

# A Comparative Study of People-to-People Recommender Algorithms in Hybrid Method

Elizabeth Irenne Yuwono<sup>1</sup>  
and Nehemia Sugianto<sup>2</sup>  
Ciputra University, Indonesia  
<sup>1</sup>eirene@ciputra.ac.id  
<sup>2</sup>nsugianto@ciputra.ac.id

**Abstract** - Based on Internet Live Stats data, more than 2 billions information has been accessed daily of internet user daily. Amatriain said that trend of searching method has been deprecated. Trend of recommendation method has taken the position nowadays. Information must be processed to be transformed into recommendation and recommendation will reveal hidden information to the right position. One of recommendation method is people-to-people recommendation because one of the most accessed media in internet is social media. This research will discuss about comparison between hybrid algorithm in people-to-people recommendation. There are three algorithms which will be compared: hybrid content-collaborative reciprocal (without using weight) algorithm, hybrid content-collaborative reciprocal (using weight) algorithm, and interactive-based + decision tree algorithm. These algorithm will be implemented in recommending workout partner using "FitParners", Android-based mobile application and used 200 respondents (active or not active workout). The result will be compared based on execution time in generating the result and accuracy level of recommendation. Interactive-based + decision tree algorithm has best execution time which is 1726 ms. Hybrid content-collaborative reciprocal (using weight) algorithm has best accuracy level which is 76,45%.

**Keywords** - Content-Collaborative Reciprocal, Hybrid Filtering, Interactive-Based and Decision Tree, Recommender

**System**

## I. INTRODUCTION

Based on Internet Live Stats, there are more than 2 billion information read on website daily. Meanwhile, Amatriain said that the trend of searching has shift to the trend of recommending. Information will be processed into recommendation, and data searching turn into the finding of new information.

Many successful business has received the benefit of recommender system; 70% of movie watched on Netflix comes from recommendation, Google News has received 38% additional reader since using recommender system, 35% Amazon sales is from recommendation, and 28% of Choicestream user buy its music of the recommendation list (Amatriain, 2014).

Although recommender system has been developed since years ago, there are still unsolved problems. One of them is in finding the right method for people-to-people recommendation. Currently there are three main methods, they are content-based, collaborative, and hybrid filtering. Collaborative filtering is recommendation method based on people interaction with item or other people. On the other hand, content-based filtering is a recommendation method based on item features. And then the result is presented in ranking list. The last approach is hybrid filtering, where two previous methods are combined to complete each other weakness.

This research compared hybrid-filtering algorithms on social media for finding workout partner. The chosen algorithms are hybrid content-collaborative reciprocal and interactive-based combined with decision tree.

## II. HYBRID FILTERING

There are couple of approaches in hybrid filtering; linear combination, sequential combination, and item-based clustering (Li, 2002). The first approach combine both content-based and collaborative filtering by ranking result. Meanwhile sequential combination integrates both filters by combining the calculation result. The third combine the two filters by incorporating item information rating to calculate items similarity.

### A. Hybrid Content-Collaborative Reciprocal

This algorithm consists of three phases:

#### 1. Content-Based Filtering

In the first phase, algorithm search for users that are similar (SU) with current user U using profile attributes and distance metric.

#### 2. Collaborative Filtering

And then algorithm examines the interaction of every SU, candidate users (CU) will be retrieved based on their interaction similarity with SU.

#### 3. Item-Based Clustering

Candidate user list is recommended to U.

This research carried out this algorithm in two manners; with and without weight. The first is to execute the algorithm with given parameters directly, and the other is to give weight to each parameter and then normalized them to 0 to 1 scale before executing the algorithm.

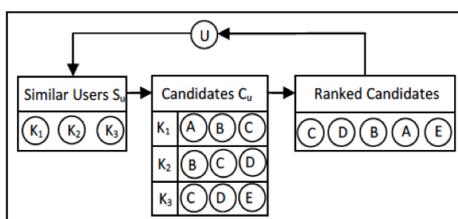


Fig. 1 The Scheme of Hybrid Content-Collaborative Reciprocal Algorithm

### B. Interaction-Based + Decision Tree

This algorithm is the optimized version of interactive-based algorithm, it consists of three phases:

#### 1. Interaction-Based

At first, algorithm will produce a list of recommended users, whereas the successful and unsuccessful interactions are ranked. The rank is determined based on implicit rating on the frequency of user conversation.

#### 2. Decision Tree

After that a new list of users will be retrieved based on user profile information using decision tree.

#### 3. Linear Combination

The final result is obtained from the combination of the first and second list.

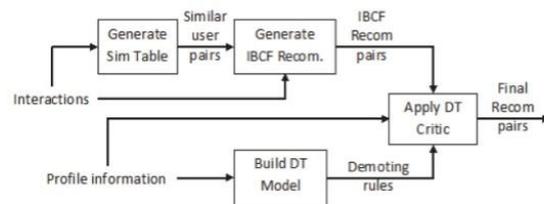


Fig. 2 The Scheme of Interactive-Based + Decision Tree Algorithm

## III. RESEARCH METHOD

The research was conducted on 200 respondents where 60% is active sportsman and 40% is inactive. All respondents are located in Surabaya and there are no gender or age boundaries. Active sportsman is defined as someone who exercises for minimum once a week, and then everyone besides that are considered as inactive. The sport activities examined are not limited in special sport venue, such as fitness center. Provided that the sport type is only limited to strength training, callisthenic, soccer, running, swimming, and badminton.

### A. Data Collection

The research data was collected online through Android based mobile application called FitPartners (Nagara, 2016). Data collection phase starts from August to

September 2016, where the data were retrieved from main database. Assuming that 200 respondents were selected randomly, therefore FitPartners users who registered before and after collection period are not considered. There are two groups of data that we collected; they are content and collaborative parameter. Table I and II shows the parameter lists, provided that they are from previous research with additional parameters to improve recommendation accuracy.

**TABLE I  
CONTENT PARAMETER GROUP**

Variable	Indicator
Physical similarity	- Age - Gender - Body type
Activity similarity	- Sport preference - Profession
Experience similarity	- Duration of sport activity (years) - Level of expertise
Location similarity	- Distance adjacency

**TABLE II  
COLLABORATIVE PARAMETER GROUP**

Variable	Indicator
Verbal communication	- Conversation intensity in chats
Non verbal warmth	- Frequency of positive emoticons in chats
Information finding	The activity of looking on: - Profile of partner candidate - Photo of partner candidate - Timeline of partner candidate
Reciprocity	- Ratio of the conversation and response frequency
Fondness	- Appraisal or positive rate on partner candidate
Shared network	- Friend of partner candidate
Communication satisfaction	- Request as friend - Challenge participation - Challenge assessment on participant

**B. Algorithm Implementation**

Three algorithms will be compared; hybrid content-collaborative reciprocal with and without weight, also interactive-based combined with decision tree. Therefore three functions are composed to represent those three using PHP language and MySQL database. In addition, every algorithm will generate 10 users as recommendation to one respondent.

**C. Validation and Evaluation**

The recommendation results of all algorithms were validated based on execution time and result accuracy. Validation by execution time was performed internally. In detail, every time the function was called, the execution time was recorded as validation data. On the other hand, the accuracy validation data was obtained by emailing every respondent. The email purpose is to obtain the number of suitable candidates of every recommendation from respondent.

**IV. RESULT AND ANALYSIS**

The validation results of execution time and recommendation accuracy are shown in table III and IV consecutively. Algorithm 1 represents hybrid content-collaborative reciprocal without weight, and then the same algorithm with weight is represented by algorithm 2, and lastly algorithm 3 is defined by interactive-based + decision tree.

**TABLE III  
ALGORITHMS COMPARISON BASED  
ON EXECUTION TIME**

Validation	Alg. 1	Alg. 2	Alg. 3
Average execution time	2842 ms	3307 ms	1726 ms

Based on execution time, algorithm 3 is the fastest in execution time; meanwhile the other two needs longer time to retrieve recommendation results. Assuming that algorithm 2 composed of interactive-based filtering and decision tree, there are three affecting factor in its short execution time: 1) dataset filtered from interactive-based is smaller (interaction data is smaller than users profile) therefore the time needed to determine

the recommendation by decision tree is shorter, 2) decision tree required shorter execution time, and 3) there is only one sorting process in decision tree. On the other hand, algorithm 1 and 2 has more complex process: 1) longer time is needed content-based filtering phase since every user similarity (both friend and not friend) has to be calculated, 2) provided that the filtering result from content-based phase is more than previous algorithm, collaborative phase need more time to filter them, and 3) additional time is required to sort the result from both filter.

**TABLE IV**  
**ALGORITHMS COMPARISON BASED ON**  
**RECOMMENDATION ACCURACY**

Validation	Alg. 1	Alg. 2	Alg. 3
Total of accepted recommendation	1453/2000 (72.65%)	1529/2000 (76.45%)	1099/2000 (54.95%)
Total of rejected recommendation	547/2000 (27.35%)	471/2000 (23.55%)	901/2000 (45.05%)
Average accepted recommendation per respondent	7/10 (70%)	8/10 (80%)	5/10 (50%)
Average rejected recommendation per respondent	3/10 (30%)	2/10 (20%)	5/10 (50%)

The result of table IV was analyzed based on accepted and rejected recommendation. Out of 200 respondent questioned during validation phase, 17 did not reply back, therefore their answer was considered as none of the candidates were accepted (accepted partner = 0).

Based on the comparison of recommendation accuracy, algorithm 2 has the best accuracy where 76, 45% recommendation was accepted. There are a couple of affecting factors, they are: 1) the first filtering is based on content similarity and not interaction similarity. Content similarity will result in possible accepted candidate based on the premise that someone will be more comfortable to be introduced to people who are similar to him or her. Moreover similar people have similar fondness normally, 2)

Prioritizing friend over stranger in producing ranked list provide stronger recommendation. This is supported by the premise that friend of our friend who is similar to us has stronger possibility to be our friend compared to stranger who is similar to us, and 3) The application of weight in calculating candidate similarity strengthens the validity of recommendation, assuming that the filtering is based on the user similarity priority. One-person similarity with others does not always matches, some similarity factors might have different weight according to the user preference.

## V. CONCLUSIONS

Based on the implementation result and analysis, we conclude that the best algorithm for people-to-people recommendation is hybrid content-collaborative reciprocal. Provided that three algorithms execution time is within normal time limit (in second), therefore the best algorithm was chosen based on recommendation accuracy. Moreover we should considerate the fact that nonetheless accuracy of recommendation is the main factor in determining whether a recommendation is acceptable or not.

And yet, this algorithm still needs some optimization in its execution time. One of the solutions is to execute content filtering before recommendation request is submitted. User contents were more reliant to changes, while on the other hand interaction or collaborative data is more likely to changes in faster rate. Therefore algorithm can be started directly to collaborative phase, where content filtered result can be retrieved before algorithm start.

## VI. ACKNOWLEDGMENT

The authors wish to acknowledge Daniel Wonohadidjojo for guiding and giving insight to us in the preliminary research. We also acknowledge Herlambang Satya Nagara for allowing us to use his product “FIT PARTNERS” as study case. And last, we would like to acknowledge Ciputra University for funding this research especially the

Research and Development Department.

## REFERENCES

**(Arranged in the order of citation in the same fashion as the case of Footnotes.)**

- [1] Akehurst, J., Koprinska, I., Yacef, K., Pizzato, L., Kay, J., and Rej, T. (2011). "CCR: a content-collaborative reciprocal recommender for online dating, in IJCAI'11". Twenty-Second International Joint Conference on Artificial Intelligence, AAAI Press, 2011, pp. 2199-2204.  
<<http://dl.acm.org/citation.cfm?id=2283765>>. Accessed 4 October 2016.
- [2] Amatriain, X. (2014). "Recommender Systems: Collaborative Filtering and Other Approaches".
- [3] Galland, A. (2010). "Recommender Systems".  
<<http://alban.galland.free.fr/Documents/Enseignements/INF396/recommendersystems-slides.pdf>>. Accessed 10 October 2016.
- [4] Huang, Z., Zeng, D., and Chen, H. "A comparison of collaborative-filtering recommendation Algorithms for e-commerce". IEEE Intelligent Systems, Vol. 22, No. 5, pp. 68-78,  
<<http://dl.acm.org/citation.cfm?id=1304521>>. Accessed 10 October 2016.
- [5] Krzywicki, A., Wobcke, W., Cai, X., Bain, M., Mahidadia, A., and Compton, P. (2012). "Using a critic to promote less popular candidates in a people-to-people Recommender system". in Twenty-Fourth Conference on Innovative Applications of Artificial Intelligence (IAA/I 2012), ResearchGate, 2012, Vol. 3, pp. 12-4863.  
<[https://www.researchgate.net/publication/243963176\\_Using\\_a\\_Critic\\_to\\_Promote\\_Less\\_Popular\\_Candidates\\_in\\_a\\_People-to-People\\_Recommender\\_System](https://www.researchgate.net/publication/243963176_Using_a_Critic_to_Promote_Less_Popular_Candidates_in_a_People-to-People_Recommender_System)>. Accessed 10 October 2016.
- [6] Nagara, H.S. and Sugianto, N. (2016). "Implementation of Recommendation System for Exercise Partners in Fitness Centre on FIT PARTNERS Design

Application Using Content-Collaborative Reciprocal Algorithm". Jakarta: ACM New York.

- [7] Noij, G.J. (2008). "Recommender Systems: An Overview". Bachelor Thesi, Pennsylvania State University, United States.