

Business Model Characterization by analyzing Business Model Components of Patent Data

Seok-Hun Kim¹, Kwang Hyuk Im^{2*}

¹Division of Computer Engineering, Mokwon University, Korea (First Author)

²Department of Electronic Commerce, Pai Chai University, Korea (Corresponding Author)

Email Address: khim@pcu.ac.kr

Received: Received May 02, 2011; Revised July 25, 2011; Accepted September 12, 2011

Published online: 1 January 2012

Abstract: The Business model describes the rationale of how an organization creates, delivers, and captures value - economic, social, or other forms of value. Although the concept of business model exists in the former, it coincided with the rapid spread of the Internet in the mid-1990s. Because concept of business model is a very abstract, there is confusion about business model definition and business model taxonomy. But, most researchers agree to the fact that business model is composed in several components which interacts with each other, and as a result, decides the characteristics of its business model. In this paper, we suggest a method to analyze the characteristics of the company by analyzing its patent based on business model components. By showing the results of analyzing Hyundai Motors, STX Shipbuilding, and Naver's patent datum with business component model, we could effectively see some characteristics about business model which allows us to verify appropriate management strategies.

Keywords: Business Model, Business Model Components, Business Model Characteristics, Patent data analysis

1 Introduction

Business model shows how a business buys and sells the product and service, and how it gains profit. However, business model can be too abstract because there are various opinions on the concept of the model, and therefore, can cause confusion. Even in existing paper, the starting point to describe the concept of a business model to the process of a business modeling is approached differently, and therefore a continuous discussion about the difference of a strategy and business model is occurred [1].

Business model has the business characteristics of a company, and to analyze this kind of business characteristics, it is needed to extract and analyze

components which a business model has. Shafer et al. analyzed the 12 representative paper about the definition of business model from the paper presented in 1998 to 2003. He classified the 42 business model components which were referred as the components of a business model, and organized the elements that were referred more than two times and created a clustering into the following 4 categories- strategic Choices, Creating Value, Capturing value, and the value network[2]. Osterwalder et al. found the 4 major areas composed of the required business model issue of a company, and reformed it into 9 business model components in detail[3].

* Corresponding Author: Kwang Hyuk Im, khim@pcu.ac.kr

Context	Timmers (1998) E-Business	Hamel (2000) Strategy	Afuah and Tucci (2001) E-Business	Amit and Zott (2001) E-Business	Well and Vitale (2001) E-Business	Dubosson-Torbay et al. (2002) E-Business	Magretta (2002) Strategy	Rayport and Jaworski (2002) E-Business	Van Der Vorst et al., 2002 E-business/SCM ^a	Hoque (2002) Technology	Chesbrough (2003) Strategy	Hedman and Kalling (2003) IS ^b and strategy
Components												
Value network (suppliers)	X	X			X	X			X		X	X
Customer (target market, scope)	X	X	X			X		X		X	X	
Resources/assets		X		X		X		X				X
Value proposition			X			X	X	X		X	X	
Capabilities/competencies		X	X	X		X						X
Processes/activities		X	X			X			X			X
Revenue/pricing	X	X	X			X					X	
Competitors							X			X		X
Cost						X		X			X	
Information flows	X			X	X							
Output (offering)				X				X				X
Product/service flows	X			X	X							
Strategy		X									X	
Branding						X				X		
Customer information		X				X						
Customer relationship		X				X						
Differentiation		X										
Financial aspects						X		X				
Mission	X										X	
Profit												
Business opportunities				X			X					
Cash flows					X							
Create value				X								
Culture										X		
Customer benefits								X				
Customer interface		X										
Economic logic							X					
Environment											X	
Firm identity											X	
Firm reputation											X	
Fulfillment and support		X										
Functionalities									X			
Implementation			X									
Infrastructure-applications									X			
Infrastructure-management						X						
Management												X
Product innovation						X						
Specific characteristics									X			
Sustainability			X									
Transaction content				X								
Transaction governance				X								
Transaction structure				X								

^a Supply chain management.
^b Information systems.

In this paper, we suggest a method to analyze a company’s characteristics by analyzing its patents based on business model components. Since the patent data of Hyundai Motors, STX Shipbuilding of the machine industry and NAVER of the e-biz industry, are based on the business model components, by showing the results of the analysis, we could effectively extract each characteristic of the industry and company about the business model. Also, based on the results of this analysis, a management strategy of an industry or a company can be established.

2 Business Model Components

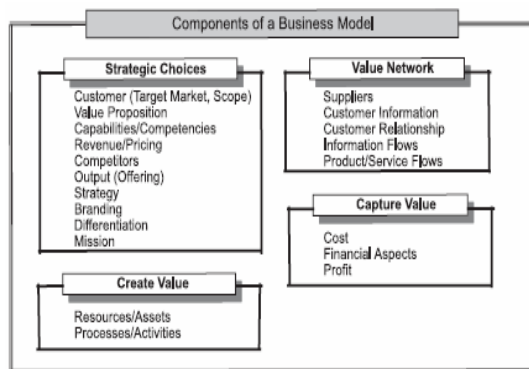


Fig.1 Components of business model affinity diagram

Shafer *et al.* analyzed the 12 representative paper about the definition of business model from the paper presented in 1998 to 2003. He classified the 42 business model components which were

referred as the components of a business model, and organized it into a table. The following [Table 1] shows the 42 organized components [2].

He classified the elements that were referred more than two times into the following 4 categories - strategic Choices, Creating Value, Capturing value, and the value network- as shown in [Figure 1]

Osterwalder *et al.* found the 4 major areas composed of the required business model issue of a company, and reformed it into 9 business model components in detail [3].

[Table 2] shows the 4 areas- Product, Customer Interface, Infrastructure Management, Financial Aspect- and 9 business components - Value proposition, Value Propositions, Customer Segments, Channels, Customer Relationships, Key Activities, Key Resources, Key Partnerships, Cost Structure, Revenue Stream- that he proposed.

Shafer *et al.* proposed 4 areas, 20 components, and Osterwalder *et al.* proposed 4 areas, 9 components. Shafer *et al.* could not propose the exact definition of the extracted 20 elements, and reached his limit because he could not clearly assert the classification standard of the 4 categories. On the other hand, Osterwalder *et al.* proposed the exact definition of the 9 components. So, therefore, in this paper, the 9 components proposed by Osterwalder *et al.* were used for the analysis.

Shafer *et al.* proposed 4 areas, 20 components, and Osterwalder *et al.* proposed 4 areas, 9 components. Shafer *et al.* could not propose the exact definition of the extracted 20 elements, and

reached his limit because he could not clearly assert the classification standard of the 4 categories. On the other hand, Osterwalder *et al.* proposed the exact definition of the 9 components. So, therefore, in this paper, the 9 components proposed by Osterwalder *et al.* were used for the analysis.

4 Area(s)	BM Component(s)	Definition
Product	Value Propositions	Consideration in giving valuable product and service to the customer
	Customer Segments	Customer(s), the company wants to provide value
Customer Interface	Channels	Ways to contact with the customer
	Customer Relationships	Bond made by the company for the relationship between the customer and the company
	Key Activities	Composition of necessary resource and activity to create value for the customer
Infrastructure Management	Key Resources	The most important assets for the business model to work smoothly and to create value for the customer
	Key Partnerships	Cooperative relation between two or more companies to create value for the customers
	Cost Structure	All cost to manage the business model
Financial Aspects	Revenue Stream	Ways to earn profit for the company through various profit model

Table.2 The Business Model Ontology

3 Clustering and Classification Methodologies

3.1 Clustering - K-means Algorithm

The most common algorithm uses an iterative refinement technique. Due to its ubiquity it is often called the *k-means algorithm*.

Given an initial set of *k* means $\mathbf{m}_1^{(1)}, \dots, \mathbf{m}_k^{(1)}$ (see below), the algorithm proceeds by alternating between two steps:

Assignment step: Assign each observation to the cluster with the closest mean.

$$S_i^{(t)} = \left\{ \mathbf{x}_j : \|\mathbf{x}_j - \mathbf{m}_i^{(t)}\| \leq \|\mathbf{x}_j - \mathbf{m}_{i^*}^{(t)}\| \text{ for all } i^* = 1, \dots, k \right\} \tag{1}$$

Update step: Calculate the new means to be the centroid of the observations in the cluster.

$$\mathbf{m}_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{\mathbf{x}_j \in S_i^{(t)}} \mathbf{x}_j \tag{2}$$

The algorithm is deemed to have converged when the assignments no longer change [4].

Commonly used initialization methods are Forgy and Random Partition. The Forgy method randomly chooses *k* observations from the data set and uses these as the initial means. The Random Partition method first randomly assigns a cluster to each observation and then proceeds to the Update step, thus computing the initial means to be the centroid of the cluster's randomly assigned points. The Forgy method tends to spread the initial means out, while Random Partition places all of them close to the center of the data set. According to Hamerly *et al.*, the Random Partition method is generally preferable [5].

As it is a heuristic algorithm, there is no guarantee that it will converge to the global optimum, and the result may depend on the initial clusters. As the algorithm is usually very fast, it is common to run it multiple times with different starting conditions. However, in the worst case, *k-means* can be very slow to converge: in particular it has been shown that there exist certain point sets, even in 2 dimensions, on which *k-means* takes exponential time, that is $2\Omega(n)$, to converge[6-7]. These point sets do not seem to arise in practice: this is corroborated by the fact that the smoothed running time of *k-means* is polynomial [8-9].

3.2 Classification - C4.5 Algorithm

C4.5 is an algorithm used to generate a decision tree developed by Ross Quinlan [10]. C4.5 is an extension of Quinlan's earlier ID3 algorithm. The decision trees generated by C4.5 can be used for classification, and for this reason, C4.5 is often referred to as a statistical classifier.

C4.5 builds decision trees from a set of training data in the same way as ID3, using the concept of information entropy. The training data is a set $S = s_1, s_2, \dots$ of already classified samples. Each sample $s_i = x_1, x_2, \dots$ is a vector where x_1, x_2, \dots represent attributes or features of the sample. The training data is augmented with a vector $C = c_1, c_2, \dots$ where c_1, c_2, \dots represent the class to which each sample belongs[11-13].

At each node of the tree, C4.5 chooses one attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. Its criterion is the normalized information gain (difference in entropy) that results from choosing an attribute for splitting the data. The attribute with the highest normalized information

gain is chosen to make the decision. The C4.5 algorithm then recurs on the smaller sublists.

This algorithm has a few base cases.

- All the samples in the list belong to the same class. When this happens, it simply creates a leaf node for the decision tree saying to choose that class.
- None of the features provide any information gain. In this case, C4.5 creates a decision node higher up the tree using the expected value of the class.
- Instance of previously-unseen class encountered. Again, C4.5 creates a decision node higher up the tree using the expected value

4 Analysis of Business Model Characteristics

4.1 Data gathering and coding

To gather data, a company targeted for gathering data has been selected. The selection standard is based on the representation of the manufacturing and service industry. As the result, the machine industry and e-biz industry were selected for the comparison analysis. STX Shipbuilding and Hyundai Motors representing the machine industry, and NHN representing the e-biz industry were selected.

We gathered the patent application data as data for analysis, and by analyzing each company and the annual number of application patent, the annual ratio was selected. Finally, 100 patents of each company were selected. The following [Table 3] shows number of selected data of each company.

company	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Total
STX Shipbuilding	0	0	0	2	8	0	51	0	89	28	178
	(0.0%)	(0.0%)	(0.0%)	(1.12%)	(4.49%)	(0.0%)	(28.65%)	(0.0%)	(50.00%)	(15.73%)	(100%)
Hyundai Motors	4009	4171	4017	3931	3878	2716	2646	3017	3027	225	31,637
	(12.67%)	(13.18%)	(12.70%)	(12.43%)	(12.26%)	(8.58%)	(8.36%)	(9.54%)	(9.57%)	(0.71%)	(100%)
NHN (NAVER)	1	2	9	68	234	133	140	234	258	46	1,125
	(0.09%)	(0.18%)	(0.80%)	(6.04%)	(20.80%)	(11.82%)	(12.44%)	(20.80%)	(22.93%)	(4.09%)	(100%)
Total	300										

Table.3 Results of Gathering Data

Experts analyze the patent application content of the selected patent data and assigned value of 0 and 1 each to 9 business component. For the standard of coding, if there is a business component related to the application, '1' is assigned, and if there is no content for the related business component, '0' is assigned.

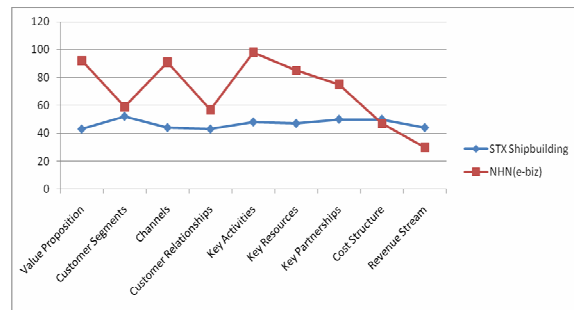
The following [Table 4] shows the analysis results of each company's 100 patent classified into 9 components. The value each component shows is the number of related patent from the 100 patents. In other words, from the 100 patents, Hyundai Motors has 86 patents describing the Value Proposition.

Company	Value Proposition	Customer Segments	Channels	Customer Relationships	Key Activities	Key Resources	Key Partnerships	Cost Structure	Revenue Stream
Hyundai Motors	86	37	61	39	98	93	66	36	33
STX Shipbuilding	43	52	44	43	48	47	50	50	44
NHN(e-biz)	92	59	91	57	98	85	75	47	30

Table.4 Component Analysis Results of Each Company

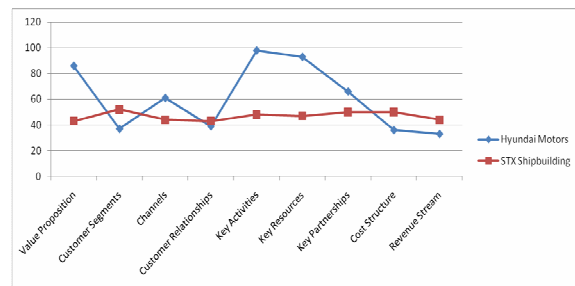
4.2 Comparison the Characteristics of Business Model

By analyzing the provided data, a comparison the characteristics of industrial business model is identified. First, to compare the characteristics of the manufacturing industry's business model and the service industry's business model, the following [Figure 2-(a)] informs the business model's



characteristics of STX Shipbuilding and NHN.

(a) Comparison of business model components of STX Shipbuilding and NHN



(b) Comparison of business model components of STX Shipbuilding and Hyundai Motor

Fig.2 Comparison of Business Model Components

Everyone can predict that manufacturing industry and service industry have differences, but it is difficult to show it in specific numbers. From [Figure 2-(a)], STX Shipbuilding has 9 components showing values of 40-50%, which is similar in importance; in contrary, NHN's business model components related to Value Proposition, Channel, Key Activities, Key Resources, Key Partnership, has greater importance. In other words, giving valuable product and service to the customer, ways to contact with the customer, composition of necessary resource and activity to create value for the customer, the most important assets for the business model to work smoothly and to create value for the customer, cooperative relation between two or more companies to create value for the customers is comparatively important in e-biz industry, and companies in the industry applies these characteristics in its business model. This kind of difference is held because of the characteristics of STX Shipbuilding which deals with order production industry in which special customers order ships instead of general customers.

So then, in comparison with the STX Shipbuilding, we want to know the business model's characteristics of manufacturing industry dealing with general customers. So, we compare STX Shipbuilding with Hyundai Motors. From [Figure 2- (b)], Hyundai Motors, compared to STX Shipbuilding, components of value proposition, channel, Key Activities, Key Resources, Key Partnerships are relatively high. This result is similar with the comparison between STX Shipbuilding and NHN. So, the following [Figure 3], shows the overall comparison of STX Shipbuilding, Hyundai Motors, and NHN.

As shown in [Figure 3], compared to STX Shipbuilding, Hyundai Motors and NHN reflect value proposition considering giving valuable product and service to the customer, various channels to contact with the customer, key activities showing the composition of necessary resource and activity to create value for the customer, and the resource of the most important assets for the business model to work smoothly and to create value for the customer as key resources in applying and managing these components. These industrial characteristics show that Hyundai Motors and NHN deal with general customers. Relatively, since STX Shipbuilding deals with special customers, they reflect less on channel, and also reflect less on value proposition because of order production. But they

show that all aspects of components should be measured for order production.

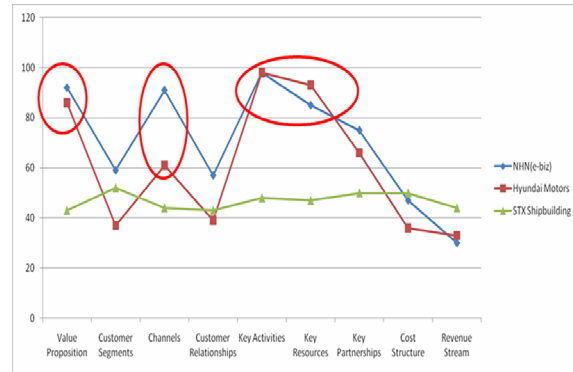


Fig.3 Comparison of Business Model Components of STX Shipbuilding, Hyundai Motors, and NHN

4.3 Comparison Methodology of the Characteristics of Business Model

When there are three targeted business from section 4.2, characteristics can be shown clearly from the graph, but if there are too many targeted business, it is hard to identify its characteristics from the graph. So, by analyzing data, it is needed the methodology for extracting and comparing business model characteristics. It uses data mining techniques - clustering and classification. With the researched patent of each company, data set were created by the value given to each business model components. To cluster the composed data set with patents of similar characteristics, clustering is conducted, and to extract the characteristics of the cluster, classification is conducted. The framework of this methodology is shown as the following [Figure 4].

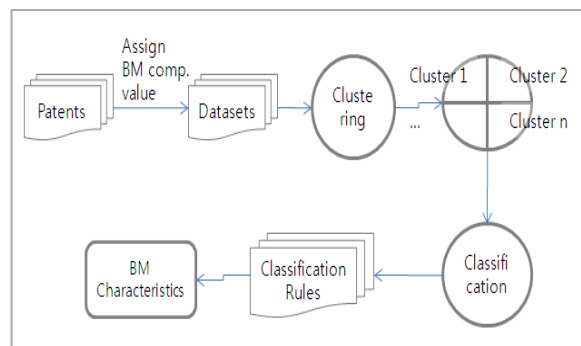


Fig.4 Total Framework of Business Model's Characteristics Analysis

The effectiveness of the methodology is proven by applying the analysis methodology shown in [Figure 4] to the classified data set in section 4.1. First, the data set classified in section 4.1 input in clustering algorithm forms the cluster. Various



algorithms exist in the clustering algorithm, but the present research used k-means algorithm, which is most frequently used. Since there are only 3 related companies, 2 clusters were conducted as the algorithm. As the result of conducting the clustering, cluster 1 and cluster 2 were made; the number of related patent for Hyundai Motors, STX Shipbuilding, and NHN is shown in the following [Table 5].

Company	Cluster 1	Cluster 2
STX Shipbuilding	73	27
Hyundai Motors	7	93
NHN	9	91
Total	89	211

Table.5 Analysis Result of Patent Data's Clustering

Classification algorithm is the algorithm used to automatically extract characteristics of cluster. Various algorithms exist in the classification algorithm like CRT, Quest, C5.0, but we used C5.0. The following shows the Classification result.

```

=====
Key Activities = 1 [ Mode: cluster-2 ]
Value Proposition = 1 [ Mode: cluster-2 ]
Key Resources = 1 [ Mode: cluster-2 ] => cluster-2
Key Resources = 0 [ Mode: cluster-2 ]
Channels = 1 [ Mode: cluster-2 ] => cluster-2
Channels = 0 [ Mode: cluster-1 ] => cluster-1
Value Proposition = 0 [ Mode: cluster-1 ]
Key Resources = 1 [ Mode: cluster-2 ]
Channels = 1 [ Mode: cluster-2 ] => cluster-2
Channels = 0 [ Mode: cluster-1 ] => cluster-1
Key Resources = 0 [ Mode: cluster-1 ] => cluster-1
Key Activities = 0 [ Mode: cluster-1 ]
Channels = 1 [ Mode: cluster-1 ]
Value Proposition = 1 [ Mode: cluster-2 ]
Key Resources = 1 [ Mode: cluster-2 ] => cluster-2
Key Resources = 0 [ Mode: cluster-1 ] => cluster-1
Value Proposition = 0 [ Mode: cluster-1 ] => cluster-1
Channels = 0 [ Mode: cluster-1 ] => cluster-1
=====
    
```

The characteristics of the cluster can be known by applying the classification rules as the above. To intuitively recognize these rules, the following decision tree form [Figure 5] is shown.

If the patent of a specific company or industry can be analyzed by using the framework proposed in this paper, a systematic analysis of the business model for the targeted company or industry is verified by applying the actual data.

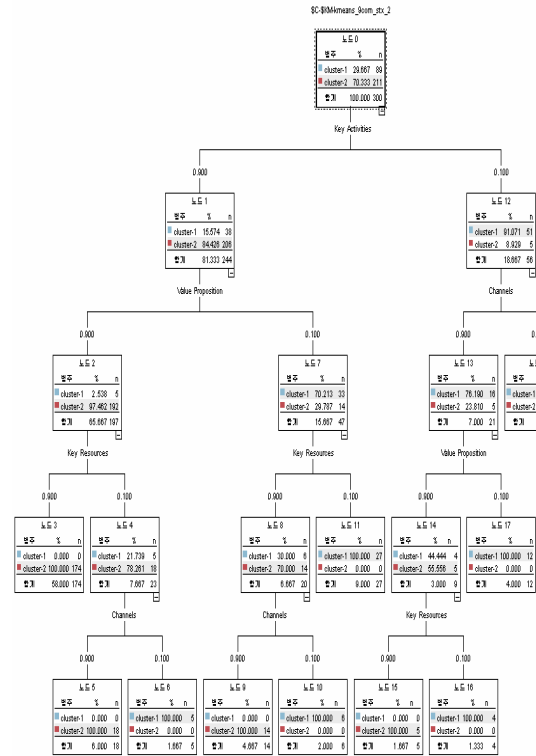


Fig.5 C5.0 Classification Tree

4.4 Strategy Establishment

If the company gets the results of the analysis by the suggested framework, with this source, it can establish a strategy to strengthen its capability. The first strategy that can be established is the strategy of maximizing its strongpoint. From the analysis results, by extracting the most confident component, and further developing that component, a strategy can be established which can increase the company's competitiveness. The second strategy that can be established is the strategy of compensating its weakness. The analysis result includes relatively weak components of itself. Therefore, instead of competitive companies, a strategy can be established by strengthening your weak components and your capability.

5 Conclusion

This paper proposed analysis framework based on business model components for the business model analysis of specially designated companies or industry group. It used patent data for analysis and experts conducted the mapping of patent data into the business model components. Mapping data for business model components by patent is used for input data for clustering. Every cluster classified as the result of clustering is formed by patents whose characteristics are similar. In order to analyze characteristics of each cluster, classification algorithm is necessary. And rules extracted by classification algorithm show characteristics of each cluster in detail. Proposed framework is practically applied for the data of STX Shipbuilding, Hyundai Motors, and NHN, extracted characteristics of each company's business model and verified performance of proposed framework. Regarding components, STX Shipbuilding has almost similar degree of importance, but Hyundai Motors and NHN show that such components as value proposition, channel, key activities and key resources are considered relatively valuable.

Existing paper emphasized definition of business model and provision of components, but this paper is meaningful to propose practical framework by using proposed business model components, analyzing actual data, extracting results and establishing strategy.

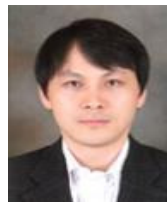
However, it needs to be further studied whether this research proposes efficient results regarding the huge data since this research proposed analysis result based on limited data. Also additional research needs to gather and analyze for further objective and various data since only the patent data has been used for the input, and there are limits on the patent data.

References

- [1] Stähler, P., Business Models as an Unit of Analysis for Strategizing, *International Workshop on Business Models, Lausanne, Switzerland*, (2002)
- [2] S. M. Shafer, H. Jeff Smith, J. C. Linder, The Power of business models, *Business Horizons*, 48, (2005) 199-207.
- [3] A. Osterwalder. and Y. Pigneur, Business Model Generation, *John Wiley & Sons, Inc.*, (2010.)
- [4] MacKay, David, Chapter 20. An Example Inference Task: Clustering, *Information Theory, Inference and Learning Algorithms. Cambridge University Press.* (2003) 284-292.
- [5] Hamerly, G. and Elkan, C., Alternatives to the k-means algorithm that find better clusterings, *Proceedings of the eleventh international conference on Information and knowledge management*, (2002).
- [6] Vattani, A., k-means requires exponentially many iterations even in the plane, *Proceedings of the 25th Symposium on Computational Geometry*, (2009).
- [7] Arthur, D., Vassilvitskii, S., How Slow is the k-means Method?, *Proceedings of the 22nd Symposium on Computational Geometry*, (2006).
- [8] Arthur, D., Manthey, B., Roeglin, H., k-means has polynomial smoothed complexity, *Proceedings of the 50th Symposium on Foundations of Computer Science*, (2009).
- [9] Hyun-Kyung Kang, Hierarchical Clustering based Personalized Feedback System for Mass Health Examination, *Journal of The Korea Knowledge Information Technology Society*, 6, 4(2011)103-112.
- [10] Quinlan, J. R., "C4.5: Programs for Machine Learning", *Morgan Kaufmann Publishers*, (1993).
- [11] Han-Kuk Hong, Feature and Countermeasure on Enterprise Crisis through Case Analysis, *Journal of The Korea Knowledge Information Technology Society*, 6, 4(2011)83-101.
- [12] Changsok Yoo, H.K. Kim, E.Y Heo, The Economic Value of Online Game Developers in Early Stages, *Journal of Future Game Technology*, 1, 1(2011) 21-34.
- [13] Kwang Hyuk Im, Seok-Hun Kim, Common cheating forms and Detecting game bots, *Journal of Future Game Technology*, 1, 1(2011) 43-48.



Seok-Hun Kim is an adjunct professor in the Computer Engineering at Mokwon University. He received the M.S., Ph.D. degree in Computer Engineering from Hannam University.



Kwang Hyuk Im is an assistant professor of Electronic Commerce at Pai Chai University. He received the B.S. degree in computer science, the M.S., Ph.D. degree in industrial engineering from Korea Advanced Institute