OPEN ACCESS
Indo-JC Indonesia Journal of Computing
ISSN 2460-9056
socj.telkomuniversity.ac.id/indojc

Ind. Journal on Computing
Vol. 4, Issue. 2, Sept 2019. pp. 21-30
doi:10.21108/indojc.2019.4.2.258

# Group Recommenders System Using Hybrid Filtering for Tourism Domain

Lutfi Ambarwati #1, Z.K.A. Baizal #2

*# Informatics Faculty, Telkom University*
*Bandung, Indonesia*

[1] lutfiambar@students.telkomuniversity.ac.id
[2] baizal@telkomuniversity.ac.id

**Abstract**

**Group recommenders system is built to overcome the different needs that exist in each of the tour group members in determining tourist destinations. This is because everyone has different desires in determining the constraints of tourist destinations, such as costs and tourist categories. This recommendation process is based on a combination of profiles of each member to produce a tourist destination that can fulfill their needs. This paper uses hybrid filtering which is a combination of collaborative filtering and knowledge-based filtering. Because the two methods have their own shortcomings. The collaborative method has a deficiency of tourist items that have never been rated by a user before, so the item cannot be recommended for active users. While in knowledge-based, tourist items must be completely described in detail so that they can be recommended to users. Domain used in this research is Bandung Raya area. The results of the evaluation in this study, users tend to feel very satisfied for the results of group travel recommendations generated the application. As for the evaluation of the recommendation algorithm, users prefer the results of the recommendations produced by collaborative and hybrid methods.**

**Keywords: Bandung Raya, collaborative filtering, group recommenders system, knowledge-based filtering, recommenders system.**

**Abstrak**

***Group recommender system* dibangun untuk mengatasi perbedaan yang ada pada masing-masing anggota kelompok dalam menentukan tujuan wisata. Hal ini disebabkan, karena setiap orang memiliki keinginan masing-masing dalam menentukan batasan tujuan wisata, seperti biaya dan kategori wisata. Hal ini disebabkan, karena setiap orang memiliki keinginan masing-masing dalam menentukan batasan tujuan wisata, seperti biaya dan kategori wisata. Makalah ini menggunakan *hybrid filtering* yang merupakan gabungan dari metode *collaborative filtering* dan *knowledge-based filtering*. Karena kedua metode tersebut memiliki kekurangan masing-masing. Metode *collaborative filtering* memiliki kekurangan dimana suatu *item* wisata yang belum pernah di *rating* pengguna sebelumnya, maka *item* tersebut tidak dapat direkomendasikan untuk pengguna aktif. Sedangkan pada metode *knowledge-based filtering, item* wisata harus benar-benar di deskripsikan secara detail supaya dapat direkomendasikan kepada pengguna. Domain yang digunakan dalam penelitian ini adalah wilayah Bandung Raya. Hasil dari evaluasi dalam penelitian ini, pengguna cenderung merasa sangat puas dengan hasil rekomendasi perjalanan kelompok yang dihasilkan aplikasi. Sedangkan untuk evaluasi algoritma rekomendasi, pengguna lebih suka hasil rekomendasi yang dihasilkan oleh metode kolaborasi dan hybrid.**

**Kata kunci: Bandung Raya, group recommender system, collaborative filtering, knowledge-based filtering, rekomendasi sistem.**

Received on August 2019. Accepted on Sept 2019

## I. INTRODUCTION

Decision making in the system that makes learning groups from each member, such as tourism categories, tourist locations, tourism facilities, and tourism costs. So that group members are satisfied with the recommendations. The recommendation system is widely used in various business areas to make decisions or recommend choices from available information [1] [2] [3]. The recommendation system in the tourism area is one example of the application of decision making to help tourists in determining their tourism destinations. The internet and the world wide web provide a wealth of information about tourism in the world. But to find this information is very difficult. Because there is no information available on the internet according to the latest conditions in tourist attractions. So that a system is needed to resolve the problem. The system is known as the Recommender System which is used for e-commerce website applications to offer and add some travel tips to users. The recommendation system can also be in the form of information and provide facilities needed in the decision making process [4] [5] [6].

Therefore, the corresponding recommendation model for the recommendation system is needed so that it can recommend tourist destinations and make it easier for users to determine tours according to the user's needs [7]. Travel is a group activity, both with family, friends, and school friends. Group recommendation system that aims to design the goals of groups with group members in according to group needs. In this research system uses hybrid filtering method which is a combination of collaborative filtering methods and knowledge-based filtering methods [8]. Because both methods have their own shortcomings, a hybrid filtering method is needed to solve the problem.

The collaborative filtering method has a lack of rating on items that have never been rated and items that have never been included before. This item cannot be given to the user. While in knowledge-based filtering systems knowledge must really know the characteristics and domain of each item for the process of finding tourist destinations to users. The system must be able to determine information that suits the needs of users with available items [9] [10].

In this paper, the author builds a system of recommender group system, where user ratings, tourist categories, and user costs are used as input in the application. Then it is processed by algorithm calculation from collaborative filtering method and knowledge-based filtering method. After that the results of the two methods are combined into the hybrid filtering calculation. The constraints in this study are as follows.

- The tourism recommendation system of this group only calculates the explicit rating of tourist items from users already in the system.
- Users who have never rated a tourist item in the system, default to 0.
- Do not handle cold-start problems on tourist items.
- Users have registered and have rated items before.

The purpose of building a group tourism recommendation system is to overcome differences between group members who want to travel together. So that produced recommendations for tourist destinations that can fulfill their wishes.

## II. RELATED WORK

### A. *Collaborative Filtering*

Collaborative filtering is a method used to predict the usefulness of an item based on the previous user's rating on the item [11]. This method recommends items that have been selected by previous users with similar model items from active users [12].

The collaborative filtering method in this study uses an item-based collaborative filtering. Item-based Collaborative Filtering is a recommendation method based on the similarity between a previous item and the selected item. From the level of similarity of items, divided by parameters the user needs to obtain the value of the use of an item. Items that have the highest usability value are recommended [13]. This approach has two stages such as in the first stage of collecting the most similar tourism items. This stage uses Pearson Correlation-based Similarity calculations to get the closeness between the two most similar tourist items. Furthermore, the second stage calculates item rating predictions that have not been rated by active users [8]. The results of rating prediction calculations are used for hybrid calculations.

*1) Pearson Correlation-based Similarity:* is used to measure how much a linear relationship between two variables [14]. Similarity values range from -1 to +1. The value -1 has a negative linear correlation, the value 0 has no correlation, and the +1 value has a positive linear correlation. $\mathrm{sim}(i,j)$ is the closeness between item i and item j. m is the number of users who give a rating on item i and item j. $Ru, i$ is the rating given by the user for the item i. $\overline{R\imath}$ is the average of users who give a rating on item i. The equation for calculating the similarity between items i and item j is as follows [15].

$$sim(i,j) = \frac{\sum_{u=1}^{m}(Ru,i-\overline{R\imath})(Ru,j-\overline{RJ})}{\sqrt{\sum_{u=1}^{m}(Ru,i-\overline{R\imath})^2}\sqrt{\sum_{u=1}^{m}(Ru,j-\overline{RJ})^2}} \quad (1)$$

*2) Rating Prediction:* After obtaining similarity values from Pearson Correlation-based Similarity, an item rating prediction has not been rated by the user. This rating prediction is used as a candidate to recommend items for users. Value on prediction rating between 1 to 5. n is a neighbor of items for users u. This neighbor consists of item j which is given an explicit rating by the user u and is similar to item i. $\hat{r}u, i$ is a rating prediction for u users on item i. This neighbor is also different for each user.

$$\hat{r}u, i = \overline{\mathrm{r}\imath} + \frac{\sum_{u=1}^{n} sim(i,j)(Ru,j-\overline{RJ})}{\sum_{u=1}^{n}|sim(i,j)|} \quad (2)$$

*B. Knowledge-based Filtering*

The system must understand more information about tourist destinations with restrictions raised by users. Knowledge-based filters out tourist destinations that are similar to what users want. The knowledge-based stage collects information only for tourist destination domains. The following sources of information are integrated into knowledge-based filtering [5] [8].

- Information on tourist destinations, such as: description of tourist attractions, address, telephone, coordinates, and facilities.
- Tourism costs, such as: entrance fee for tourist sites.
- Tourist categories, such as: nature and outbound tours, educational tours, park and playground tours, swimming pools and waterfalls, religious and cultural tours, and shopping tours.
- Tourist areas, such as: West Bandung, Bandung City, Kab. Bandung, and Cimahi.

$$\hat{r}u, i = \frac{\sum_{k\epsilon D} wk * sc(i,k)}{\sum_{k\epsilon D} wk} \quad (3)$$

$$sc(i,k) = 1 - \frac{(expected_{cost} - \max\_budget, 0)}{\max\_budget} \quad (4)$$

$$wk = 1 \quad (5)$$

wk is the weight of the knowledge-based recommender method. $\mathrm{sc}(i,k)$ is a scoring for items i.

### C. Hybrid Filtering

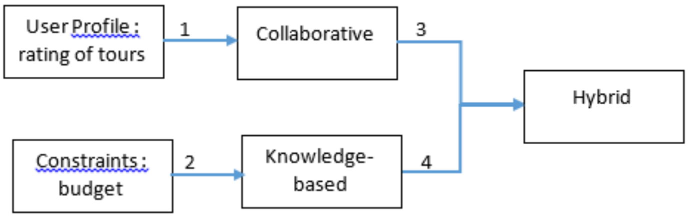



Fig. 1. Block diagram of Hybrid Recommendations [8]

Hybrid filtering is a recommendation system method for combining several recommendation methods to produce tourist destination output [12]. Some recommendation systems use hybrid methods to combine collaborative and knowledge-based methods to help the constraints found in both methods [1]. From the results of the two rating prediction calculations combined into hybrid predictions by adding multiplication results from the two rating predictions with weight on the hybrid. The weight of the hybrid for both methods is 1/2. Because the weight on the hybrid if added should be 1 [8]. wcf is a weight for collaborative filtering. r̂cf is a rating prediction for collaborative filtering.

$$\hat{r}hybrid = wcf * \hat{r}cf + wkb * \hat{r}kb \tag{6}$$

$$wcf + wkb = 1 \tag{7}$$

### D. Anova Test

Analysis of variance (Anova) test is a statistical procedure used to test a hypothesis by comparing the average number of samples of population variance. There are two hypotheses used in Anova test, such as:

- H0 (Null Hypothesis), the initial hypothesis that the average of two groups of variance is not different. H0 has a significance level of $(\alpha) \leq 0.05$
- H1 (Alternative Hypothesis), namely the alternative hypothesis that the average of the two groups of variance is different. H1 has a significance level of $(\alpha) > 0.05$

## III. RESEARCH METHOD

### A. Dataset

The tourism dataset consists of details of tourist attractions and rating of tourist attractions resulting from the questionnaire. Below is a table I of tourist attractions used in this study.

TABLE I
DATASET OF TOURISM

| No. | Category of Tourism | Country of Tourism (Total) | | | |
|---|---|---|---|---|---|
| | | Bandung Barat | Kab. Bandung | Kota Bandung | Cimahi |
| 1. | natural and outbound tourism | 18 | 15 | 1 | 3 |
| 2. | swimming pool and waterboom | 1 | 8 | 2 | 2 |
| 3. | tourist parks and playgrounds | 3 | 1 | 12 | 2 |
| 4. | educational tours | 8 | 1 | 11 | 0 |
| 5. | religious and cultural tourism | 1 | 0 | 3 | 1 |
| 6. | shopping tour | 0 | 0 | 4 | 0 |

*B. Group Recommendation*

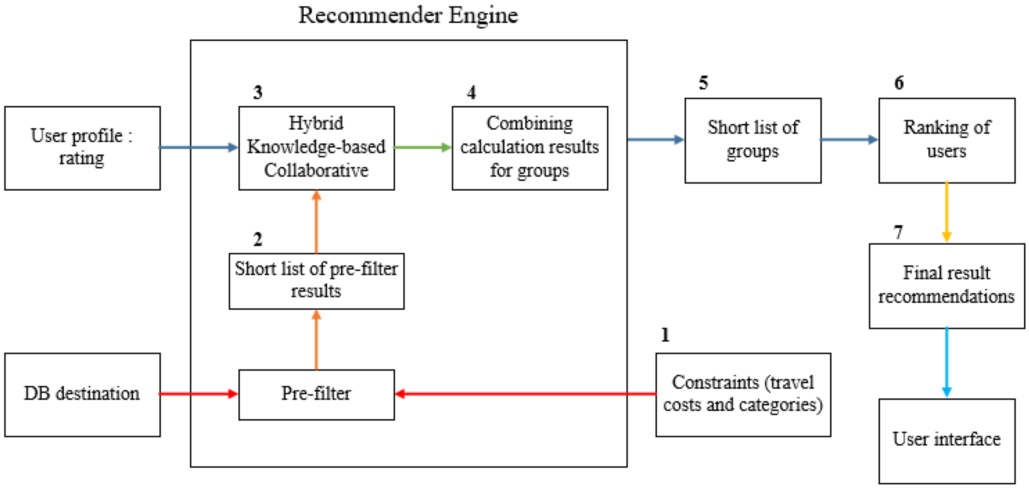



Fig. 2. Block diagram of Group Recommendations [8]

Fig. 2. shows that group recommendations require input consisting of user profiles in the form of rating of tourist attractions, tourist categories, and costs [16]. The following stages can distinguish information flows according to fig. 2.

- Users enter tourism categories and costs on the application.
- Then the system performs a pre-filter process from the database of tourist attractions to user input in the form of tourist categories. This process produces a short list of tourist attractions to be used in the method calculation process.
- Short list of pre-filter results is calculated using collaborative methods (for rating), knowledge methods (for costs), and hybrid methods (combining collaborative and knowledge).
- Results from the calculation of the hybrid method of each group member are combined using the recommendation aggregation technique [17]. The recommendation aggregation technique uses an "average" strategy [18] which is calculating the average hybrid rating prediction of each group member who has the same tourist destination. But if tourism destinations are different, hybrid rating predictions will immediately become a short list of groups [8].
- Short list of groups are a list of initial tourist destinations for groups.
- After that, the results short list of groups are carried out a ranking process for each group member. This process uses the Borda calculation method [18] to determine the winner of a tourist destination by giving a number of weights to each candidate for a tourist destination. The highest weight is given to the first rank, while the lowest weight is given to the last rank [19]. Ranking is given between 1 to 10 [8], where the weight for the first rank is 10 and the rating of 10 is 1 [19].
- The results of the Borda calculation process become a list of final recommendations.

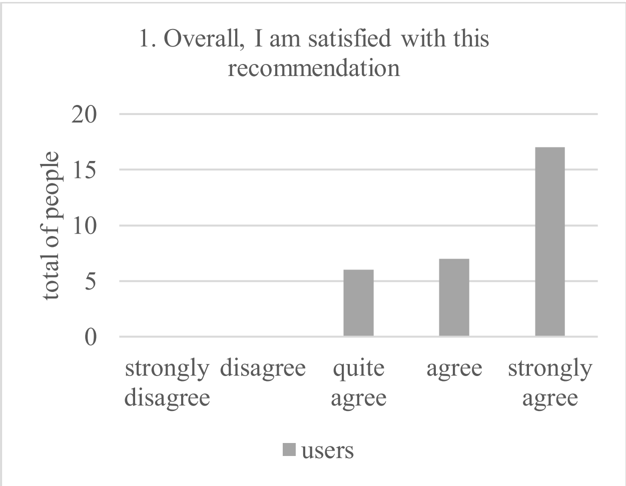
## IV. RESULTS AND DISCUSSION

*A. Results of Testing User Satisfaction Levels on Group Recommendations*

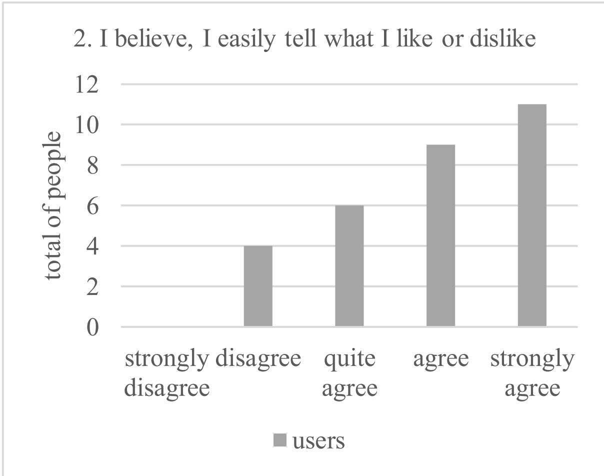
In the results of testing the level of user satisfaction on the group recommendations, testing was conducted in the form of a questionnaire to determine the level of user satisfaction in using this application. Data collected from 30 users through questionnaires such as fig. 3. Reference questions and users from previous paper references.

TABLE II
USER DISTRIBUTION IN THE QUESTIONNAIRE

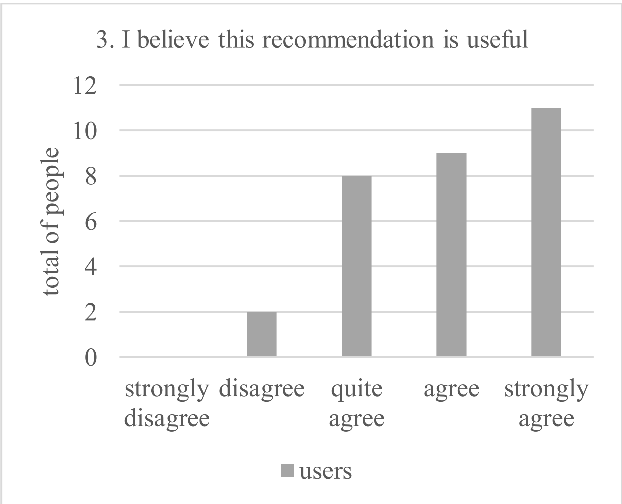
| No. | Gender | Age | Job | Total |
|---|---|---|---|---|
| 1. | Male | 22 years and above | college student | 14 |
| 2. | Female | 22 years and above | college student | 16 |



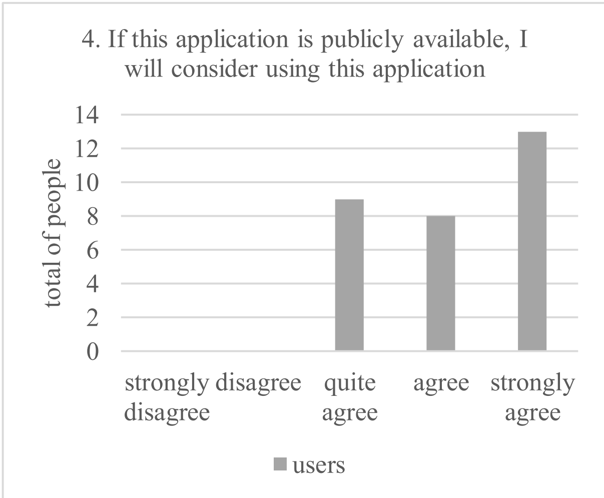

(a)




(b)




(c)




(d)

Fig. 3. User Satisfaction on Group Recommendations (a), preference for recommendations (b), usefulness of recommendations (c), application usage in the future (d)

Based on fig. 3. the level of user satisfaction on group recommendations produced by the application for the strongly agree category has the highest percentage value of 56.667% and the category agrees with the highest percentage of 30% from of 30 users in table II.

*B. Results of Evaluation of Algorithm Recommendations*

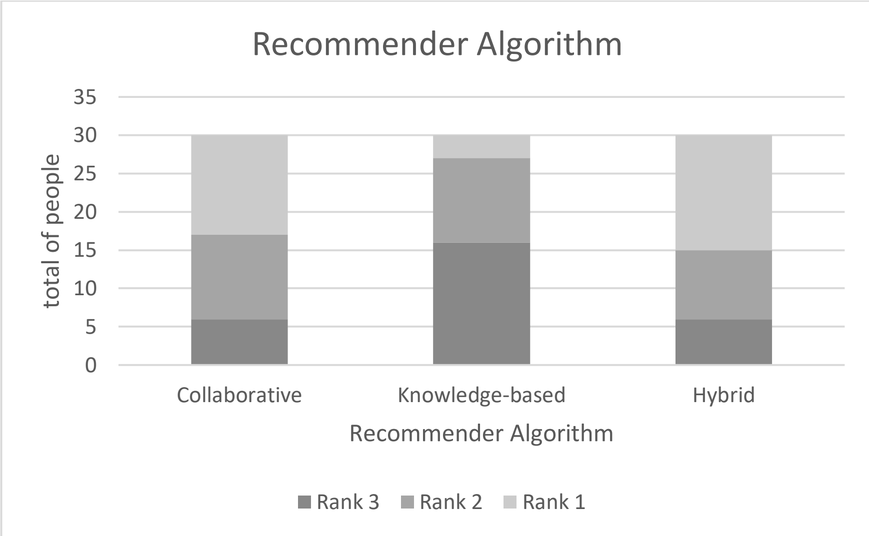



Fig. 4. Results of Evaluation of Algorithm Recommendations

Based on fig. 4. users prefer the results of recommendations on collaborative and hybrid. Because based on the data in fig. 4 the percentage of rank 1 in both methods shows a value of 43,333% for collaborative filtering methods and 50% for hybrid filtering methods. But the knowledge-based filtering method for ranking 1 is only 10%. Because the knowledge-based method of the recommendation results does not have many tourist destinations. While in collaborative and hybrid methods, users get many choices of tourist destinations. The resulting value is still low because the user testing is only 30 users and questions from the questionnaire submitted based on the previous paper which was later developed by the author.

*C. Results of Anova Test*

**ANOVA**

Ranking

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | ,000 | 2 | ,000 | ,000 | 1,000 |
| Within Groups | 154,000 | 6 | 25,667 | | |
| Total | 154,000 | 8 | | | |

Fig. 5. Results of Anova Test

Based on fig. 5. statistical analysis carried out by using Anova One Way test to test the superiority of algorithm recommendations with 95% confidence level and HO value of 0.05, that Anova test output for significance value (sig.) is 1. This means that the sig value > HO value. So it can be concluded that the method variable does not significantly influence the difference in ranking average of the three methods used.

## V. Conclusion

Based on the test results data in chapter IV, both testing the level of user satisfaction on group recommendations, testing algorithm recommendations, and Anova One Way testing. Users tend to strongly agree with the results of group tourism recommendations presented by the application. With this application, it is expected to be able to meet the wishes of users in determining tourist destinations in groups. Users can choose restrictions such as the costs and tourist categories they want.

The algorithm recommendations that are used have various recommendations in accordance with input from users. Collaborative and hybrid methods tend to be favored by users rather than knowledge-based methods. Because collaborative and hybrid methods produce more choices of tourist destinations than knowledge-based methods. However, in testing the statistical analysis with the Anova One Way test to test the superiority of recommendation algorithms, it is statistically insignificant on the difference in ranking average of the three methods used.

It is expected that further research will further develop and explore the three methods used. Can also be added to other methods, such as content-based filtering methods. So that the resulting recommendations are far more varied. Can also be added with other domains, such as restaurants, routes, and inns to get more complex recommendations.

## REFERENCES

[1] Sharda, N. (Ed.). (2009). *Tourism Informatics: Visual Travel Recommender Systems, Social Communities, and User Interface Design: Visual Travel Recommender Systems, Social Communities, and User Interface Design*. IGI Global.

[2] Baizal, Z. A., Widyantoro, D. H., & Maulidevi, N. U. (2016, October). Query refinement in recommender system based on product functional requirements. In *Advanced Computer Science and Information Systems (ICACSIS), 2016 International Conference on* (pp. 309-314). IEEE.

[3] Abdurahman Baizal, Z. K., Rahmawati, A. A., Lhaksmana, K. M., Mubarok, M. Z., & Qadrian, M. (2018). Generating Travel Itinerary Using Ant Collony Optimization. *Telkomnika*, *16*(3).

[4] Ricci, F., Rokach, L., & Shapira, B. (2015). Recommender systems: introduction and challenges. In *Recommender systems handbook* (pp. 1-34). Springer, Boston, MA.

[5] Baizal, Z. K. A., Iskandar, A., & Nasution, E. (2016, May). Ontology-based recommendation involving consumer product reviews. In *Information and Communication Technology (ICoICT), 2016 4th International Conference on* (pp. 1-6). IEEE.

[6] Baizal, Z. A., Widyantoro, D. H., & Maulidevi, N. U. (2016, October). Design of knowledge for conversational recommender system based on product functional requirements. In *Data and Software Engineering (ICoDSE), 2016 International Conference on* (pp. 1-6). IEEE.

[7] Mcginty, L., & Smyth, B. (2006). Adaptive selection: An analysis of critiquing and preference-based feedback in conversational recommender systems. *International Journal of Electronic Commerce*, *11*(2), 35-57.

[8] De Pessemier, T., Dhondt, J., Vanhecke, K., & Martens, L. (2015). TravelWithFriends: a hybrid group recommender system for travel destinations. In *Workshop on Tourism Recommender Systems (TouRS15), in conjunction with the 9th ACM Conference on Recommender Systems (RecSys 2015)* (pp. 51-60).

[9] Tran, T. (2007). Combining collaborative filtering and knowledge-based approaches for better recommendation systems. *Journal of Business and Technology*, *2*(2), 17-24.

[10] Baizal, Z. A., Widyantoro, D. H., & Maulidevi, N. U. (2016). Factors Influencing User's Adoption of Conversational Recommender System Based on Product Functional Requirements. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, *14*(4), 1575-1585.

[11] Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge & Data Engineering*, (6), 734-749.

[12] Sebastia, L., Garcia, I., Onaindia, E., & Guzman, C. (2009). e-Tourism: a tourist recommendation and planning application. *International Journal on Artificial Intelligence Tools*, *18*(05), 717-738.

[13] Bogers, T., & Van den Bosch, A. (2007, October). Comparing and evaluating information retrieval algorithms for news recommendation. In *Proceedings of the 2007 ACM conference on Recommender systems* (pp. 141-144). ACM.

[14] Kim, B. M., Li, Q., Kim, J. W., & Kim, J. (2004, August). A new collaborative recommender system addressing three problems. In *Pacific Rim International Conference on Artificial Intelligence*(pp. 495-504). Springer, Berlin, Heidelberg.

[15] Tan, X., & Pan, P. (2012). A contextual item-based collaborative filtering technology. *intelligent Information Management Journal*, *4*, 85-88.

[16] Borràs, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert Systems with Applications*, *41*(16), 7370-7389.

[17] De Pessemier, T., Dooms, S., & Martens, L. (2014). Comparison of group recommendation algorithms. *Multimedia tools and*

*applications*, *72*(3), 2497-2541.

[18] Masthoff, J. (2004). Group modeling: Selecting a sequence of television items to suit a group of viewers. In *Personalized digital television* (pp. 93-141). Springer, Dordrecht.

[19] Setia, I. P. (2017). IMPLEMENTASI ANALITICAL HIERARCHY PROCESS (AHP)-TOPSIS DAN BORDA PEMILIHAN TANAMAN OBAT UNTUK PENYAKIT BATUK. *Jurnal Elektronik Nasional Teknologi dan Ilmu Komputer*, *1*(01).