

# Application of a Two-Level Self Organizing Map for Korean Online Game Market Segmentation

Sang-Chul Lee<sup>1</sup>, Jae-Young Moon<sup>2</sup>, Jae-Kyeong Kim<sup>2</sup>, and Yung-Ho Suh<sup>2</sup>

<sup>1</sup> Department of Management Information Systems, Korea Christian University,  
San 204, Hwagok 6-Dong, Kangseo-Ku, Seoul 157- 722, South Korea  
leecho@kcu.ac.kr

<sup>2</sup> School of Business Administration, Kyung Hee University,  
1 Hoegi-Dong, Dongdaemoon-Gu, Seoul 130-701, South Korea  
{moonlight, jaek, suhy}@khu.ac.kr

**Abstract.** Online game industry has encountered higher competition in global market. To survive successfully in today's competitive online game markets, they need to determine who the target customers are and what motivates them. The purpose of our research is to identify the critical variables and to implement a new methodology for online game market segmentation using self organizing map. Our research tested the model with Korean online game users.

## 1 Introduction

A new revolutionary period of e-commerce has begun since the early 2000's [21, 22]. Recently, the global online game industry has been grown rapidly and has been developed into the core of the world cultural industries. With the rapid growth, many online game companies hoped that the first mover would be successful and recklessly entered into online game markets without understanding the core needs of those audiences. However, the lack of consideration has forced many online game companies to fail to survive in game market [16]. To survive in today's competitive markets, online game companies need to determine who the target customers are and what motivates them. This process is called market segmentation, by which companies are able to understand their loyal customers and concentrate their limited resources into them [20].

However, previous research had problems of both methodologies and variables. First, the traditional methodology for market segmentation was based mainly on statistical clustering techniques; hierarchical and partitive approaches. However, hierarchical method can not provide a unique clustering because a partitioning to cut the dendrogram at certain level is not precise. This method ignores the fact that the within-cluster distance may be different for different clusters [6, 29]. Partitive method predefines the number of clusters, before performing it. It can be part of the error function and can not identify the precise number of clusters [7, 25, 29]. Additionally, these algorithms are known to be sensitive to noise and outliers [4, 5, 29].

To settle these problems, we segment Korean online game market using a two-level Self-Organizing Map (SOM): SOM training and clustering [29]. Instead of clustering the data directly, a large set of prototypes is formed using the SOM. The prototypes can be interpreted as proto-cluster, which are combined in the next phase from the actual clusters. The benefit of using this method is to effectively reduce the

complexity of the reconstruction task and to reduce the noise. Our research implements this method into marketing research field.

Secondly, variables of previous studies could not be accepted since they have been conducted mainly from the technological and psychological perspective. The main concern of technological research on online game is to design and develop a more attractive and effective online game environment [1, 26, 31]. However, no matter how sophisticated the technologies applied, users would not revisit the game site if it failed to reflect their needs. Our research identifies the primary factors for online game from a business perspective.

The purpose of our research is to identify the critical variables and to implement a new methodology for online game market segmentation. Our research approach is categorized into two phases. The first phase is using a statistical approach (Structural Equation Model: SEM) to find the critical segmentation factors. The second phase is conducted by a two-level Self Organizing Map (SOM) to indicate the actual clusters. To implement our methodology, Korean online game data was analyzed because Korean online game market was located in the center of those trends. Therefore, research about Korean online game markets will be helpful for other countries to understand the change of global game markets.

## 2 Theoretical Background

### 2.1 Determinant Variables

In online game, more emphasis is being placed on the impact of a flow using a business perspective [2, 9, 10, 15, 18, 23, 28]. Through the review of the relevant literature, we identify the primary factors for online game from a business perspective as follows: the convenience of operator, the suitability of feedback, the reality of design, the precision of information and the involvement of virtual community. Our research hypothesizes that these determinants have a positive effect on flow.

The convenience of the operator was defined as the manipulatability of operators to play games [27]. Operator is an important determinant of influencing interaction between users and games [2, 12, 30]. Feedback is the reaction from online games [3, 10]. For example, when players kill a monster within NCsoft's Lineage, they receive feedback upgrading their level. The reality of design is defined as the design of interface making gamers feel online games as part of the real world [1, 26, 32]. Information is the contents from online game to achieve the stated goals. Gamers who received more precise information about how to play the games tended to achieve online game goals and experience flow easier [10, 24]. Virtual community is defined as computer-mediated spaces with potential for integration of member-generated content and communication [14]. Online game users should solve problems together interacting with other users in virtual communities [10].

### 2.2 A Two-Level SOM

Vesanto and Alhoniemi proposed a two-level SOM: SOM training and clustering. A two-level SOM was combined SOM, K-means and DB Index [29]. In the first level (SOM training), the data were clustered directly in original SOM to form a large set

of prototypes. SOM was developed by Kohonen, which was very suitable for clustering in that it implemented an ordered dimensionality-reducing mapping of the training data and has prominent visualization properties [19].

In the second level (SOM clustering), the prototypes of SOM are clustered using k-means and the validity of clusters is evaluated using DB index. K-means clustering is to partition a data set into a set of group, minimizing distances within and maximizing distances between clusters. To select the best one among different partitioning, each of these can be evaluate using some kind of validity index. Generally, there are four validity indices; DB (Davies-Bouldin) index [11], Dunn’s index [13], CH (Calinski-Harabasz) index [8] and index  $I$  [25]. DB index is suitable for evaluation of k-means partitioning because it gives low values, indicating good clustering results for spherical cluster [29]. Our research used DB index.

DB index is a function of the ratio of the sum of within-cluster scatter to between-cluster separation. The DB index is defined as equation (1).

$$DB(U) = \frac{1}{c} \sum_{i=1}^c \max_{i \neq j} \left\{ \frac{\Delta(X_i) + \Delta(X_j)}{\delta(X_i, X_j)} \right\} \tag{1}$$

The scatter within the  $i$ th cluster,  $\Delta X_i$ , is computed as equation (2), the distance between cluster  $X_i$  and  $X_j$ , is denoted as equation (3). Here,  $Z_i$  represents the  $i$ th cluster centre.

$$\Delta X_i = \frac{1}{|C_i|} \sum_{x \in C_i} \{ \|x - z_i\| \} \tag{2}$$

$$\delta(X_i, X_j) = \|z_i - z_j\| \tag{3}$$

Conclusively, the proper clustering is achieved by minimizing the DB index.

### 3 Research Methods

#### 3.1 Research Framework

To segment the online game market and develop marketing strategies, our research approach is categorized into two phases. Firstly, the confirmatory factor analysis (CFA) and structural equation model (SEM) are used to identify the critical segmentation variables for clustering. Secondly, a two-level SOM is used to segment online game market. The first level develops the prototypes from large data set and the actual clusters are developed from the prototypes in the second level.

After segmentation of the markets, we use ANOVA to recognize the characteristics of sub-divided clusters. Finally, we target a segment market with the highest customer loyalty, and used those results as the starting point for the marketing strategies.

#### 3.2 Data and Measurement

To test the model, a Web-based survey was employed. We developed the web-questionnaire page using a common gateway interface (CGI). We sent a mail to

customer within OZ intermedia in Korea, which explained the objectives of the research and contained the link to the Web-Survey. Conclusively, the 1704 complete data is available for analysis, after elimination of missing data.

Our research developed multi-item measures for each construct. Twenty-one items for five determinants are selected. We asked respondents to indicate on a five point Likert scale to what extent the determinants influence on flow in online game. We used CFA to evaluate convergent validity for five constructs and 15 items remained within our model. All the fit statistics of the measurement model were acceptable.

## 4 Results

### 4.1 Identification of Critical Factors

To find the critical factors for segmentation, we used AMOS 4.0 in structural equation modeling (SEM). The structural model was well converged. The results indicated that the chi-square of the model was 295.82 with d.f. of 104, the ratio of chi-square to d.f. was 2.844, GFI was 0.982, AGFI was 0.973, RMSR was 0.031 and NFI was 0.975; all the fit statistics were acceptable. Additionally, the squared multiple correlations ( $R^2$ ) indicated that the present model explains 51 % of the variance in flow.

**Table 1.** The results of Stuctural Equeation Model

	Path		Estimate	S.E.	t	p
O	-->		0.037	0.024	1.456	0.145
FB	-->		0.108**	0.032	3.482	0.000
IF	-->	F	0.081*	0.035	2.499	0.012
D	-->		0.274**	0.036	8.798	0.000
C	-->		0.437**	0.031	14.794	0.000

\* p<0.05, \*\* p<0.01

O: The convenience of operator,

FB: The suitability of feedback

IF: The precision of information,

D: Reality of Design

C: The involvement of virtual community,

F: Flow,

Four of the five paths were statistically significant and the path from the convenience of operator to flow was insignificant, as shown in Table 1. The critical variables for marketing segmentation are the suitability of feedback, the reality of design, the precision of information and the involvement of virtual community.

### 4.2 Market Segmentation

To segment the Korean online game market, our research was conducted using a two-level SOM. In the experiments, the first level was SOM training. 1704 data samples of the Korean were collected using the test variables: the suitability of feedback, the precision of information, the reality of design and virtual community except the convenience of operator. A SOM was trained using the sequential training algorithm for

Korean data samples. A neighborhood width decreased linearly 5 to 1 using the Gaussian function. A map was used by 19\*11 matrix and 209 prototypes were developed.

The second level was SOM clustering. The partitive clustering of 209 SOM's prototypes was carried out using batch K-means algorithm. The K-means ran multiple times for each k. The DB index was used to select the best clustering in Fig. 2. The analysis of the DB index resulted in the development of ten market segments.

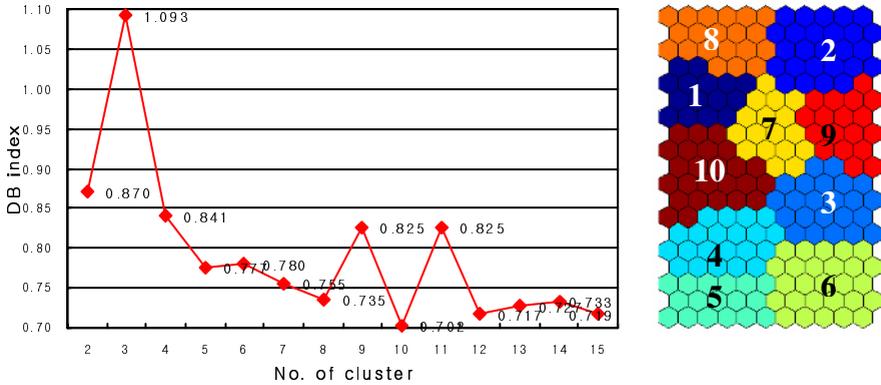


Fig. 2. DB Index and visualization of clusters

### 4.3 Determination of Target Market

After segmenting the markets, we used ANOVA to recognize the variable characteristics of each cluster. According to results of ANOVA, all variables (components) were significant;  $F=154.34$  to  $588.19$  and  $p=0.00$ . To precisely recognize the variable characteristics of clusters, we categorized the effectiveness of the variables into 3 levels; high, middle and low. The middle level ranged between  $3 \pm 2.5$  because our research measurement was used on a five point Likert scale. The high score suggested that the cluster was influenced by the variables positively, the middle score was normal, the low score was negative.

Additionally, to identify the structure of the clusters, we conducted on the analysis of the demographic and behavioral variables: gender, age, income level,  $i\_year$  (how long did gamers use the Internet),  $i\_day$  (how many hours did gamer use the Internet per day), and  $g\_day$  (how many hours did gamer play online games per day). The characteristics and structure of clusters are summarized in Table 2.

The analysis of customer loyalty indicated that the ranks of clusters are as follows: cluster 6 (4.02 for average) > cluster 5 (3.94) > cluster 3 (3.76) > cluster 4 (3.69) > cluster 9 (3.54) > cluster 7 (3.48) > cluster 10 (3.28) > cluster 1 (3.24) > cluster 2 (3.07) > cluster 8 (2.86) in Table 2. The other analysis of the intention of revisit and WOM (Word of Mouth) indicated the same results. As a result, cluster 6 was indicated as the primary target market.

**Table 2.** Profiles of clusters

	C 1 (n=116)	C 2 (n=204)	C 3 (n=130)	C 4 (n=130)	C 5 (n=176)	C 6 (n=224)	C 7 (n=103)	C 8 (n=196)	C 9 (n=161)	C 10 (n=264)
FB	2.72 (L*)	1.52 (L)	2.26 (L)	3.10 (M)	3.72 (H)	2.72 (L)	2.44 (L)	1.91 (L)	1.80 (L)	3.03 (M)
IF	2.69 (L)	2.50 (L)	3.17 (M)	3.22 (M)	3.76 (H)	3.63 (H)	2.72 (L)	2.18 (L)	3.01 (M)	2.92 (M)
D	2.77 (M)	2.47 (L)	3.56 (H)	3.25 (M)	3.60 (H)	3.80 (H)	2.97 (M)	2.55 (L)	3.22 (M)	3.08 (M)
C	2.68 (L)	3.45 (H)	3.59 (H)	3.49 (H)	3.67 (H)	4.00 (H)	3.35 (H)	2.38 (L)	3.56 (H)	3.05 (M)
Gender	male	male	female	female	male	female	both	male	female	female
Age	26-30	21-25	21-25	26-30	21-25	26-30	21-30	26-30	26-30	26-30
Income (\$)	1,001- 2,000	- 500	1,001- 2,000	1,001- 2,000	501- 1,000	501- 1,000	- 500	- 500	1,001- 2,000	501- 1,000
i_year	3	3	2-4	3	3	2-4	3	4	2-5	2-3
i_day	2	5	3-5	3-5	5	5, 10	5	3	5	3
G_day	1	2	3	1-2	1	2	1	1	1	1
Revisit	3.29	3.15	3.77	3.72	3.94	4.02	3.51	2.96	3.61	3.33
WOM	3.18	3.00	3.75	3.67	3.93	4.02	3.45	2.75	3.47	3.22
Loyalty**	3.24	3.07	3.76	3.69	3.94	4.02	3.48	2.86	3.54	3.28
Rank	8	9	3	4	2	1	6	10	5	7

\* L=Low, M=Middle, H=High

\*\* Loyalty is estimated by average of revisit and WOM

O: The convenience of operator,

FB: The suitability of feedback

IF: The precision of information,

D: Reality of Design

C: The involvement of virtual community,

F: Flow,

## 5 Implication

The results of our research have the following implications for Korean online game companies. To attract the primary target audiences, companies should develop strategies depending on the effectiveness of the variables and the demographic and behavioral characteristics of cluster 6.

The characteristics of target audiences indicate that the members are positively influenced by the suitability of feedback, the reality of design and the involvement of virtual community. The strategies of the reality of the suitability of feedback proposed that companies should provide gamers with a higher level faster, items and more cybermoney, when gamers completed their mission. The strategies of the reality of design proposed that companies should make an interface where the game site looks real. For example, the interface of recent games changed 2D such 'Lineage' into 3D such as 'MU', 'Lagnarok' and 'Laghaim'. For virtual community, companies need to provide a Role Playing Game (RPG) where the gamer cooperates with each other rather than shooting games where the gamer compete with each other. Furthermore, the different villages and guilds which were harmonized with customer needs were provided. For example, 'Lineage' provided 15 villages to satisfy the different gamers' needs.

As to the demographic information, there are more female members in the target group than male members are. Popular ages range from 26 to 30 in the group. Monthly income level of the members of the group ranges from \$500 to \$ 1,000. They

have used the Internet for over 2-4 years, use the Internet for 5 or 10 hours per day, and play online games for more than 2 hours per day.

The result indicated that online game companies should develop diverse types of online games considering the extension of the age of online game users. The number of female users is growing fast and the needs of online game users become diversified [21, 33]. To better satisfy their needs, online game companies should cluster similar customers into specific market segments with different demands and then develop marketing strategy based on their properties. Especially, our research shows that the middle-aged and female users are classified as target customers as well as adolescents. This finding is consistent with the statistics in the Korean Game White Paper, which indicates that female users increased from 31% of the game population in 2001 to 47% in 2003 and the middle-aged users increased from 2% in 2001 to 21% in 2003.

These implications were proven to be true through NCsoft's example, which is the primary Korean online game company. They recognized that online game customers' needs have been changed and encountered higher competition with foreign online game competitors. To survive in this changing environment, they developed the games for male and female separately. For instances, the background of the recent game 'Lineage' was medieval, the type was combatable, and their target audiences were adolescents and younger male, while 'Shining Lore' is developed to target female customers who might be more interested in sweet and exciting stories [16].

## 6 Conclusion and Limitation

The results of our study have several contributions to academia and business world. Our research identifies the new primary factors for online game markets which may not be found in the previous researches from the technological perspectives. Additionally, our research proposes a new methodology for market segmentation using a two-level SOM and marketing strategies for the survival in competitive online game market.

Even though our research is conducted on Korean online game market, these implications are able to be applied into those of other countries because Korean online game market is the frontiers of global online game market. And the research about Korean online game market is thought to be helpful for other countries to understand the change of their own online game markets. However, other countries are able to develop their own marketing strategies more exactly using our methods with considering and adjusting their market environment, instead of accepting the results of our research.

For further study, more demographic and behavioral variables might be necessary to segment the markets more precisely. Secondly, a cross-national analysis can be added to our research in order to better understand the loyal customers in different countries.

## References

1. Ackley, J.: Roundtable Reports: Better Sound Design. Gamasutra. (1998). [http://www.Gamasutra.com/features/gdc\\_reports/cgdc\\_98/ackley.htm](http://www.Gamasutra.com/features/gdc_reports/cgdc_98/ackley.htm)
2. Agarwal, R., Karahanna, E.: Time Files When You're Having Fun: Cognitive Absorption and Beliefs About Information Technology Usage. *MIS Quarterly*, Vol.24, No.4. (2000) 665-694

3. Baron, J.: Glory and Shame: Powerful Psychology in Multiplayer Online Games. Gamasutra. (1999). [http://www.gamasutra.com/features/19991110/Baron\\_01.htm](http://www.gamasutra.com/features/19991110/Baron_01.htm)
4. Bezdek, J. C.: Some New Indexes of Cluster Validity. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, Vol.28, No.3. (1998) 301-315
5. Blatt, M., Wiseman, S., Domany, E.: Super-paramagnetic Clustering of Data. *Physical Review Letters*, Vol.76, No.8. (1996) 3251-3254
6. Boudaillier, E., Hebrail, G.: Interactive Interpretation of Hierarchical Clustering. *Intelligent Data Analysis*, Vol.2, No.3. (1998) 41-
7. Buhmann, J., Kühnel, H.: Complexity Optimized Data Clustering by Competitive Neural Networks. *Neural Computation*, Vol.5, No.3. (1993) 75-88
8. Calinski, R. B., Harabasz, J.: A Dendrite Method for Cluster Analysis. *Communication in Statistics*, Vol.3. (1974) 1-27
9. Cho, N. J., Back, S. I., Ryu, K. M.: An Exploratory Investigation of Player Loyalty to Online Games. *Journal of the Korean Operations Research and Management Science Society*, Vol.26, No.2. (2001) 85-97
10. Choi, D. S., Park, S. J., Kim, J. W.: A Structured Analysis Model of Customer Loyalty in Online Games. *Journal of MIS Research*, Vol.11, No.3. (2001) 1-20
11. Davies, D. L., Bouldin, D. W.: A Cluster Separation Measure. *IEEE Transactions Pattern Analysis and Machine Intelligence*, Vol.1. (1979) 224-227
12. Davis, F. D., Bagozzi, R. P., Warshaw, P. R.: Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. *Journal of Application Society Psychology*, Vol.22, No.14. (1992) 1111-1132
13. Dunn, J. C.: A Fuzz Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, Vol. 3. (1973).32-57
14. Hagel, J., Armstrong, A.: *Net Gain: Expanding Markets through Virtual Communities*. Harvard Business School Press, Boston. (1997)
15. Hoffman, D. L., Novak, T. P.: Marketing in Hypermedia Computer Mediated Environments: Conceptual Foundations. *Journal of Marketing*, Vol.60, No.3. (1996) 50-68
16. ICA: 2003 Cases of Information and Telecommunication Export, ICA, Seoul. (2003)
17. KGDI: 2003 Korean White Game Paper. KGDI, Seoul. (2003)
18. Kim, N.H., Lee, S.C., Suh, Y.H.: Strategy of Market Penetration in Japanese Internet Market: Comparing Online Game Loyalty between Korea and Japan with MSEM. *Journal of the Korea Society for Quality Management*, Vol.31, No.1. (2003) 21-41
19. Kohonen, T.: Self-organizing Map. *Proceedings of the IEEE*, Vol.78, No.9. (1990) 1469-1480
20. Kotler, P.: *Marketing Management: Analysis, Planning, Implementation and Control*, 9<sup>th</sup> edn. A Simon and Schuster Co, New Jersey. (1997)
21. Laudon, K. C., Traver, C. G.: *E-Commerce: Business, Technology, Society*. Addison Wesley, Boston. (2002)
22. Lee, H. K.: Overview and Problems within the Korean Online Game Industry. *Information and Telecommunication Policy*, Vol.13. (2000) 20-37
23. Lee, S. C., Kim, N. H., Suh, Y. H.: The Effect of Flow and Addiction upon Satisfaction and Customer Loyalty in Online Games. *Korean Management Review*, Vol.32, No.3. (2003) 1479-1501
24. Lewinski, J. S.: *Developer's Guide to Computer Game Design*. Wordware Publishing Inc, Texas. (2000)
25. Maulik, U., Bandyopadhyay, S.: Performance Evaluation of Some Clustering Algorithms and Validity Indices. *IEEE Transactions of Pattern Analysis and Machine Intelligence*, Vol.24, No.12. (2002) 1650-1654

26. Sanchez-Crespo, D.: 99 from a Game Development Perspective. Gamasutra. (1999). [http://gamasutra.com/features/19990802/siggraph\\_01.html](http://gamasutra.com/features/19990802/siggraph_01.html)
27. Spector, W.: Remodeling RPGs for the New Millennium. Gamasutra. (1999). [http://www.gamasutra.com/features/game\\_desing/19990115/remodeling\\_01.htm](http://www.gamasutra.com/features/game_desing/19990115/remodeling_01.htm)
28. Sujan, H., Weitz, B. A., Kumar, N.: Learning, Orientation, Working Smart, and Effective Selling. *Journal of Marketing*, Vol.58, No.3. (1994) 39-52
29. Vesanto, J., Alhoniemi, E.: Clustering of the Self-organizing Map. *IEEE Transactions on Neural Networks*, Vol.11, No.3. (2000) 586-600
30. Webster, J., Martocchio, J. J.: Micro Computer Playfulness: Development of a Measure with Workplace Implications. *MIS Quarterly*, Vol.16, No.2. (1992) 201- 226
31. Wells, J. D., Fuerst, W. L., Choobinceh, J.: Managing Information Technology for One to One Customer Interaction. *Information and Management*, Vol.35, No.1. (1999) 53-62
32. Woodcock, W.: Game AI: the State of the Industry. Gamasutra. (1999). [http://www.gamasutra.com/features/19990820/game\\_ai\\_01.html](http://www.gamasutra.com/features/19990820/game_ai_01.html)
33. Yu, S. S.: Overview of Information and Telecommunication Industries: Software and Internet Contents. *Information and Telecommunication Policy* (2002) 131-144