DeLone과 McLean의 정보시스템 성공 모형을 통한 추천시스템 성공 요인 재구성

Reconfiguration of Recommender System Success with DeLone and McLean's Model of IS Success

권 오 병 경희대학교 경영대학 (obkwon@khu.ac.kr)

ABSTRACT

Recommender system is a core component of e-commerce. Correspondingly, metrics to evaluate the system performance have been developed and applied. However, even though we have lots of applications that have tried to adopt recommender systems, the dearth of successfully installed recommender systems for more than a decade leads us to a skeptical thinking that current metrics do not sufficiently indicate the recommender system success in business viability point of view. Hence, the purpose of this paper is to reconfigure measures for recommender system success. Adopting DeLone and McLean's amended model of information system success as the underlying framework, content analysis with intellectual properties on recommender systems was conducted to modify the currently used metrics. Then a model of recommender system success is proposed based on the newly identified metrics are compared with traditional metrics.

Keywords : Closed Innovation, Open Innovation, Schumpeterian Competition

Ⅰ. Introduction

Since recommender systems have emerged as an independent research area in the mid-1990s, recommender system's strategic impact has been discussed in e-commerce context (Schafer et al., 1999). After then, plenty of ideas have been proposed to build technically and at the same time economically acceptable recommendation systems.

Thanks to lots of efforts to improve the recommendation performance in terms of accuracy for more than a decade, not a few outcomes have been considered practically in industry. For example, the Google's Recommended Sources feature analyzes the user's feeds via Reader Trends and Web History to find feeds the user might like. CDnow.com was another successful example of collaborative-filtering-based recommendation systems in the entertainment industry. CDnow keeps track of customer's web browsing history such as reviewed or searched product list to provide individualized recommendation. By using eBay's recommender system called

This work was supported by Mid-career Researcher Program through NRF grant funded by the MEST (R01-2008-000-20696-0).
논문접수일 : 2010년 10월 13일; 게재확정일 : 2010년 11월 19일

Active Buyer's Guide, people can get recommendations to fit their needs and individual preferences. Other than these, lots of recommender systems have been installed in e-commerce sites and supported the customers who intended to purchases goods and services.

However, considering the volume of efforts and the number of studies on recommender systems, so far only a few recommender systems or algorithms have been successfully satisfied the business viability point of view. This may be caused by the fact that user test is quite expensive and actual recommendation samples are hard to be acquired (Mooney and Roy, 1999). Nevertheless, as Adomavicius and Tuzhilin (2005) stated, technical measures such as accuracy or precision often are not adequate to evaluate the acceptance and quality of recommender systems. When we are regarding adoption in the industry as the actual success of recommender system, finding what causes such scarcity of success should contribute to the right direction for enriching recommender system research. More generally speaking, this lack of standardized metrics is damaging to the progress of knowledge related to recommender systems (Herloker et al., 2004). Correspondingly, this is a hindrance to recommender system success in terms of business viability. However, to our knowledge, empirical or theoretical studies to examine what affect the success of recommender system have been very few.

Hence, the purpose of this paper is to propose a model of recommender system success in terms of business viability. DeLone and McLean's Model of Information System Success is adopted as the base model. To explore candidate constructs and measures of the model, we performed content analysis using the patents as intellectual properties which were enrolled from 2000 to 2008, because our idea is that granted, not just pended, patents are most relevant and objective to show the business viability of the idea of invented recommendation systems. As Herlocker et al. (2004) already have pointed out, most metrics used in recommender system literature still focus on accuracy, and ignore issues to what extent the system is viable actually.

The remaining part of this paper is organized as follows. In chapter 2, DeLone and McLean's model of information system success and legacy metrics to evaluate recommender system are reviewed as baseline theories. How the content analysis to extract metrics is performed using the patents and the results are delineated in chapter 3. Based on the findings, in chapter 4, we propose a theoretical model of recommender system success is proposed with implications. Lastly, we conclude in chapter 5.

II. Theoretical Background

1 DeLone and McLean's Model of Information Systems Success

There is no doubt that the first yet robust model which illustrates information system success is DeLone and McLean's model (1992). In the model, they pointed out two constructs, information quality and system quality, lead to user satisfaction and system use. Then user satisfaction and system use affect individual impact and then organizational impact.

However, there have been some argues and suggestions for modifications on DeLone and McLean's original model. Seddon attempted to modify DeLone and McLean's model by seeing system use as behavior, and additionally incorporate perceived usefulness developed by Davis (Davis et al., 1989).

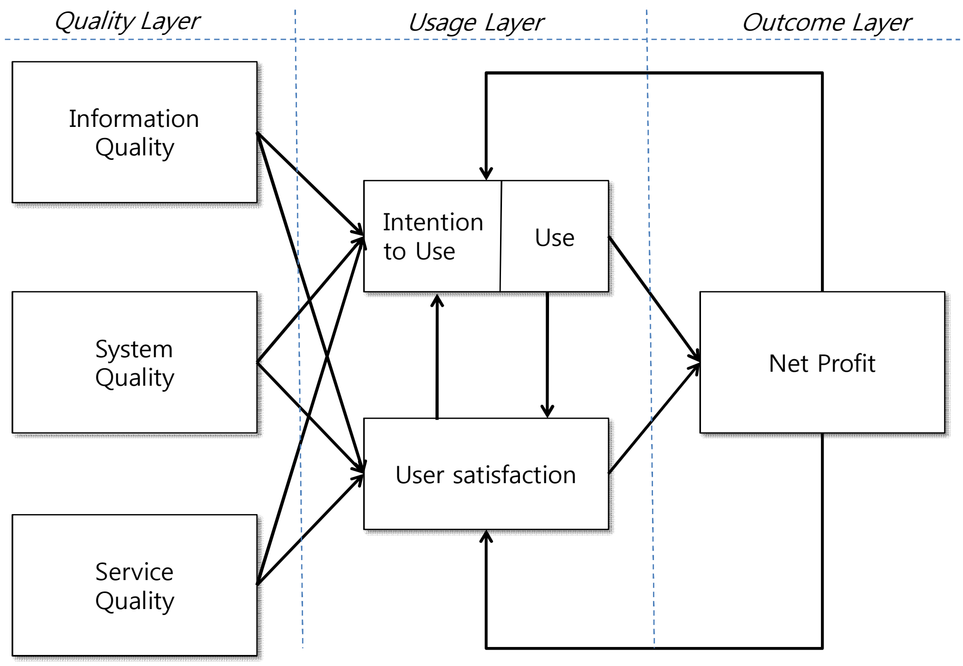
In the year of 2003, to cope with the debates, DeLone and McLean theoretically proposed extended model of information system success as shown in Figure 1. DeLone and McLean's model consists of three layers: quality layer, usage layer and outcome layer. In the quality layer, the main extension is that service quality, inspired from SERVQUAL, was newly added to the quality dimensions. Service quality is the overall support delivered by the service provider (Zeithmal et al., 1990). In SERVQUAL, reliability, assurance, tangibles, empathy, and responsiveness are proposed as main measures for rating services. However, these are not perfectly applicable to information systems. For example, services in information systems can be more likely to be intangible: physical facilities, equipment, staff appearance are not visible to the users in information system services. Therefore, DeLone and McLean's amended model adopted only three measures: assurance, empathy and responsiveness. The definitions are as follows:

- Assurance - the competence of the system and its security, credibility and courtesy
- Responsiveness – the willingness to help customers and provide prompt service
- Empathy – the ease of access, approachability and effort taken to understand customers' requirements

This viewpoint is consistent to other information system literatures (Jiang et al., 2002; Kwon et al, 2007).

Meanwhile, individual impact and organizational impact have been unified as net benefit. DeLone and McLean's original model has focused on management information system for organizational operation and control, not business-to-customer applications. However, under e-commerce setting, decomposing outcomes as organizational impact and individual impact is not quite appropriate because the customer as individual explicitly has nothing to do with the organization. Hence, in their modified model, under e-commerce context, they changed the focus from employee to customers.

Along with the shift of main focus, the measures have been changed. System quality includes usability, availability, reliability, adaptability and response time that are valued by users. Information quality contains personalized, complete, relevant, easy to understand, and secure information enough to let the information system users feel comfortable in using the recommender systems. Usage measures everything from a visit to the information system. Under e-commerce context, the quantity and quality of use can be estimated by observing the



〔Figure 1〕 DeLone and McLean's model of information system success (2003) and three layers

user's navigation patterns, number of site visits, and number of transactions executed. User satisfaction covers the entire customer experience cycle from information retrieval through purchase, payment, receipt, and service. Lastly, DeLone and McLean (2003) suggested cost savings, expanded markets, incremental additional sales, reduced search costs and timesavings as net benefit.

DeLone and McLean's model gives us another important implication: user acceptance in terms of intention to use, use and user satisfaction comes from quality layer and at the same time outcome layer. As for quality layer – usage layer link, DeLone and McLean (2003) and corresponding studies have shown lots of evidences that quality affects user satisfaction and intention to use. At the same time, outcome layer – usage layer link is found when net profit plays a prominent role of reinforcement or expectation to the user for further successful use of information systems. If it is true that net profit could be perceived, then well-organized information which makes the user convince that a certain information system will give the user desirable net profit may increase the intention to use. Actually the similar term 'perceived benefit' has been considered to explain the intention to use e-commerce system. For example, as Lee (2009) pointed out, perceived risk theory, five specific risk facets - financial, security/privacy, performance, social and time risk have been regarded as the elements in perceived benefit.

2 Evaluation of Recommender Systems

Recommender algorithms are usually classified into three categories: content-based recommendations, collaborative recommendations, and hybrid approaches. Content-based recommendations inform some items similar to the ones the user preferred in the past. Collaborative recommendations give a set of information on items that other people who have similar preferences positively perceived in the past. Since Tapestry system, collaborative filtering has been regarded as the most applied recommendation method (Goldberg et al., 1992). As a matter of course, hybrid approaches try to combine collaborative and content-based methods to take both advantages.

To show the feasibility of the invented recommendation systems, scholars have suggested a sort of performance metrics. Table 1 shows some popularly used metrics. Traditional metrics to evaluate the performance of recommender systems are accuracy, recall, precision, and efficiency (Olmo and Gaudioso, 2008). Accuracy metrics measure the quality of nearness to the truth or the true value achieved by a system. Perhaps accuracy may be the most frequently used measure in recommender system community. Recall represents the coverage of useful items the recommender system can obtain. In other

〈Table 1〉 Metrics found in the literature

Metrics	Representative references	Layer
Accuracy	Melamed et al., 2007 Sheng, 2007 van Setten et al., 2004 Olmo and Gaudioso, 2008	Quality
Prediction shift	Mobasher, et al., 2007	Quality
Hit ratio	Mobasher, et al., 2007 O'Mahony et al.,2004	Quality
Recall	Zanker, et al., 2007 Zanker and Jessenitschnig, 2009 Baeza-Yates and Ribeiro-Neto, 1999 Olmo and Gaudioso, 2008	Quality
Efficiency (computational cost)	Olmo and Gaudioso, 2008	Quality
Correlation	Olmo and Gaudioso, 2008	Quality
Privacy	Sheng, 2007	Quality
Reliability (false alarm rate)	Sheng, 2007	Quality
System Usability	Ricci, 2002 Hurley et al., 2007	Quality
Precision	Baeza-Yates and Ribeiro-Neto, 1999 Zanker and Jessenitschnig, 2009	Quality
Time decay	van Setten et al., 2004	Quality
Trust	Komiak and Benbasat, 2006 Wang and Benbasat, 2007 Komiak and Benbasat, 2008	Quality
Quality of decision	Ho and Tam, 2005 Tam and Ho, 2006	Quality
User satisfaction	Liang et al., 2007	Usage
User participation	Melamed et al., 2007	Outcome

words, this metrics measures the capacity of obtaining all the useful items present in the pool. Precision metric gives the share of successful recommendations from the total number of computed recommendations, while the Recall metric computes the ratio of hits and the theoretical maximum number of hits due to the testing set size. Efficiency refers to the computational cost of CF algorithms. Privacy-preserving CF provides protection against divulgence of personal information. Ratings (and even their existence) can reveal information about individuals' personal preferences. Reliability is the ability to detect and prevent malicious attacks that might make specific items appear more or less popular than they truly are (Sheng, 2007). Out of these metrics, novelty, serendipity, robustness, scalability, space coverage and sales diversity can be considered to evaluate the feasibility and applicability of recommender systems.

However, the focal parts of these metrics stress more on technical viability. Even though some studies have mentioned the importance of user acceptance issue or business viability, they seem to be based on practical experiences, not a theoretical background.

III. Content Analysis for Identifying Model of Recommender System Success

1 Patent search and criteria for inclusion

The main research question of this paper is why the current recommender systems are hardly accepted as actual business applications. Intellectual property such as patent can be regarded as a highly approved method that is more likely to be operated by business units. On behalf of the users as executing organizations who carefully consider the business and technical feasibility and then are willing to purchase the idea, the certified national organization evaluates the quality of the employed idea and then selectively approves them. The patent contains not only technical idea but also market feasibility in terms of novelty and excellence. In this regard, we believe the patent can bridge the gap which makes currently suggested recommender systems more acceptable by the users and system operators.

Hence, we performed content analysis to identify some metrics which are suitable for model of recommender system success in terms of business viability. To do so, relevant patents published in Korea between 2000 and 2008 were identified using computer and manual searches. Computerized searches were conducted using K-PEG (Korea Patent Evaluation & Grading) system databases using the following terms: recommender, recommendation, and method. K-PEG, which is officially run by a Korea governmental organization named Korea Intellectual of Patent Information (KIPI), is a computer-based system which estimates the quality of patent and searches patent information from patent database.

In keeping with the stated objectives of the present study we performed exhaustive search and initially identified 64 patents on recommender methods and/or systems. Fortunately, since all patents were available as it is, KIPI provided us with the whole documents as pdf format. To be included in the present review, several criteria had to be met. First, the metadata for patent articles had to be data based. Therefore, we requested KIPI agent to deliver spreadsheet which includes records about patent ID, classification code, patent application number, date of application, and name for each patent. Second, since some patents of which title include 'recommendation' or 'recommender' were substantially different from recommender system or recommendation method, they were excluded in the content analysis. For example, the initially considered patents contained patents on recommendation marketing, on-line recommendation for employment, recommendation word generation method, and even channel recommendation data delivery method. As a result, 12 patents were excluded for consideration. In addition, some patents had more than one classification code. Actually, there are three classification codes on recommender systems, G06F, G06Q and H04B, according to what the patent is focusing on. If a patent's content includes more than one focus, then the patent can be assigned into more than one classification. Hence, 3 patents were disregarded. Consequently, 49 patents were retained and selected for content analysis. Table 2 shows the number of articles identified in each classification code.

〈Table 2〉 Recommender system patents

Code	Area	Frequency	Percentage
G06F	Information processing system and software	24	48.9%
G06Q	Business Model using E-Commerce	24	48.9%
H04B	Mobile devices	1	2.2%

2 Overview of the content analysis process

The content analysis was conducted with the coding of identified patents. A multistage process was applied. First, information was obtained from each patent by two coders familiar with the business method and recommender systems and entered into a database. Next, we developed a coding scheme which includes identified metrics. Two raters with expertise in the recommender methods and business methods developed the coding scheme using guidelines set forth by Weber (1990) and Krippendorff (1980). Each rater reviewed metrics coded in stage one and independently developed a coding system consisting of broad categories of metrics. Similarities among the independently generated categories were noted, and after several iterations, consensus was reached on the final coding categories. Discrepancies were resolved through discussion. High agreement was obtained between coders with 82% agreement. Finally, the coding scheme consists of 5 categories, with 19 metrics as listed in Table 3.

3 Results of the content analysis

A total of 19 metrics were catalogued in the present review. Table 4 shows results of the content analysis. Overall, based on content analysis, we confirmed that DeLone and

〈Table 3〉 Coding taxonomy used in the review

Category	Metric
Recommending items	Tangibles Intangibles
System quality	Implementation level Ease of use Automatic acquisition of user data Response time Server overloading
Information quality	Accuracy Level of detail N recommendation (recommendation coverage) Information coverage
System quality	Responsiveness Assurance Empathy
Net profit	Marketability Novelty Maintainability Financial profit Increasing sales volume

McLean's amended model of information systems success could be applied to explain the recommender system. The most commonly studied metrics for recommender system success include system quality (30.8%), information quality (35.2%), service quality (6.9%), and net profit (27.1%). Among those, information quality is mostly considered metrics mainly at the mercy of information accuracy. The technology driven metrics on system quality and information quality proposed by literatures are still considered to show the system success. Therefore, this result is quite consistent with the fact that most of research papers conduct performance evaluations with accuracy or precision. Table 4 also shows the most highly represented metrics within each category. This information provides a more fine-grained assessment of the specific types of metrics considered in the patents.

Showing marketability seems to be very important for approving the pending patents. 37.2% of the patents which contain net profit issues apparently address marketability. To a certain extent, most patents identify target market and show the revealed or potential needs: 74.3% of the patents clearly identify recommending items such as multimedia/music, advertisement, keyword for information search on the web sites, SW components, antibiotic, and TV channels.

Novelty is one of the core values of computer-based systems (Zolt, 2007). Novel recommender system will include new combination of system components, bringing together new algorithms or participants, new way of distributing incentives to the participants, new transaction method, delivering information with new devices or new user interfaces, and having any potential to leapfrog the recommender service. More than a quarter (27.1%) of the patents, which address net profit issues, also support the novelty issue.

Maintenance cost accounts for the great part of information system development cost. Maintainability which results in reducing maintenance cost is related to efficient collection of recommendation knowledge. Lots of commercial web sites have reported that uploading abundant and relevant information in a prompt manner is very costly and hence a key factor toward manageable web sites. In the same manner, recommender systems need sufficient volume of valid recommendation knowledge to get rid of cold start problem and hence increase information quality. Technically speaking, many studies have attempted to resolve this problem by applying artificial intelligence techniques. However, when it comes to patents, they proposed a business or behavioral model approach to effectively acquire recommendation knowledge, mainly using an incentive system for the customers as knowledge donors, as well as relying on new knowledge acquisition techniques. This implies business method is apparently able to support technical method for implementing profitable recommender system.

Increasing sales volume (18.6%) mainly has something to do with enriching cross sales. Most of successful recommender systems display some related items to have users get motivated to purchase more. For example, Amazon.com provides the site visitors with a variety of ways to help them purchase associated items. For example, 'Customers Who Bought This Item Also Bought' is a typical cross selling service based on people's purchase history. 'Frequently Bought Together' highlights top-ranked cross selling items from the purchase history. 'Other Customers Suggested These Items' conducts relaxed query based on what the user type in the query input box. 'Customers Viewing This Page May Be Interested in These Sponsored Links' extends 'Customers Who Bought This Item Also Bought' recommendation to other web resources other than Amazon.com. 'What Do Customers Ultimately Buy After Viewing This Item?' and 'Your Recent History' shares other's and self experience with the user, respectively. 'Look for Similar Items by Category / Subject' is a keyword based recommendation service. These services contribute to increase the possibility of purchasing the clicked items and at the same time other related recommended items.

IV. Proposed Model of Recommender System Success

1 Amended model

By organizing the metrics identified in the patents, we developed a model of recommender system success in terms of business viability. The model is based on DeLone and McLean's model of information system success (DeLone and McLean, 2003). However, on top of the model, the followings are newly considered to the model of recommender system success.

First of all, we attempted to match the legacy metrics of recommender systems with DeLone and McLean's model. As they stated, "In the DeLone and McLean's model of information system success, systems quality measures technical success; information quality measures semantic success, and use, user satisfaction, individual impacts, and organizational impacts measure effectiveness success." (DeLone and McLean, 2003).

Second, the model focuses more on net profit perceived by the patent investors than actual net profit. One of the primary goals of patents is to be purchased as intellectual

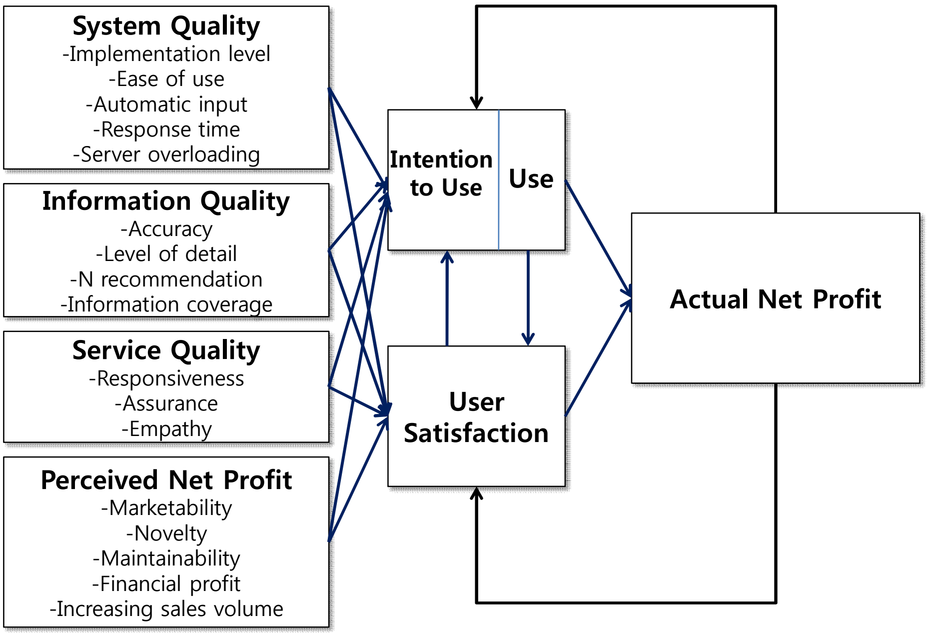
〈Table 4〉 Metrics found in patent

Category	Metric	Frequency	Percentile
Recommending items	Tangibles		
	Products	12	26.7%
	Intangibles		
	Multimedia / Music	12	26.7%
	Advertisement	4	8.9%
	Keywords	12	26.7%
	Misc. (SW components, Antibiotic, TV channel, place)	5	11.1%
System quality			30.8%
	Implementation level	7	14.3%
	Ease of use	6	12.2%
	Automatic acquisition of user data	17	34.7%
	Response time	16	32.7%
	Server overloading	3	6.1%
Information quality			35.2%
	Accuracy	27	48.2%
	Level of detail	7	12.5%
	N recommendation (recommendation coverage)	21	37.5%
	Information coverage	1	1.8%
Service quality			6.9%
	Responsiveness	2	18.2%
	Assurance	4	36.4%
	Empathy	5	45.5%
Net profit			27.1%
	Marketability	16	37.2%
	Novelty	12	27.9%
	Maintainability	3	7.0%
	Financial profit	4	9.3%
	Increasing sales volume	8	18.6%

Note: Subtotal of the metrics for each category do not need to equal to N=49 because one patent may have zero or more than one metric for each category.

property by investors. When a patent document is revealed by a potential investor, he/she will estimate the expected net profit if the recommender systems or methods are actually implemented and operated. Hence, to be actually used and satisfied by the end users, corresponding net profit must be perceived beforehand the profit is actually occurred. Hence, perceived net profit is the previous step of usage layer which contains intention to use and user satisfaction (See Figure 2).

Second, most of the conventional metrics shown in the original success model are consistently valid in recommender system cases. In general, system quality is measured with ease-of-use, functionality, reliability, flexibility, data quality, portability, integration, and importance. Information quality was measured in terms of accuracy, timeliness, completeness, relevance, and consistency (DeLone and McLean, 2003).



〔Figure 2〕 Metrics identified through content analysis

Third, we found that outcome consistency in measuring information quality cannot be directly reusable. More often than not, users may lose interest in unchanging recommendation and prefer the results which are different from the previous trial. Hence, consistency here should mean consistency with the user's present preference, not simply consistent result with the past trials.

Fourth, successful use of recommender system contains not only hit ratio, but also commitment to purchase the recommended items. Since system use is the "use of computer-generated reports by top management", recommendation system use can be "use of recommendation outcomes by the customer" (DeLone, 1983). From this viewpoint, the fact that lots of literatures have invited accuracy as a proxy measure to evaluate the usage performance may misleads the actual recommendation system success. Accuracy does not always mean that the user finally and really got recommended by the recommendation list. When we remind that "the nature of system use could be addressed by determining whether the full functionality of a system is being used for the intended purposes (DeLone and McLean, 2003)," recommender system success should not be isolated in just system use as hit ratio, but extended to system use as actual buying behavior.

Fifth, increasing sales volume is considered in perceived net profit construct. To do so, conformation with other business process should be considered as a facilitating condition. Commitment to purchase the recommended goods is possible when recommender system is logically and actually well incorporated in a whole business process. In general, the previous step of recommendation is to input the user profile or preference in a manual and/or automated manner to run a recommendation algorithm. Then N recommendations are listed for the user by running the recommendation algorithm. If a sort of N recommendations list has been displayed to the user, then the commerce system naturally leads the user move to ‘add to cart’ or ‘check out’ for payment. Rate of actual payment from ‘check out’ step may not the recommender system’s responsibility.

Sixth, maintainability is identified as a main metric of net benefits in the model. Maintainability of the recommendation knowledge is very crucial for making recommender system more sustainable and profitable by reducing the operating cost. As identified in the context analysis, and at the same time similar to e-commerce success, acquiring a sufficient volume of user preference, recommendation knowledge as cases and attacking resources is very costly. Main challenging issues such as cold start problem and sparcity problem come out from maintainability issue. First, cold start problem is the issue that the recommender algorithms cannot draw any inferences for users or items about which it has not yet gathered sufficient information. Cold start problem happens when providing meaningful advice and recommendations to anonymous or first-time users still remains a major challenge. This is quite behavioral, rather than technical, issue because the user tends to be reluctant to dedicating an amount of effort to make the recommender system sufficiently intelligent and hence usable. Let users rate on a set of products, collect click streams, infer from a content of webpage that the user saw, and just conduct user survey to let the user explicitly declare the preference in a manual manner have been suggested to cope with the cold start problem (Goldberg et al., 2001; Pierrakos et al., 2003; Schein et al., 2002; Rafter and Smyth, 2005; Zanker and Jessenitschnig, 2009). Meanwhile, legacy recommender system development efforts disregard path from customer satisfaction to customer use based on Sabherwal et al. (2006)’s advice.

It has not been sufficiently considered that human factors are crucial for recommender system success. A recommender system can be more acceptable if it explains how the system contributes to business success in the business model. For example, web users by nature seldom expose their preferences and profile because of privacy concern. This tendency prohibits showing explicitly their satisfaction on the service they have experienced. It is interesting that this tendency occurs from the scarcity of information, which in turn results in cold start problem. One candidate idea to cope with this concern is to reward those who give feedback to the recommender system. Actually, three out of 49 patents deal with the method how to encourage the users to get self-motivated to contribute to the recommender systems by providing recommending knowledge, which will reduce the maintenance cost.

A recommender system can be more acceptable when it is logically integrated with the other commercial systems simply because recommender system per se seldom adds value. Customers can more easily purchase a recommended item when payment and/or delivery modules are connected with recommender system in a transparent and seamless way. More often than not, recommender systems are transferred not just the recommender system itself but a bundle of business modules. For example, NHN, one of the biggest e-commerce enterprises in Korea, has developed a bunch of patents for a keyword-based advertisement content recommendation solution, which include automated keyword generation, as well as keyword recommendation.

As for accuracy, not a few actual reasons of lower accuracy are actually apart from technical issues. Rather, it is caused by user's attitude not to allow third party usage of their visiting history. As well known, user model or preference model is very important to increase the accuracy of recommendation, especially in case of incorporating people-to-people collaborative filtering method.

Trust or confidence issue needs to be taken care of as assurance dimension of service quality. From time to time, users are curious about the recommendation results. In particular, they want to know why the items are recommended, make sure if any other excellent items are omitted, or know what happens if the user's static or contextual preference changes. These concerns are hard to be expired only with recommender algorithms, simply because trust and confidence are very perceptual and attitudinal.

2 Implications to research

To observe how the focal points of recommender system performance are different from research and patent, we firstly found two outstanding papers which performed meta analysis: Herlocker et al. (2004) and Del Olmo and Gaodioso (2008)'s. Second, the metrics identified in the papers were

〈Table 5〉 Number of references for each metric in the literature

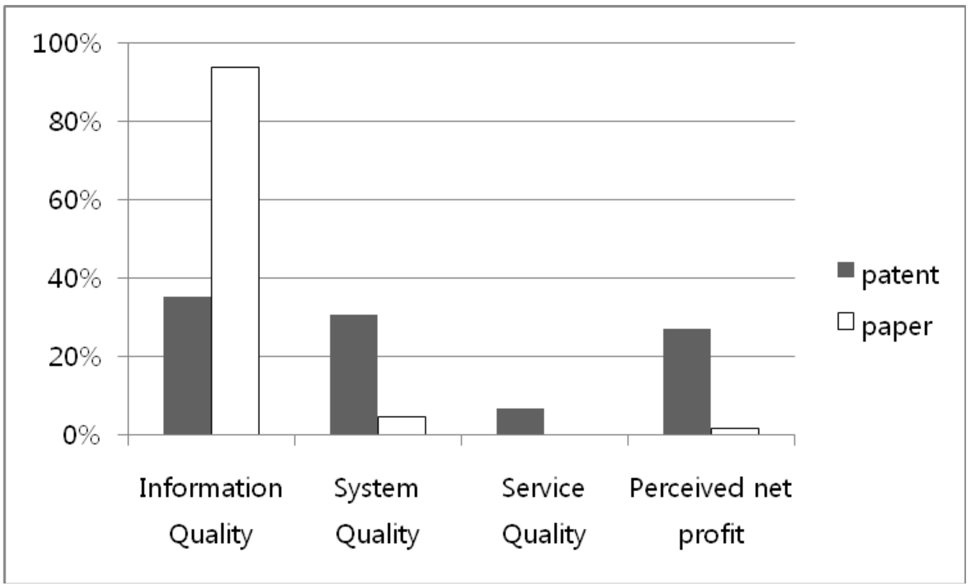
Selected paper	Metric	Number of references	Variables
Herlocker et al. 2004	Accuracy	18	Information quality
	Precision / Recall	13	Information quality
	Learning rate	1	System quality
	Novelty / Serendipity	3	Information quality
	Confidence	1	Information quality
	Usability	2	System quality
	User participation	7	Net profit
	Rank	2	Information quality
Del Olmo and Gaodioso, 2008	Accuracy	10	Information quality
	Precision / Recall	7	Information quality
	Rank	6	Information quality
	N Recommendations	1	Information quality
	Quality of decision	1	Net profit

enumerated. Then, third, we conducted content analysis to count the number of references which address each metric. Frequencies for each metric appeared in the papers are shown in Table 5. Note that the notion of user participation itself is not directly connected with net profits in terms of financial feasibility; rather it could be indirectly related. A successful outcome requests the users to act on the system's recommendations, and then actually purchase the recommended items (Herlock et al., 2004).

Based on the results in Table 4 and Table 5 on patent and literature respectively, frequencies are compared as shown in Figure 3. The figure clearly shows that research studies have relied much more on information quality such as recommendation accuracy, while what the patents have described are relatively more balanced. This could be natural if the readers are different then what they expect while reading and utilizing will be different: patent developers will focus on investors rather than end users who will be get recommended. In comparison with what the patents stress, lots of research papers, which mainly aim to show technical excellence, might sting to recommender system users. Nevertheless, the set of metrics that have been invited in the literature need to be more balanced by taking system quality, service quality and perceived net profit into account. Making compound metric can be an alternative to provide a balanced metric. Other than graphical comparison like Figure 3, we will not perform any statistical reasoning simply because the differences are very clear.

In sum, there are several discrepancies between patents and research studies.

(1) Patents show the balance between quality layer – usage layer and outcome layer – usage layer link, while research studies focus more on just quality layer – usage layer.



〔Figure 3〕 Comparison of percentiles of metrics

(2) As for net profit, patents address direct outcome such as cost and benefit, while research studies indirect outcome such as quality of decision and user participation.

(3) Target items to be recommended are more specific and realistic when it comes to the approved patents.

(4) Patents clearly show logical link to related business components such as user interface, payment and advertisement modules.

(5) Patents are more sensible to service quality, while research studies pay more attention to system quality and information quality only.

V. Concluding Remarks

Again, accurate recommendations alone do not guarantee success of recommender systems an effective and satisfying experience. Instead, systems are useful to the extent that they help users complete their tasks (Herlocker et al., 2004). We propose a theoretical model of recommender system success based on DeLone and McLean's model of information system success by conducting content analysis and comparative study with literature survey.

The result of content analysis gave us an insight of well-balanced compound metric. The balanced metric needs to contain system quality, service quality and perceived net profit, as well as information quality such as accuracy-related metrics. Even though accuracy is an essential ingredient to make recommender systems acceptable, accuracy itself hard to guarantee the success of recommender systems. On top of accuracy, recommender systems should convince end users or investors that the proposed recommender systems are substantially profitable. This includes promising market, maintainability for reducing cost, logical linkage with other business modules such as payment and cross sell marketing, novelty, and more.

Meanwhile, recommender system is regarded as an information service to the user. Hence, efforts for showing high-level responsiveness, assurance and empathy will increase usage level and hence net profit. Correspondingly, metrics must adopt this angle to have recommender system performance evaluation activity be more tightly linked to system success. These efforts will be profitable to make more comprehensive metrics to evaluate recommender systems.

We considered the Korean patent cases in this paper, even though metrics for system success are collected globally. Hence, let us note that the statistical results will be extended to mention the generality of the proposed IS success model. Moreover, an empirical study will be necessary to prove the applicability of the proposed model.

참 고 문 헌

[1] Adomavicius, G. and 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 and Data*

Eng., 17(6), 734–749.

[2] Ardissono, L., Goya, A., Petronea, G., Segnana, M., and Torassoa, P. (2003), Intrigue: Personalized Recommendation of Tourist Attractions for Desktop and Handset Devices, *Applied Artificial Intelligence*, 17(8-9), 687–714.

[3] Baeza-Yates, R. and Ribeiro-Neto, B. (1999), *Modern information retrieval*, Essex: Addison Wesley.

[4] Davis, F.D., Bagozzi, P., and Warshaw, P.R. (1989), User acceptance of computer technology: A comparison of two models, *Management Science*, 35(8), 982-1001.

[5] DeLone, W.H. and McLean, E.R. (1992), Information system success: The quest for the dependent variable, *Information System Research*, 3(1), 60-95.

[6] DeLone, W.H. and McLean, E.R. (2003), The DeLone and McLean Model of Information Systems Success: A Ten-year Update, *Journal of Management Information Systems*, 19(4), 9-30.

[7] DeLone, W.H.(1983), Determinants of Success For Small Business Computer Systems, Ph.D. Dissertation, UCLA.

[8] Goldberg, D., Nichols, D., Oki, B.M., and Terry, D. (1992), Using collaborative filtering to weave an information tapestry, *Communications of the ACM*, 35(12), 61-70.

[9] Goldberg, K., Roeder T., Gupta, D., and Perkins, C. (2001), Eigentaste: a constant time collaborative filtering algorithm, *Information Retrieval*, 4(2), 133–151.

[10] Ho, S.Y. and Tam, K.Y. (2005), An empirical examination of the effects of web personalization at different stages of decision making, *International Journal of Human–Computer Interaction*, 19(1), 95–112.

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

[12] Hurley, N.J. O'Mahony, M.P., and Silvestre, G.C.M. (2007), Attacking Recommender Systems: A Cost-Benefit Analysis, *IEEE Intelligent Systems*, 22(3), 64-68.

[13] Jiang, J.J., Klein, G., and Carr, C.L. (2002), Measuring Information System Service Quality: SERVQUAL from the Other Side, *MIS Quarterly*, 26(2), 145-166.

[14] Komiak, S.Y.X. and Benbasat, I. (2008), A two-process view of trust and distrust building in recommendation agents: a process-tracing study, *Journal of the Association for Information Systems*, 9(12), 727–747.

[15] Komiak, S.Y.X. and Benbasat, I. (2006), The effects of personalization and familiarity on trust and adoption of recommendation agents, *MIS Quarterly*, 30(4), 941–960.

[16] Kwon, O and Kim, J. (2007), Applying Ubi-SERVQUAL to Assessing the Quality of Ubiquitous Service Scenarios, *KORMS*, 23(1), 1-13.

[17] Lee M.C. (2009), Factors influencing the adoption of Internet banking: An in-

tegration of TAM and TPB with perceived risk and perceived benefit, *Electronic Commerce Research and Applications*, 8(3), 130-141.

[18] 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.

[19] Linden, G., Smith, B., and York, J. (2003), Amazon.com Recommendations: Item-to-Item Collaborative Filtering, *IEEE Internet Computing*, 7(1), 76-80.

[20] Melamed, D., Shapira, B., and Elovici, Y. (2007), MarCol: A Market-Based Recommender System, *IEEE Intelligent Systems*, 22(3), 74-78.

[21] Mobasher, B., Burke, R., Bhaumik, R., and Sandvig, J.J. (2007), Attacks and Remedies in Collaborative Recommendation, *IEEE Intelligent System*, 22(3), 56-63.

[22] Mooney, R. J., & Roy, L. (2000), Content-based book recommending using learning for text categorization, Proceedings of the ACM International Conference on Digital Libraries, 195–204.

[23] O'Mahony, M., Hurley, N., Kushmerick, N., and Silvestre, G. (2004), Collaborative Recommendation: A Robustness Analysis, *ACM Transaction on Internet Technology*, 4(4), 344–377.

[24] Olmo, F.H. and Gaudioso, E. (2008), Evaluation of recommender systems: A new approach, *Expert Systems With Applications*, 35(3), 790-784.

[25] Pierrakos, D., Paliouras, G., Papatheodorou, C., and Spyropoulos, C.D. (2003), Web usage mining as a tool for personalization: a survey, *User Modeling and User-Adapted Interaction*, 13(4), 311–372.

[26] Rafter, R. and Smyth, B. (2005), Conversational collaborative recommendation an experimental analysis, *Artificial Intelligence Review*, 24(3–4), 301–318.

[27] Resnick, P. and Varian, H.R. (1997), Recommender Systems, *Communications of the ACM*, 40(3), 56-58.

[28] Ricci, F. (2002), Travel Recommender Systems, *IEEE Intelligent Systems*, 17(6), 55-57.

[29] Schafer, J.B., Konstan, J., and Riedl, J. (1999), Recommender Systems in E-Commerce, Proceedings of the First ACM Conference on Electronic Commerce, 158-166.

[30] Schein, A.I., Popescul, A., Pennock, D.M., and Ungar, L.H. (2002), Methods and metrics for cold-start recommendations, Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval. 253–260.

[31] Seddon, P.B. (1997), A respecification and extension of the DeLone and McLean model of IS success, *Information Systems Research*, 8(3), 240-254.

[32] Sheng, Z. (2007), Building Trustworthy Recommender Systems, Ph.D. Dissertation,

Dartmouth College.

[33] Tam, K.Y. and Ho, S.Y. (2006), Understanding the impact of web personalization on user information processing and decision outcome, *MIS Quarterly*, 30(4), 865–890.

[34] van Setten, M., Pokraev, S. and Koolwaaij, J. (2004), Context-Aware Recommendations in the Mobile Tourist Application COMPASS, Proceedings of the Third International Conference on Adaptive Hypermedia and Adaptive Web-based Systems, 235–244.

[35] Wang, W. and Benbasat, I. (2007), Recommendation agents for electronic commerce: effects of explanation facilities on trusting beliefs, *Journal of Management Information Systems*, 23(4), 217–246.

[36] Zanker, M., Jannach, D., Gordea, S., and Jessenitschnig, M. (2007), Comparing Recommendation Strategies in a Commercial Context, *IEEE Intelligent Systems*, 22(3), 69-73.

[37] Zanker, M. and Jessenitschnig, M. (2009), Case-studies on exploiting explicit customer requirements in recommender systems, *User Modeling and User-Adapted Interaction*, 19(1-2), 133–166.

[38] Zeithaml, V. A., Berry, L. L., and Parasuraman, A. (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*, New York: Free Press.

[39] Zolt, C. (2007), Business model design and the performance of entrepreneurial firms, *Organization Science*, 18(2), 181-334.

저 자 소 개

권 오 병 (Kwon, Ohbyung)

1988년 서울대학교에서 경영학사를 19990년과 1995년에 한국과학기술원에서 각각 공학석사와 공학박사 학위를 취득하였다. 현재는 경희대학교 국제경영대학 교수로 재직 중이다. Decision Support Systems 등 다수의 저널에 수십 편의 논문을 발표하였다. 관심분야는 유비쿼터스 컴퓨팅, 의사결정지원시스템 등이다