Asia Pacific Journal of Information Systems
Vol. 29 No. 2 (June 2019), 203-216

ISSN 2288-5404 (Print) / ISSN 2288-6818 (Online)
https://doi.org/10.14329/apjis.2019.29.2.203

Assessing Personalized Recommendation Services Using Expectancy Disconfirmation Theory

Il Young Choi[a], Hyun Sil Moon[b], Jae Kyeong Kim[c,*]

[a] Visiting Professor, Graduate School of Business Administration & AI Research Center, Kyung Hee University, Korea
[b] Lecturer, Graduate School of Business Administration & AI Research Center, Kyung Hee University, Korea
[c] Professor, School of Management, Kyung Hee University & AI Research Center, Kyung Hee University, Korea

ABSTRACT

There is an accuracy-diversity dilemma with personalized recommendation services. Some researchers believe that accurate recommendations might reinforce customer satisfaction. However, others claim that highly accurate recommendations and customer satisfaction are not always correlated. Thus, this study attempts to establish the causal factors that determine customer satisfaction with personalized recommendation services to reconcile these incompatible views. This paper employs statistical analyses of simulation to investigate an accuracy-diversity dilemma with personalized recommendation services. To this end, we develop a personalized recommendation system and measured accuracy, diversity, and customer satisfaction using a simulation method. The results show that accurate recommendations positively affected customer satisfaction, whereas diverse recommendations negatively affected customer satisfaction. Also, customer satisfaction was associated with the recommendation product size when neighborhood size was optimal in accuracy. Thus, these results offer insights into personalizing recommendation service providers. The providers must identify customers' preferences correctly and suggest more accurate recommendations. Furthermore, accuracy is not always improved as the number of product recommendation increases. Accordingly, providers must propose adequate number of product recommendation.

Keywords: Expectancy Disconfirmation Theory, Customer Satisfaction, Disconfirmation, Accuracy, Diversity, Personalized Recommendation Service

Ⅰ. Introduction

Although a great many new products launch every year, customers have difficulty finding products that they like. For example, there are approximate 34 million books available on Amazon. So, personalized recommendation service is considered as a solution for helping its customers navigate this volume.

*Corresponding Author. E-mail: jaek@khu.ac.kr Tel: 8229619355

Internet leaders such as Amazon (Linden et al., 2003), Google (Das et al., 2007) and Netflix (Bennett and Lanning, 2007) offer personalized recommendation services as an important method of maintaining a sustainable competitive advantage. In general, personalized recommendation services provide customers with recommendation lists based on their preferences to broaden their purchasing patterns (Lawrence et al., 2001). However, recommendations that do not meet customer expectations may lead to rejection of the recommendation and even contempt for the recommendation service (Fitzsimons and Lehmann, 2004).

Most studies on personalized recommendation services have focused on improving the accuracy of the recommendations, which typically generate top-N recommendation lists based on similarities among customer preferences (Ahn et al., 2006; Cho and Kim, 2004; Cho et al., 2002; Herlocker et al., 2000; Kim et al., 2009; Kim et al., 2010; Moon et al., 2017; Shardanand and Maes, 1995; Sohn and Suh, 2006; Suh et al., 2014). However, highly accurate recommendation lists may not help extend customers' purchase patterns because more accurate recommendations indicate narrower recommendation lists (Adomavicius and Kwon, 2008; Adomavicius and Kwon, 2012; Chandra et al., 2006; Gan and Jiang, 2013; Zhou et al., 2010). In other words, there is an accuracy-diversity dilemma with personalized recommendation services.

Nonetheless, good personalized recommendation services satisfy customers (Chen et al., 2010; Jiang et al., 2010). Although it has been assumed that personalized recommendation services satisfy customers by recommending products that suit their preferences, there is an argument about the relationship between the accuracy of recommendations and customer satisfaction. Liang et al. (2007) believe that accurate recommendation might reinforce customer satisfaction. However, others claim that highly accurate recommendations and customer satisfaction are not always correlated (McNee et al., 2002; Willemsen et al., 2011; Ziegler et al., 2005). Furthermore, Ziegler et al. (2005) claim that diverse recommendations can also influence customer satisfaction.

It is not clear whether accuracy and diversity affect customer satisfaction in reviewing the literature. We believe it is critical to address this issue because the adoption of recommender systems in enterprises continues to increase, which serves as a driver for increasing sales (Thongpapanl and Ashraf, 2011). Thus, this study attempts to establish the causal factors that determine customer satisfaction with personalized recommendation services to reconcile these incompatible views. To explore this question, we first employ expectancy disconfirmation theory (EDT). According to EDT, customer satisfaction is decided by the difference between the quality level expected by customers and the actual quality of the purchased products (Athiyaman, 1997; Bitner, 1990; Chong and Wong, 2005; Hill, 1995; Maxham, 2001; Zhao and Lu, 2012). In other words, customers are satisfied if the quality of the recommended products is higher than or equal to the quality level they expect. Second, we measure accuracy, diversity, and customer satisfaction using simulation experiments that are common in research on personalized recommendation services. Notably, previous studies have measured accuracy and diversity using simulation methods by asking a customer how satisfied he or she is with the recommendation list directly. Finally, we statistically analyzed the simulation output data to determine which factors affect customer satisfaction.

Our experiment results indicate that accuracy positively affects customer satisfaction, whereas diversity negatively affects customer satisfaction. Thus, we claim that accurate recommendations are important for improving customer satisfaction.

Ⅱ. Research Background

2.1. Expectancy Disconfirmation Theory

Expectancy disconfirmation theory (EDT) is widely used to elucidate information system (IS) continuance (Bhattacherjee, 2001; Roca et al., 2006). IS continuance intention is influenced by customer satisfaction, which is determined by the difference between perceived quality and expectation levels. Consequently, customer satisfaction has a positive effect on repurchase intentions and word-of-mouth.

The EDT model passes through five phases as follows (Oliver, 1980). First, customers form pre-purchase expectations regarding a specific product or service. Second, they purchase the product or service and then assess its quality. Third, they compare the perceived quality with their pre-purchase expectations and determine whether these expectations were confirmed. If the perceived quality is higher than or equal to their expectations, these expectations are either positively disconfirmed or confirmed. However, if the expectations are higher than the perceived quality, the expectations are negatively disconfirmed. Fourth, customers experience satisfaction or dissatisfaction based on the level of their disconfirmation. Finally, satisfied customers form repurchase intentions and spread positive word-of-mouth, whereas dissatisfied customers avoid subsequent repurchase and spread negative word-of-mouth.

Many customers post star ratings on retailers' websites regarding the products they have purchased. Star ratings are important cues for predicting initial expectation levels for recommended products from a ratings-based personalized recommendation service because recommendation services predict customers' purchase likelihood scores based on their neighbors' star ratings. Additionally, star ratings are important for measuring perceived post-purchase product quality because high and low ratings indicate positive and negative views of products, respectively (Mudambi and Schuff, 2010). Therefore, we compute disconfirmation as the average of the differences between predicted ratings and actual ratings. Disconfirmation is computed as follows:

$$Disconfirmation = \frac{1}{m}\sum_{i=1}^{n} y_i - f_i,$$

where m is the total number of the recommended products, and y_i and f_i, are the actual star ratings and the predicted star ratings, respectively.

2.2. Accuracy and Diversity Metrics of Personalized Recommendation Services

Various metrics are used to evaluate personalized recommendation services. These metrics are broadly classified as accuracy metrics and diversity metrics. The widely used accuracy metrics are recall, precision, and F1 (Moon et al., 2013; Sarwar et al., 2000; Tsai and Hung, 2012; Yu et al., 2004). Recall means how many actually purchased products are recommended products, whereas precision means how many of the recommended products belong to the actual customer purchase list. However, there is a trade-off between recall and precision when increasing the size of the recommendation set. $F1$ is used to produce universally comparable evaluation results and is defined as the harmonic mean of recall and precision. Recall, precision and $F1$ are computed as follows;

$$Recall = \frac{|\,Purchased\ Products \cap Recommended\ Products\,|}{|\,Purchased\ Products\,|}$$

$$Precision = \frac{|\,Purchased\ Products \cap Recommended\ Products\,|}{|\,Recommended\ Products\,|}$$

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

In recent years, several studies have measured diversity to evaluate the effectiveness of personalized recommendation services (Javari and Jalili, 2015; Lee and Lee, 2015; Moon et al., 2013). The well-known diversity metrics are Shannon entropy, Simpson concentration, and Renyi entropy (Moon et al., 2013). These metrics compute the percentage of products that are recommended as follows:

$$Shannon\ \ entropy = -\sum_{i=1}^{n} p_i \log p_i$$

$$Simpson\ \ concentration = \sum_{i=1}^{n} p_i^2$$

$$Renyi\ entropy\ = -\log \sum_{i=1}^{n} P_i^2$$

where p_i indicates the percentage of the recommendation list containing i^{th} products and n is the total number of products.

Ⅲ. Hypotheses Development

Customer satisfaction involves the assessment that customers make after purchasing the products or services (Calvo-Porral and Lévy-Mangin, 2015; Deng et al., 2010; Gerpott et al., 2001). It is important for companies to satisfy customers because satisfied customers are likely to repurchase a company's products and services and disseminate positive word-of-mouth.

In the literature, algorithms for personalized recommendation services were developed on the assumption that customer satisfaction increases as accuracy increases. some studies (Abdel-Hafez et al., 2014; Christoffel et al., 2015; Zhou et al., 2012) have shown that accurate recommendations lead to customer satisfaction. Moreover, Liang et al. (2007) have empirically verified that accuracy has a significant effect on customer satisfaction. In other words, an accurate recommendation increases the probability that a customer will find products that suit his or her preferences, which will theoretically lead to increased customer satisfaction. We thus hypothesize as follows:

H1: Accurate recommendations as a function of the number of recommended products positively influence customer satisfaction.

However, some studies have claimed that accuracy was not the only consideration when measuring the quality of the recommendation (Herlocker et al., 2004; Kaminskas and Bridge, 2017; McNee et al., 2006; Pu et al., 2011). Furthermore, McGinty and Smyth (2003), McNee et al. (2006), Smyth and McClave (2001), Ziegler et al. (2005), argue that a more diverse recommendation list increases the probability that products will be chosen by the customer; as a result, diverse recommendations will improve customer satisfaction. We thus hypothesize as follows:

H2: Diverse recommendations as a function of the number of recommended products positively influence customer satisfaction.

Ⅳ. Empirical Study

4.1. Dataset and Experiment Design

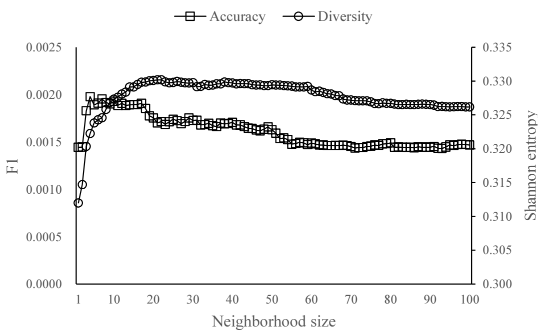
We acquired a movie review dataset from Amazon.com spanning from Jan 1, 2011 to Dec 31, 2011. The dataset contains star ratings (1 to 5 stars) that reflect customer-perceived product quality. Following simulation experiments widely used in personalized recommendation research (Im and Hars, 2007), we divided the dataset into a training set and a test set. The time between Jan 1, 2011 and Sep 31, 2011 was set as the training period, and the time between August 1, 2011 and Dec 31, 2011 was set as test period. The training set contained 569,879 transactions of 109,648 customers on 108,509 movies, and the test set contained 233,305 transactions of 47,987 customers on 67,823 movies. All experiments were performed on a PC running Windows 8.1 with an Intel Core i7-4500U processor running at 1.80 GHz, 8 GB RAM, 250 GB SSD and 500 GB HDD. All programs were implemented in SQL Server 2014.

To test the hypotheses, we developed a collaborative filtering (CF) – based recommender system known as one of the most successful personalized recommendation services (Cho and Kim, 2004; Cho et al., 2002; Hill et al., 1995; Kim et al., 2009; Konstan et al., 1997; Koren, 2010; Liu et al., 2014; Resnick et al., 1994; Zheng et al., 2010). We also used F1 and Shannon entropy to measure accuracy and diversity. In particular, we measured disconfirmation-based customer satisfaction through simulation experiments, although previous studies have used questionnaires to measure customer satisfaction regarding personalized recommendation services (Cremonesi et al., 2011; Ekstrand et al., 2014; Hijikata et al., 2009; Liang et al., 2007; Zins et al., 2004). Here, disconfirmation is computed as the average of the differences between predicted ratings and actual ratings.

In general, a CF system is characterized by three phases. In the first phase, a customer profile is created using past purchases, transactions or ratings. In the second phase, similarities between customers are computed. In the final phase, a top-N recommendation list is generated from products that a target customer's neighbors previously purchased. Thus, we determined the optimal neighborhood size and product recommendation size prior to testing the hypothesis because neighborhood size and product recommendation size affect the recommendation performance (Im and Hars, 2007; Shardanand and Maes, 1995). We then measured accuracy, diversity and customer satisfaction for each optimal size and tested the hypothesis.

4.2. Impact of Neighborhood Size

To determine the optimal neighborhood size for accuracy and diversity, several experiments were performed by varying neighborhood sizes from 1 to 100. <Figure 1> shows the results of our experiments. The results show that the recommendation performance (including for accuracy and diversity) increased as the number of neighbors increased. However, after a certain peak, the improvement gains diminished and the quality became worse. Accuracy and diversity were highest when the number of neighbors was 4 and 22, respectively. Thus, we performed various experiments to determine the optimal number of product recommendations when there are 4 and 22 neighbors.



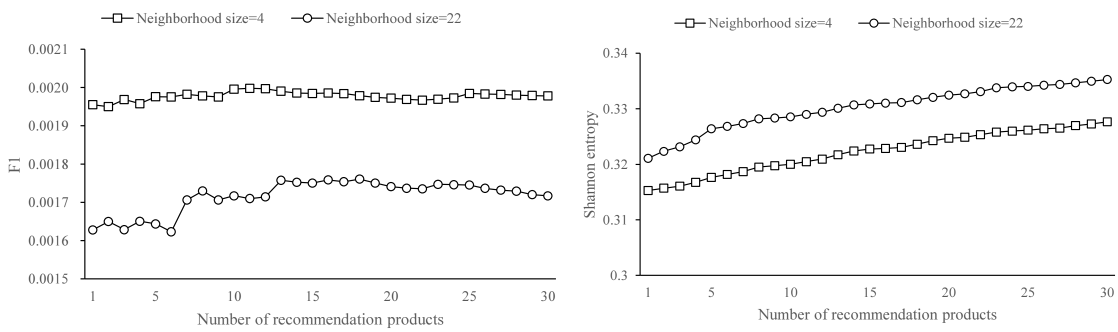
<Figure 1> Impact of Neighborhood Size

4.3. Impact of the Number of Product Recommendations

To find the optimal accuracy and diversity, simulations were conducted on a number of product recommendations that varied from 1 to 30 at the neighborhood size of 4 and 22. The results are shown in <Figure 2>. Accuracy generally increased with the number of product recommendations but then decreased after a certain peak. Accuracy was highest at the neighborhood size of 4 and 22 when the recommendation size was 11 and 18, respectively, whereas diversity increased as the number of recommended products increased. In other words, the total number of distinct products increased as the number of recommended products increased. Diversities at the neighborhood size of 4 and 22 were highest when the recommendation size was 30. Furthermore, accuracy at the neighborhood size of 4 was slightly higher than at the neighborhood size of 22, as the average of the F1 values at the neighborhood size of 4 and 22 were 0.002 and 0.0017, respectively. Additionally, diversity at the neighborhood size of 22 was higher than at the neighborhood size of 4, as the average of the Shannon entropy values at the neighborhood size of 4 and 22 were 0.32 and 0.33, respectively. These results are natural because accuracy and diversity are highest at the neighborhood size of 4 and 22, as shown in <Figure 2>. Therefore, we tested the hypothesis at the optimal neighborhood sizes and product recommendation sizes, respectively.

4.4. Experimental Results

The mean and standard deviation for accuracy, diversity, and customer satisfaction at (1) a neighborhood size of 4 and product recommendation size of 11, (2) a neighborhood size of 4 and product recommendation size of 30, (3) a neighborhood size of 22 and product recommendation size of 18, and (4) a neighborhood size of 22 and product recommendation size of 30 are listed in <Table 1>. The means for accuracy and diversity were between 0.0035 and 0.0049 and between 0.0131 and 0.0202, respectively. Additionally, the means of customer satisfaction were between -4.9160 and -4.8628. Accuracy at a neighborhood size of 4 and product recommendation size of 11 was highest (0.0049) and accuracy at a neighborhood size of 22 and product recommendation size of 18 was lowest (0.0035). Diversity at a neighborhood size of 22 and product recommendation size of 30 was highest (0.0202) and diversity at a neighborhood



(a) Accuracy (b) Diversity

<Figure 2> Accuracy and Diversity

size of 4 and product recommendation size of 11 was lowest (0.0131). Customer satisfaction at a neighborhood size of 4 and product recommendation size of 30 was highest (-4.8628), and customer satisfaction at neighborhood size of 4 and product recommendation size of 11 was lowest (-4.9160). Especially, products with highly predicted ratings are recommended regardless of the actual purchase. So, customer satisfaction is negative because it is defined as the average of the differences between predicted ratings and actual ratings.

To test H1 and H2, multiple regressions were performed under the four prior conditions. <Table 2> summarizes the results of multiple linear regressions for hypotheses H1 and H2. The table shows the unstandardized regression coefficient, the standardized regression coefficient, t-value, tolerance, and variance inflation factor (VIF) of each predictor and R^2, ad-

<Table 1> Descriptive Statistics of Accuracy, Diversity, and Customer Satisfaction

Condition	Variables	Mean	Standard deviation
(1) Neighborhood size = 4, Product recommendation size = 11 (*n* = 1,055)	Accuracy Diversity Customer satisfaction	0.0049 0.0131 -4.9160	0.0255 0.0118 0.3044
(2) Neighborhood size = 4, Product recommendation size = 30 (*n* = 1,055)	Accuracy Diversity Customer satisfaction	0.0048 0.0134 -4.8628	0.0232 0.0117 0.3870
(3) Neighborhood size = 22, Product recommendation size = 18 (*n* = 1,055)	Accuracy Diversity Customer satisfaction	0.0035 0.0195 -4.9017	0.0202 0.0168 0.3315
(4) Neighborhood size = 22, Product recommendation size = 30 (*n* = 1,055)	Accuracy Diversity Customer satisfaction	0.0036 0.0202 -4.8775	0.0205 0.0169 0.3723

<Table 2> Results of Multiple Regression Analysis

Condition	Dependent	Unstandardized Beta	Standardized Beta	*t*-value	Tolerance	VIF
(1)	Accuracy (H1)	3.155	.264	9.176**	.995	1.005
	Diversity (H2)	-6.042	-.234	-8.121**	.995	1.005
	R^2 = 0.133, Adjusted R^2 = 0.131, F = 80.643**, Durbin-Watson = 2.050					
(2)	Accuracy (H1)	3.119	.187	6.496**	.995	1.005
	Diversity (H2)	-10.029	-.302	-10.506**	.995	1.005
	R^2 = 0.134, Adjusted R^2 = 0.132, F = 81.453**, Durbin-Watson = 2.045					
(3)	Accuracy (H1)	3.215	.196	6.789**	.998	1.002
	Diversity (H2)	-5.617	-.284	-9.844**	.998	1.002
	R^2 = 0.124, Adjusted R^2 = 0.122, F = 74.552**, Durbin-Watson = 2.034					
(4)	Accuracy (H1)	3.203	.176	6.164**	.997	1.003
	Diversity (H2)	-7.058	-.321	-11.206**	.997	1.003
	R^2 = 0.140, Adjusted R^2 = 0.139, F = 85.833**, Durbin-Watson = 2.035					

Note: ** $p < 0.01$, * $p < 0.05$

<Table 3> One-way ANOVA of Customer Satisfaction

One-Way ANOVA	Sum of Squares	*df*	Mean Square	*F*	Sig.
Between Groups	1.801	3	0.600	4.890**	0.002
Within Groups	517.481	4216	0.123		
Total	519.281	4219			
condition			Mean Difference	Std. Error	Sig.
Condition (1)	Condition (2)		-0.0532**	0.0153	0.007
	Condition (3)		-0.0143	0.0153	0.832
	Condition (4)		-0.0385	0.0153	0.095
Condition (2)	Condition (1)		0.0532**	0.0153	0.007
	Condition (3)		0.0389	0.0153	0.090
	Condition (4)		0.0147	0.0153	0.819
Condition (3)	Condition (1)		0.0143	0.0153	0.832
	Condition (2)		-0.0389	0.0153	0.090
	Condition (4)		-0.0242	0.0153	0.471
Condition (4)	Condition (1)		0.0385	0.0153	0.095
	Condition (2)		-0.0147	0.0153	0.819
	Condition (3)		0.0242	0.0153	0.471

Note: ** $p < 0.01$, * $p < 0.05$

justed R^2, F, and Durbin-Watson of each model in linear regression analysis. Under the four conditions, each model is statistically significant ((1) $F = 80.643$, $p < 0.01$, (2) $F = 81.453$, $p < 0.01$, (3) $F = 74.552$, $p < 0.01$, and (4) $F = 85.833$, $p < 0.01$ for conditions (1), (2), (3) and (4), thereby predicting 13.3%, 13.4%, 12.4%, and 14.0% of the variance in customer satisfaction, respectively). Moreover, there is no multicollinearity between accuracy and diversity under the four condition ((1) tolerance = .995 and VIF = 1.005, (2) tolerance = .995 and VIF = 1.005, (3) tolerance = .998 and VIF = 1.002, and (4) tolerance = .997 and VIF = 1.003 for conditions (1), (2), (3) and (4)). The results show that accuracy positively and significantly affects the disconfirmation ($p < 0.01$), supporting H1. However, diversity negatively and significantly affects the disconfirmation ($p < 0.01$), thus H2 is not supported. These findings suggest that it is important to improve accuracy to satisfy the customer.

In addition, a one-way analysis of variance (ANOVA) was conducted to determine whether there was a significant difference in customer satisfaction under the four prior conditions. The Scheffe Post Hoc Test was used to identify multiple comparisons of group means. The results are presented in <Table 3> and indicated that there was a significant customer satisfaction difference between conditions ($F = 4.890$, Sig. = 0.002). The significant mean difference was found between condition (1) and condition (2) (mean difference = -0.0532, Sig. = 0.007), which indicated that customer satisfaction was associated with the recommendation product size when neighborhood size was optimal in accuracy.

Ⅴ. Discussion and Conclusion

A personalized recommendation service is a tool used to maintain competitive advantages. Many Internet leaders, such as Amazon, Google, and Netflix, offer the service to their customers. However, there are not only trade-offs between accurate recommendations and diverse recommendations but also continuing debates over which factor—accurate recommendations or diverse recommendations—has a more significant impact on customer satisfaction. Thus, we investigated which factors affect customer satisfaction by statistical analyses of simulation output data.

This study finds the following regarding adopting personalized recommendation services. First, we employed EDT to measure customer satisfaction with personalized recommendation services for the first time. In particular, we computed simulation-based disconfirmation as a metric for measuring customer satisfaction because the accuracy and diversity of personalized recommendation services were measured by a simulation method. Second, we investigated how neighborhood size impacts accuracy and diversity and then identified how product recommendation sizes impact the accuracy and diversity at the optimal neighborhood size. We found that both accuracy and diversity increased with neighborhood size, but they also decreased after a certain peak. In addition, we determined that accuracy increased gently and then decreased gently as the number of product recommendations increased at the optimal neighborhood size, whereas diversity increased consistently as the number of product recommendations increased at the optimal neighborhood size. Third, we identified which factors provide customer satisfaction. The results showed that accurate recommendations positively affected customer satisfaction, which is consistent with previous studies (Liang et al., 2007). However, diverse recommendations negatively affected customer satisfaction. Finally, we compared the mean of customer satisfaction at the optimal neighborhood size and product recommendation size and demonstrated that customer satisfaction increased as the product recommendation size increased at the optimal neighborhood size in accuracy.

Thus, these results offer insights into service providers. First, the providers will be able to increase sales volume by offering the products which suit customer preferences because accurate recommendations cause customer satisfaction. Second, the providers must propose an adequate number of product recommendation to the customers because accuracy is not always improved as the number of product recommendation increases.

This study has the following limitations. First, our experiments were conducted in laboratory environments, which are substantially different from the context of real-world information seeking. Therefore, more work must be performed to know whether the results hold true in the real world. Second, we used data from Amazon.com. More research with other product domains is needed to examine whether the findings in this study can be generalized. Third, the recommendation method adopted in this research was collaborative filtering. We are not sure whether other methods, such as content-based filtering and hybrid recommender systems, would result in the same findings. The comparison between content-based filtering and collaborative filtering in different domains may also be worth investigating in the future.

<References>

[1] Abdel-Hafez, A., Tang, X., Tian, N., and Xu, Y. (2014). *A reputation-enhanced recommender system*. Advanced Data Mining and Applications, Springer, Cham.

[2] Adomavicius, G., and Kwon, Y. (2008). Overcoming accuracy-diversity tradeoff in recommender systems: a variance-based approach. In *Proceedings of the 18th Workshop on Information Technology and Systems*.

[3] Adomavicius, G., and Kwon, Y. (2012). Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Transactions on Knowledge and Data Engineering, 24*(5), 896-911.

[4] Ahn, H., Han, I., and Kim, K. J. (2006). The product recommender system combining association rules and classification models: the case of *g* internet shopping mall. *Information Systems Review, 8*(1), 181-201.

[5] Athiyaman, A. (1997). Linking student satisfaction and service quality perceptions: the case of university education. *European Journal of Marketing, 31*(7), 528-540.

[6] Bennett, J., and Lanning, S. (2007). The Netflix prize. In *Proceedings of KDD Cup and Workshop*.

[7] Bhattacherjee, A. (2001). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly, 25*(3), 351-370.

[8] Bitner, M. J. (1990). Evaluating service encounters: the effects of physical surroundings and employee responses. *Journal of Marketing, 54*(2), 69-82.

[9] Calvo-Porral, C., and Lévy-Mangin, J. (2015). Switching behavior and customer satisfaction in mobile services: analyzing virtual and traditional operators. *Computers in Human Behavior, 49*, 532-540.

[10] Chandra, A., Chen, H., and Yao, X. (2006). *Trade-off between diversity and accuracy in ensemble generation*. Multi-Objective Machine Learning, Springer Verlag, Berlin.

[11] Chen, D. N., Hu, P. J. H., Kuo, Y. R., and Liang, T. P. (2010). A Web-based personalized recommendation system for mobile phone selection: Design, implementation, and evaluation. *Expert Systems with Applications, 37*(12), 8201-8210.

[12] Cho, Y. H., and Kim, J. K. (2004). Application of web usage mining and product taxonomy to collaborative recommendations in e-commerce. *Expert Systems with Applications, 26*(2), 233-246.

[13] Cho, Y. H., Kim, J. K., and Kim, S. H. (2002). A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications, 23*(3), 329-342.

[14] Chong, B., and Wong, M. (2005). Crafting an effective customer retention strategy: a review of halo effect on customer satisfaction in online auctions. *International Journal of Management and Enterprise Development, 2*(1), 12-26.

[15] Christoffel, F., Paudel, B., Newell, C., and Bernstein, A. (2015). Blockbusters and wallflowers: accurate, diverse, and scalable recommendations with random walks. In *Proceedings of the 9th ACM Conference on Recommender Systems, ACM*, 163-170.

[16] Cremonesi, P., Garzotto, F., Negro, S., Papadopoulos, A. V., and Turrin, R. (2011). *Looking for "good" recommendations: a comparative evaluation of recommender systems*. Human-Computer Interaction-INTERACT, Springer Verlag, Berlin.

[17] Das, A. S., Datar, M., Garg, A., and Rajaram, S. (2007). Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th International Conference on World Wide Web, ACM*, 271-280.

[18] Deng, Z., Lu, Y., Wei, K. K., and Zhang, J. (2010). Understanding customer satisfaction and loyalty: an empirical study of mobile instant messages in China. *International Journal of Information Management, 30*(4), 289-300.

[19] Ekstrand, M. D., Harper, F. M., Willemsen, M. C., and Konstan, J. A. (2014). User perception of differences in recommender algorithms. In *Proceedings of the 8th ACM Conference on Recommender Systems, ACM*, 161-168.

[20] Fitzsimons, G. J., and Lehmann, D. R. (2004).

Reactance to recommendations: when unsolicited advice yields contrary responses. *Marketing Science, 23*(1), 82-94.

[21] Gan, M., and Jiang, R. (2013). Constructing a user similarity network to remove adverse influence of popular objects for personalized recommendation. *Expert Systems with Applications, 40*(10), 4044-4053.

[22] Gerpott, T. J., Rams, W., and Schindler, A. (2001). Customer retention, loyalty, and satisfaction in the German mobile cellular telecommunications market, Telecomm. *Policy, 25*, 249-269.

[23] Herlocker, J. L., Konstan, J. A., and Riedl, J. (2000). Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, ACM*, 241-250.

[24] Herlocker, J. L., Konstan, J. A., Terveen, L. G., and Riedl, J. T. (2004). Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems (TOIS), 22*(1), 5-53.

[25] Hijikata, Y., Shimizu, T., and Nishida, S. (2009). Discovery-oriented collaborative filtering for improving user satisfaction. In *Proceedings of the 14th International Conference on Intelligent User Interfaces, ACM*, 67-76.

[26] Hill, F. M. (1995). Managing service quality in higher education: the role of the student as primary consumer. *Quality Assurance in Education, 3*(3), 10-21.

[27] Hill, W., Stead, L., Rosenstein, M., and Furnas, G. (1995). Recommending and evaluating choices in a virtual community of use. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM Press/Addison-Wesley Publishing Co., 194-201.

[28] Im, I., and Hars, A. (2007). Does a one-size recommendation system fit all? the effectiveness of collaborative filtering based recommendation systems across different domains and search modes. *ACM Transactions on Information Systems, 26*(1), 4.

[29] Javari, A., and Jalili, M. (2015). A probabilistic model to resolve diversity–accuracy challenge of recommendation systems. *Knowledge and Information Systems, 44*(3), 609-627.

[30] Jiang, Y., Shang, J., and Liu, Y. (2010). Maximizing customer satisfaction through an online recommendation system: a novel associative classification model. *Decision Support Systems, 48*(3), 470-479.

[31] Kaminskas, M., and Bridge, D. (2017). Diversity, serendipity, novelty, and coverage: A survey and empirical analysis of beyond-accuracy objectives in recommender systems. *ACM Transactions on Interactive Intelligent Systems (TiiS), 7*(1), 2.

[32] Kim, H. K., Kim, J. K., and Ryu, Y. U. (2009). Personalized recommendation over a customer network for ubiquitous shopping. *IEEE Transactions on Services Computing, 2*(2), 140-151.

[33] Kim, H., Ji, A., Ha, I., and Jo, G. (2010). Collaborative filtering based on collaborative tagging for enhancing the quality of recommendation. *Electronic Commerce Research and Applications, 9*(1), 73-83.

[34] Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., and Riedl, J. (1997). GroupLens: applying collaborative filtering to Usenet news. *Communications of the ACM, 40*(3), 77-87.

[35] Koren, Y. (2010). Factor in the neighbors. *ACM Transactions on Knowledge Discovery from Data, 4*(1), 1-24.

[36] Lawrence, R. D., Almasi, G. S., Kotlyar, V., Viveros, M., and Duri, S. S. (2001). *Personalization of supermarket product recommendations*. Applications of Data Mining to Electronic Commerce, Springer, New York.

[37] Lee, K., and Lee, K. (2015). Escaping your comfort zone: a graph-based recommender system for finding novel recommendations among relevant items. *Expert Systems with Applications, 42*(10), 4851-4858.

[38] Liang, T., Lai, H., and Ku, Y. (2007). Personalized content recommendation and user satisfaction: theoretical synthesis and empirical findings. *Journal of Management Information Systems, 23*(3), 45-70.

[39] Linden, G., Smith, B., and York, J. (2003). Amazon.com recommendations: item-to-item collaborative filtering. *IEEE Internet Computing, 7*(1), 76-80.

[40] Liu, H., Hu, Z., Mian, A., Tian, H., and Zhu, X. (2014). A new user similarity model to improve the accuracy of collaborative filtering. *Knowledge-Based Systems, 56*, 156-166.

[41] Maxham, J. G. (2001). Service recovery's influence on consumer satisfaction, positive word-of-mouth, and purchase intentions. *Journal of Business Research, 54*(1), 11-24.

[42] McGinty, L., and Smyth, B. (2003). On the role of diversity in conversational recommender systems. In *International Conference on Case-Based Reasoning* (pp. 276-290). Springer, Berlin, Heidelberg.

[43] McNee, S. M., Albert, I., Cosley, D., Gopalkrishnan, P., Lam, S. K., Rashid, A. M., Konstan, J. A., and Riedl, J. (2002). On the recommending of citations for research papers. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, ACM*, 116-125.

[44] McNee, S. M., Riedl, J., and Konstan, J. A. (2006). Being accurate is not enough: how accuracy metrics have hurt recommender systems. In *CHI'06 extended abstracts on Human factors in computing systems* (pp. 1097-1101). Acm.

[45] Moon, H. S., Kim, J. K., and Ryu, Y. U. (2013). A sequence-based filtering method for exhibition booth visit recommendations. *International Journal of Information Management, 33*(4), 620-626.

[46] Moon, H. S., Yoon, J. H., Choi, I. Y., and Kim, J. K. (2017). An exploratory study of collaborative filtering techniques to analyze the effect of information amount. *Asia Pacific Journal of Information Systems, 27*(2), 126-138.

[47] Mudambi, S. M., and Schuff, D. (2010). What makes a helpful review? A study of customer reviews on Amazon.com. *MIS Quarterly, 34*, 185-200.

[48] Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research, 17*(4), 460-469.

[49] Pu, P., Chen, L., and Hu, R. (2011). A user-centric evaluation framework for recommender systems. In *Proceedings of the fifth ACM conference on Recommender systems* (pp. 157-164). ACM.

[50] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J. (1994). GroupLens: an open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, ACM*, 175-186.

[51] Roca, J. C., Chiu, C., and Martínez, F. J. (2006). Understanding e-learning continuance intention: an extension of the technology acceptance model. *International Journal of Human-Computer Studies, 64*(8), 683-696.

[52] Sarwar, B., Karypis, G., Konstan, J., and Riedl, J. (2000). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd ACM Conference on Electronic Commerce, ACM*, 158-167.

[53] Shardanand, U., and Maes, P. (1995). Social information filtering: algorithms for automating "word of mouth". In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM Press/Addison-Wesley Publishing Co., 210-217.

[54] Smyth, B., and McClave, P. (2001). *Similarity vs. diversity*. Case-Based Reasoning Research and Development, Springer Verlag, Berlin.

[55] Sohn, J., and Suh, Y. M. (2006). Using degree of match to improve prediction quality in collaborative filtering systems. *Information Systems Review, 8*(2), 139-154.

[56] Suh, K. S., Lee, S., Suh, E. K., Kang, H., Lee, S., and Lee, U. K. (2014). Comparisons of popularity-and expert-based news recommendations: similarities and importance. *Asia Pacific Journal of Information Systems, 24*(2), 191-210.

[57] Thongpapanl, N., and Ashraf, A. R. (2011). Enhancing online performance through website content and personalization. *Journal of Computer Information Systems, 52*(1), 3-13.

[58] Tsai, C., and Hung, C. (2012). Cluster ensembles

in collaborative filtering recommendation. *Applied Soft Computing, 12*(4), 1417-1425.

[59] Willemsen, M. C., Knijnenburg, B. P., Graus, M. P., Velter-Bremmers, L. C., and Fu, K. (2011). *Using latent features diversification to reduce choice difficulty in recommendation lists*. RecSys 2011, 14-20.

[60] Yu, K., Schwaighofer, A., Tresp, V., Xu, X., and Kriegel, H. (2004). Probabilistic memory-based collaborative filtering. *IEEE Transactions on Knowledge and Data Engineering, 16*(1), 56-69.

[61] Zhao, L., and Lu, Y. (2012). Enhancing perceived interactivity through network externalities: an empirical study on micro-blogging service satisfaction and continuance intention. *Decision Support Systems, 53*(4), 825-834.

[62] Zheng, N., Li, Q., Liao, S., and Zhang, L. (2010). Which photo groups should I choose? A comparative study of recommendation algorithms in Flickr. *Journal of Information Science, 36*(6), 733-750.

[63] Zhou, T., Kuscsik, Z., Liu, J. G., Medo, M., Wakeling, J. R., and Zhang, Y. C. (2010). Solving the apparent diversity-accuracy dilemma of recommender systems. *Proceedings of the National Academy of Sciences of the United States of America, 107*(10), 4511-4515.

[64] Zhou, X., Xu, Y., Li, Y., Josang, A., and Cox, C. (2012). The state-of-the-art in personalized recommender systems for social networking. *Artificial Intelligence Review, 37*(2), 119-132.

[65] Ziegler, C. N., McNee, S. M., Konstan, J. A., and Lausen, G. (2005). Improving recommendation lists through topic diversification. In *Proceedings of the 14th International Conference on World Wide Web, ACM*, 22-32.

[66] Zins, A. H., Bauemfeind, U., Missier, F., Venturini, A., and Rumetshofer, H. (2004). *An experimental usability test for different destination recommender systems*. Information and Communication Technologies in Tourism, Springer Verlag, New York.

◆ About the Authors ◆

Il Young Choi

Il Young Choi is a research professor at Graduate School of Business Administration, Kyung Hee University. He obtained his MS and Ph.D. in MIS, and his BS in Economics from Kyung Hee University. His current research interests focus on personalized service, bigdata analysis, social network analysis, and business analytics. He has published numerous papers which have appeared in Information Technology and Management, International Journal of Information Management, Online Information Review, and so on..

Hyun Sil Moon

Hyun Sil Moon is a senior researcher at AI Business Research Center, Kyung Hee University. He obtained his M.S. and Ph.D. in MIS, and his B.S. in Business Administration from KyungHee University. His current research interests focus on big data analysis, recommender systems based on deep learning methods, and unstructured data mining. He has published numerous papers which have appeared in International Journal of Information Management, Online Information Review, Journal of Electronic Commerce Research, Asia Pacific Journal of Information System, and so on.

Jae Kyeong Kim

Jae Kyeong Kim is a professor at School of Management, Kyung Hee University. He obtained his M.S. and Ph.D. in Management Information Systems (MIS) from KAIST (Korea Advanced Institute of Science and Technology), and his B.S. in Industrial Engineering from Seoul National University. His current research interests focus on business analytics, recommender systems, big data analysis, and deep learning. He has published numerous papers which have appeared in IEEE transactions on Systems, Men and Cybernetics, IEEE transactions on Services Computing, Information & Management, International Journal of Information Management, Applied Artificial Intelligence, Artificial Intelligence Review, International Journal of Human-Computer Studies, Expert Systems with Applications, and Group Decision and Negotiations.

Submitted: January 14, 2019; 1st Revision: April 2, 2019; Accepted: April 23, 2019