta)}}$ is the mean of personality traits vector of friends that their harmony rating given by the target user is greater than the minimum rating threshold β, and γ is the minimum personality similarity threshold.

V. Experiment Details

The main phases of the experiment are presented in Fig. 4.

A. Data

To validate the proposed system's accuracy, we have conducted an online experiment, in which we have created an SNS named PersoNet.[1] The participants have used the site for a period of 3 months. The experiment duration is divided into two phases. The first is the data collection phase, and it lasted for 2 months, in which each participant was befriended with a set of friends that have different personality traits. The recommendation at this stage was based only on neighbors' personality traits. At the end of this phase, participants were asked to rate the harmony level with their friends. The second phase is the testing phase, and it had lasted for 1 month, in which the participants were giving a chance to communicate with the recommended friends. To measure the recommendation system's accuracy, at the end of the experiment, given the list of friends that were recommended by the system, the participants were asked to rate the correctness of these recommendations.

[1] www.personet.online

TABLE III
Participants' Demographics

Parameter	Value
Number of volunteers	149
Number of active participants	123
Countries	12
Male	91 (61%)
Female	58 (39%)
Average age	23 years

B. Participants

To conduct the experiment, a total of 149 participants were solicited, most of them were undergraduate/graduate international students from different Chinese universities, and the other few were Chinese students. The benefit of using international students for this study is to conduct a holocultural study, as the participants were of different countries and even from different continents with different languages and cultural backgrounds. However, we have ensured that all participants can speak English fluently. In addition, the participants were of different ages and genders. Table III details the participants' demographics. The participants were volunteers and no compensation was made. From a total of 149 participants, only 123 active participants were considered in the later steps of the experiment, as the other 26 participants did not meet the minimum requirement of the experiment, because they did not use the system regularly, including such inactive users in neighbor formation process would negatively affect the accuracy of the recommendation.

C. Personality Measurement

After registering on the site, the users were asked to take (International Personality Item Pool Representation of the NEO PI-R)-60 personality questionnaire, and their big-five traits scores were recorded. The IPIP-NEO-60 is ideal for our situation. First, because it is open access, and second, because we did not want the volunteers to feel bored by answering a long questionnaire such as NEO-PI-R. Similarly, we have avoided using short questionnaires (five to ten questions) like BFI-10 [10]. Although they are easier to fill, their measurements are usually inaccurate, simply because the taker may underestimate the answer of one of the ten questions.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

D. Data Collection Phase

In the data collection phase, during 2 months, the participants were befriended with 30 people with different dominant traits, 6 people from each trait. The participants were encouraged to chat with their default friends as much as possible; the chat was done via PersoNet's integrated messaging system. To ensure that the participants' harmony rating is based on personality traits and to reduce the influence of homophily, the participants were asked not to reveal their real identities or any other information that would influence their rating later on, such as location, religion, age, sex, and political views. However, the participants were strongly recommended to discuss their views about other topics.

E. Harmony Rating

After 2 months of the data collection phase, the participants were asked to rate the harmony level of their friends. Each participant was asked, based on his knowledge of a given friend, whether their personality types are matched and his desire to see people like him in his recommendations. The harmony rating was scaled as Likert scale [40]. These ratings were used to determine the user's neighbors.

F. Friend Recommendations

After all participants finished the harmony rating of their friends, all participants' usernames were changed, and each participant is friended with 30 friends, the top five most recommended users by BOF, CF, and PersoNet (a friend might be recommended by more than one system), in addition to these recommendations, each participant is also friended with the top five least recommended users.

G. Testing Phase

The second phase is the testing phase, and it lasts for 1 month, in which the participants were giving the chance to communicate with the recommended friends. To measure the robustness of the three recommendation systems, at the end of the experiment, the participants were asked to rate the correctness of the recommendations. The participants did not know any details about the used recommendation systems. Based on the recommendation correctness ratings that were given by the participants, PersoNet, BOF, and CF systems were evaluated.

VI. Performance Evaluation

In this section, we present the evaluation that we conducted to validate the proposed system.

A. Implementation

To validate the proposed system's accuracy, a personality-based SNS that uses the proposed FRS named PersoNet was implemented, in which the online experiment was conducted to study users' satisfaction about the site's recommendation system. PersoNet was implemented using PHP, and the database management using MySQL, and the front-end interface using the bootstrap framework, see Fig. 5.

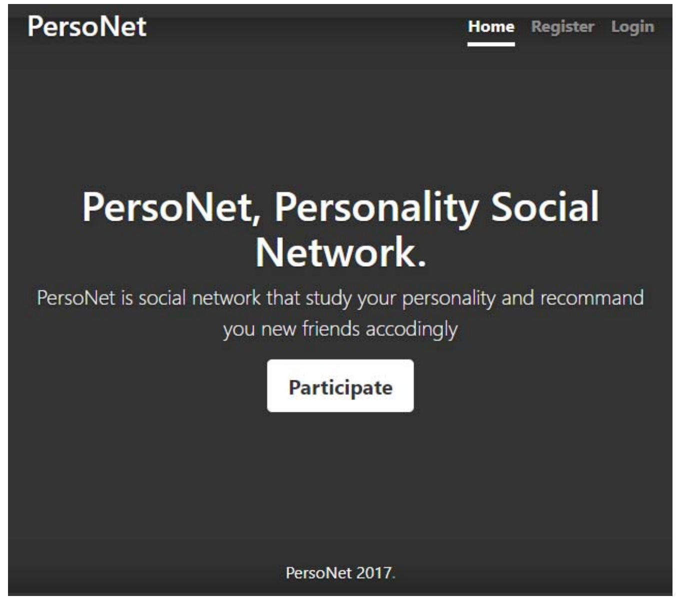

Fig. 5. PersoNet social network.

B. Evaluation Metrics

As shown in [41], recommendation systems are evaluated based on their ability to identify the relevant items for a given user from all available items. Four groups of decisions are yielded from the confusion matrix that lists the correct and incorrect recommendations: 1) true positives (TPs): the recommended friends that were rated as successful recommendations; 2) true negatives (TNs): the least recommended friends that were rated as unsuccessful recommendations; 3) false positives (FPs): the most recommended friends that were rated as unsuccessful recommendations; and 4) false negatives (FNs): the least recommended friends that were rated as successful recommendations. We have evaluated the three systems based on the following metrics.

1) Precision (P): It is the fraction of confirmed recommendations among the total recommended users by the recommendation system and is computed using the following equation:

$$P = \frac{TP}{TP + FP}. \tag{8}$$

2) Recall (R): It is the fraction of confirmed recommendations over the total confirmed recommendations by all systems and is computed using the following equation:

$$R = \frac{TP}{TP + FN}. \tag{9}$$

3) F-Measure: A combination of precision and recall in a single numerical value, it is also known as F-score and is computed using the following equation:

$$F = \frac{2\ P\ R}{P + R}. \tag{10}$$

4) Result Discussion: The mean values of precision, recall, and F-measure of the three systems are presented in Fig. 6. As we can see, BOF has the worst performance in terms of precision (0.55) and recall (0.58), as it considers only personality trait similarity measurement and ignores the user's

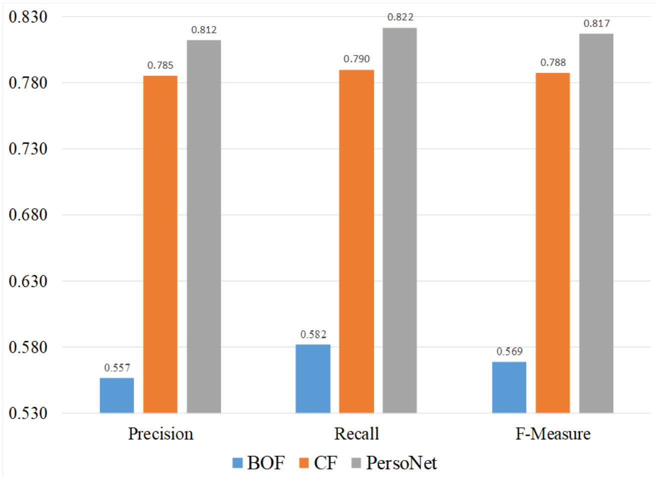

Fig. 6. Systems evaluation with rating-based similarity.

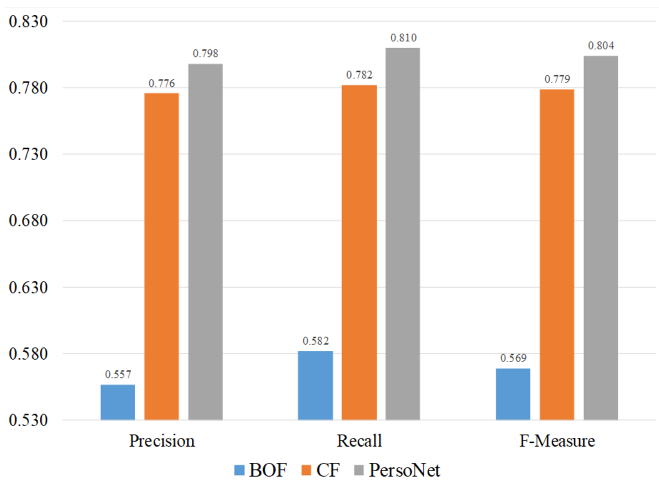

Fig. 7. Systems evaluation with message-exchange-based similarity.

previous harmony ratings, while CF scores are much better than BOF with precision (0.78) and recall (0.79). However, PersoNet has the highest precision and recall values among the three systems, with precision (0.81) and recall (0.82), that is, because it incorporates personality traits in similarity measurement without neglecting the user's preferences.

While harmony rating can be a strong indicator of how the friends are matched. However, to further validate the proposed system, we have analyzed the text exchanges between friends regarding their personalities. In this regard, we have computed the neighbor set based on the exchanged messaged between the participants rather than the harmony rating that was given by the users; the text exchange similarity is computing using the following equation:

$$\mathrm{SimM}(u_x, u_y) = \frac{\sum_{i \in F_x \cap F_y} (M_{x,i} - \overline{M_x})(M_{y,i} - \overline{M_y})}{\sqrt{\sum_{i \in F_x \cap F_y} (M_{x,i} - \overline{M_x})^2 (M_{y,i} - \overline{M_y})^2}} \quad (11)$$

where $M_{x,i}$ and $M_{y,i}$ are the average of daily exchanged messages between u_i with u_x and u_y, respectively. $\overline{M_x}$ and $\overline{M_y}$ are the mean of average of daily exchanged messages between u_x and u_y with their friends F_x and F_y, respectively.

The precision, recall, and F-measure of the three systems when similarity computed based on exchanged messages are presented in Fig. 7. As we can observe, BOFs precision (0.557) and recall (0.582) are the same as harmony rating-based similarity, that is, because BOF depends only on personality similarity between the potential friend and previous friends (content filtering). Although CF performs better than BOF in terms of precision (0.776) and recall (0.782), however, PersoNet still has the upper hand with precision (0.798) and recall (0.810).

VII. Conclusion

In this paper, a novel FRS based on the big-five personality traits model and hybrid filtering was presented and evaluated, in which the friend recommended process is based on personality traits and users' harmony rating. To validate the proposed system's accuracy, a personality-based social network site that uses the proposed system named PersoNet was implemented. The experimental results have proved that PersoNet performs better than the legacy CF-based system in terms of precision and recall. However, many aspects that could improve the effectiveness of Personet have not been discussed in this paper, such as follows.

1) In this paper, the subjects' personality traits measurement was done through questionnaires. However, PersoNet could be further improved by implementing automatic personality recognition scheme, which measures the user's personality traits based on its posted content, without the need for personality test.
2) The effectiveness of PersoNet was evaluated based on the recommendations accuracy that was validated by the users' rating. Extending the experiment by comparing PersoNet's recommendations accuracy to other schemes, such as graph-based and semantic-based recommendations, is our future direction.
3) The proposed recommendation system is based on the big-five personality traits model. Extending the model to incorporate other personality traits models such as Myers–Briggs type indicator is one of the future works.
4) The population of the experiment is relatively small ($n = 126$). Conducting the experiment on a large size population ($n > 1000$) from all ages is a future direction.

Acknowledgment

The authors would like to thank all the volunteers who have donated their time for PersoNet's experiment.

References

[1] S. KEMP. (2018). *Number of Social Media Users Passes 3 Billion With no Signs of slowing*. https://thenextweb.com/contributors/2017/08/07/number-social-media-user%s-passes-3-billion-no-signs-slowing/
[2] H. Zhao, H. Zhou, C. Yuan, Y. Huang, and J. Chen, "Social discovery: Exploring the correlation among three-dimensional social relationships," *IEEE Trans. Computat. Social Syst.*, vol. 2, no. 3, pp. 77–87, Sep. 2015.
[3] T. Altmann, S. Sierau, and M. Roth, "I guess you're just not my type," *J. Individual Differences*, vol. 34, no. 2, pp. 105–117, 2013. doi: 10.1027/1614-0001/a000105.

[4] J. Li and M. Chignell, "Birds of a feather: How personality influences blog writing and reading," *Int. J. Hum.-Comput. Stud.*, vol. 68, no. 9, pp. 589–602, 2010.

[5] W. Youyou, D. Stillwell, H. A. Schwartz, and M. Kosinski, "Birds of a feather do flock together: Behavior-based personality-assessment method reveals personality similarity among couples and friends," *Psychol. Sci.*, vol. 28, no. 3, pp. 276–284, 2017.

[6] P. J. Corr and G. Matthews, *The Cambridge Handbook of Personality Psychology*. New York, NY, USA: Cambridge Univ. Press, 2009.

[7] W. Fleeson and E. Jayawickreme, "Whole trait theory," *J. Res. Pers.*, vol. 56, pp. 82–92, Jun. 2015.

[8] P. T. Costa and R. R. McCrae, *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Lutz, FL, USA: Psychological Assessment Resources, 1992.

[9] O. P. John, E. M. Donahue, and R. L. Kentle, "The big five inventory–versions 4a and 54," Univ. California, Inst. Personality Social Res., Berkeley, CA, USA, 1991.

[10] B. Rammstedt and O. P. John, "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German," *J. Res. Pers.*, vol. 41, no. 1, pp. 203–212, 2007.

[11] X. Yang *et al.*, "Collaborative filtering-based recommendation of online social voting," *IEEE Trans. Computat. Social Syst.*, vol. 4, no. 1, pp. 1–13, Mar. 2017.

[12] Z. Wang, J. Liao, Q. Cao, H. Qi, and Z. Wang, "Friendbook: A semantic-based friend recommendation system for social networks," *IEEE Trans. Mobile Comput.*, vol. 14, no. 3, pp. 538–551, Mar. 2015.

[13] X. Yu, A. Pan, L.-A. Tang, Z. Li, and J. Han, "Geo-friends recommendation in GPS-based cyber-physical social network," in *Proc. Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Jul. 2011, pp. 361–368.

[14] N. B. Silva, I.-R. Tsang, G. D. C. Cavalcanti, and I.-J. Tsang, "A graph-based friend recommendation system using genetic algorithm," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jul. 2010, pp. 1–7.

[15] M. N. Hamid, M. A. Naser, M. K. Hasan, and H. Mahmud, "A cohesion-based friend-recommendation system," *Social Netw. Anal. Mining*, vol. 4, no. 1, p. 176, 2014.

[16] L. Bian and H. Holtzman, "Online friend recommendation through personality matching and collaborative filtering," in *Proc. UBICOMM*, 2011, pp. 230–235.

[17] L. Bian, H. Holtzman, T. Huynh, and M.-J. Montpetit, "MatchMaker: A friend recommendation system through TV character matching," in *Proc. IEEE Consum. Commun. Netw. Conf. (CCNC)*, Jan. 2012, pp. 714–718.

[18] A. Vinciarelli and G. Mohammadi, "A survey of personality computing," *IEEE Trans. Affective Comput.*, vol. 5, no. 3, pp. 273–291, Jul./Aug. 2014.

[19] V. Kaushal and M. Patwardhan, "Emerging trends in personality identification using online social networks—A literature survey," *ACM Trans. Knowl. Discovery Data*, vol. 12, no. 2, p. 15, 2018.

[20] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *J. Artif. Intell. Res.*, vol. 30, pp. 457–500, Nov. 2007.

[21] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft, "Our Twitter profiles, our selves: Predicting personality with twitter," in *Proc. IEEE 3rd Int. Conf. Privacy, Secur., Risk Trust (PASSAT)*, Oct. 2011, pp. 180–185.

[22] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from Twitter," in *Proc. IEEE 3rd Int. Conf. Privacy, Secur., Risk Trust (PASSAT), IEEE 3rd Int. Conf. Social Comput. (SocialCom)*, Oct. 2011, pp. 149–156.

[23] Y.-C. J. Wu, W.-H. Chang, and C.-H. Yuan, "Do Facebook profile pictures reflect user's personality?" *Comput. Hum. Behav.*, vol. 51, pp. 880–889, Oct. 2015.

[24] A. Eftekhar, C. Fullwood, and N. Morris, "Capturing personality from Facebook photos and photo-related activities: How much exposure do you need?" *Comput. Hum. Behav.*, vol. 37, pp. 162–170, Aug. 2014.

[25] Z. Halim, M. Atif, A. Rashid, and C. A. Edwin, "Profiling players using real-world datasets: Clustering the data and correlating the results with the big-five personality traits," *IEEE Trans. Affective Comput.*, to be published. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8036211. doi: 10.1109/TAFFC.2017.2751602.

[26] N. C. Worth and A. S. Book, "Dimensions of video game behavior and their relationships with personality," *Comput. Hum. Behav.*, vol. 50, pp. 132–140, Sep. 2015.

[27] A. Bean and G. Groth-Marnat, "Video gamers and personality: A five-factor model to understand game playing style.," *Psychol. Popular Media Culture*, vol. 5, no. 1, pp. 27–38, 2016.

[28] S. Bai, T. Zhu, and L. Cheng. (2012). "Big-five personality prediction based on user behaviors at social network sites." [Online]. Available: https://arxiv.org/abs/1204.4809

[29] H. Ning, S. Dhelim, M. A. Bouras, A. Khelloufi, and A. Ullah, "Cyber-syndrome and its formation, classification, recovery and prevention," *IEEE Access*, vol. 6, pp. 35501–35511, 2018.

[30] M. Tkalcic, M. Kunaver, J. Tasic, and A. Košir, "Personality based user similarity measure for a collaborative recommender system," in *Proc. 5th Workshop Emotion Hum.-Comput. Interact.-Real World Challenges*, 2009, pp. 30–37.

[31] R. Hu and P. Pu, "Enhancing collaborative filtering systems with personality information," in *Proc. 5th ACM Conf. Recommender Syst.*, 2011, pp. 197–204.

[32] B. Ferwerda, M. Tkalcic, and M. Schedl, "Personality traits and music genre preferences: How music taste varies over age groups," in *Proc. 1st Workshop Temporal Reasoning Recommender Syst. (RecTemp) 11th ACM Conf. Recommender Syst.*, Como, Italy, Aug. 2017, pp. 16–20.

[33] B. Ferwerda and M. Schedl, "Personality-based user modeling for music recommender systems," in *Proc. Joint Eur. Conf. Mach. Learn. Knowl. Discovery Databases*. Riva del Garda, Italy: Springer, 2016, pp. 254–257.

[34] M. Kleć, "The influence of listener personality on music choices," *Comput. Sci.*, vol. 18, no. 2, pp. 163–178, 2017.

[35] J. Golbeck and E. Norris, "Personality, movie preferences, and recommendations," in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining (ASONAM)*, Aug. 2013, pp. 1414–1415.

[36] S. Dhelim, H. Ning, M. A. Bouras, and J. Ma, "Cyber-enabled human-centric smart home architecture," in *Proc. IEEE SmartWorld, Ubiquitous Intell. Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Oct. 2018, pp. 1880–1886.

[37] M. Nasim, R. Charbey, C. Prieur, and U. Brandes, "Investigating link inference in partially observable networks: Friendship ties and interaction," *IEEE Trans. Computat. Social Syst.*, vol. 3, no. 3, pp. 113–119, Sep. 2016.

[38] F. O. Isinkaye, Y. O. Folajimi, and B. A. Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egyptian Inform. J.*, vol. 16, no. 3, pp. 261–273, 2015.

[39] H. Liu, Z. Hu, A. Mian, H. Tian, and X. Zhu, "A new user similarity model to improve the accuracy of collaborative filtering," *Knowl.-Based Syst.*, vol. 56, pp. 156–166, Jan. 2014.

[40] A. Joshi, S. Kale, S. Chandel, and D. K. Pal, "Likert scale: Explored and explained," *Brit. J. Appl. Sci. Technol.*, vol. 7, no. 4, p. 396, 2015.

[41] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Trans. Inf. Syst.*, vol. 22, no. 1, pp. 5–53, 2004.

Huansheng Ning (SM'13) received the B.S. degree from Anhui University, Anhui, China, in 1996, and the Ph.D. degree from Beihang University, Beijing, China, in 2001.

He is currently a Professor and the Vice Dean of the School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing. His current research interests include Internet of Things and general cyberspace. He has authored or co-authored more than 100 journal/conference papers and authored five books.

Dr. Ning serves as the Steering Committee Member for the IEEE INTERNET OF THINGS JOURNAL since 2016. He is the Founder and the Chair of the Cyberspace and Cybermatics International Science and Technology Cooperation Base. He has presided many research projects, including the Natural Science Foundation of China and the National High Technology Research and Development Program of China (863 Project). He serves as an Associate Editor for the IEEE SYSTEMS JOURNAL since 2013 and the IEEE INTERNET OF THINGS JOURNAL from 2014 to 2018.

Sahraoui Dhelim received the B.S. degree in computer science from the University of Djelfa, Djelfa, Algeria, in 2012, and the master's degree in networking and distributed systems from the University of Laghouat, Laghouat, Algeria, in 2014. He is currently pursuing the Ph.D. degree with the University of Science and Technology Beijing, Beijing, China.

His current research interests include social computing, personality computing, user modeling, recommendation systems, and intelligent transportation systems.

Nyothiri Aung received the master's degree in information technology from Mandalay Technological University, Mandalay, Myanmar, in 2012. She is currently pursuing the Ph.D. degree with the University of Science and Technology Beijing, Beijing, China.

From 2008 to 2010, she was a Tutor with the Department of Information Technology, Technological University of Meiktila, Meiktila, Myanmar. From 2012 to 2015, she was with System Analyst of ACE Data System, Yangon, Myanmar. Her current research interests include intelligent transportation system and personality computing.